
User Experience in Cross-Cultural Contexts

Dissertation

an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität München

eingereicht von

FLORIAN LACHNER

München, den 31.10.2018

Erstgutachter: Prof. Dr. Andreas Butz
Zweitgutachter: Prof. Dr. Sarah Diefenbach
Drittgutachter: Prof. Ph.D. Coye Cheshire

Tag der mündlichen Prüfung: 01.02.2019

ABSTRACT

The field of Human-Computer Interaction (HCI) experienced manifold changes since it emerged at the end of the past century. While product functionality mainly defined the early days of HCI, the diffusion of graphical user interfaces naturally led to further usability considerations. Especially since the turn of the century, users increasingly expect technological products to provide a pleasant experience. The resulting sub-field of User Experience (UX) generally studies the consideration and evaluation of such experiences. Nowadays, various perspectives and skills from a variety of disciplines, such as psychology, computer science, design, or engineering, are required to properly investigate the field of UX due to the complexity of experiences and associated emotions or feelings. As a consequence, UX has evolved in a multi-disciplinary area of tension, driven by the influence of diverse epistemological standpoints and varying perceptions of the relevant terminology as well as applicable methods. Nevertheless, researchers from both academia and industry agree that many aspects need to be taken into consideration to holistically address UX, e.g., gender, age, personality, and culture. Against the background of the still ongoing globalization, the latter becomes more and more important for globally acting companies. In sum, two key issues that the field of UX currently needs to address is how culture influences methodological approaches in UX research & evaluation and how different disciplines shape the consensus on how UX is understood and measured in academia and industry. Both issues are addressed in this thesis. But how can we best measure UX? How can we consider cultural differences during the UX process? And how can cross-functional UX teams efficiently work together to holistically address UX? To bring us closer to answering these questions, this dissertation makes three key contributions:

First, this dissertation presents an extensive analysis of the current state of research of UX evaluation methods. This analysis serves as a basis for the subsequent parts of this dissertation. One hundred papers from eight established and prestigious HCI venues have been analyzed to identify common patterns in UX evaluation, particularly in regard to mixed methods approaches, i.e., method triangulation. The study reveals that the majority of UX studies in academic literature is based on a small set of established UX methods. In addition, it provides researchers with an overview of established evaluation approaches, currently less studied aspects in UX, and helps to guide future work in the field of UX methods, particularly regarding method triangulation.

Second, three case studies to study how culture can be considered during the UX process represent the core of this thesis based in the preceding analysis of UX evaluation methods. Each case study pursues a different design approach, i.e., a measuring approach based on log analysis, an empathic approach based on ethnographic interviews, a drawing task, and observation, as well as a pragmatist approach based on cultural theory. These types of methodological UX approaches were derived from literature and help to study cross-cultural aspects in UX from different perspectives. The ultimate goal was to identify which UX approach is suitable for different research objectives. As a synthesis of the three case studies, I introduce the term *culturally sensitive design* in this thesis to emphasize that my aim is not to blindly apply potentially biased design guidelines in cross-cultural contexts but to properly consider cultural particularities for UX design, i.e., to design with culture in mind. In addition, I present a framework that embeds the three UX approaches in a cultural context.

Third, based on the study of UX methods, I introduce two concepts for digital UX tools that help cross-functional teams to holistically measure and communicate UX and to maintain empathic relationships with users. As UX teams, nowadays, usually consist of people from diverse educational

and technical backgrounds, single UX methods can rarely cope with consequentially diverse research mindsets and requirements. The evaluation of the measuring-focused tool revealed that a quantified UX method provides a good overview for future research efforts and helps to benchmark a product's UX. The tool that focused on empathic relationships with users showcased that a user journey dashboard including a communication feature to interact with users can effectively support UX evaluation over time. In this context, both concepts serve as a starting point to further develop the method toolkit of cross-functional UX teams.

The contribution of this dissertation centers around the ongoing discussion about a common understanding of methodological approaches and a joint language at the intersection of cultural and UX research. More precisely, this dissertation contributes a state-of-the-art meta review on UX evaluation methods and two software tool artifacts for cross-functional UX teams. In addition, it provides empirical insights in the differing usage behaviors of a website plug-in of French, German and Italian users, website design preferences of Vietnamese and German users, and learnings from a field trip that focused on studying privacy and personalization in Mumbai, India. Finally, based on these empirical insights, this work introduces the concept *culturally sensitive design* that goes beyond traditional cross-cultural design considerations in HCI that do not compare different approaches to consider culturally sensitive product aspects in user research. I hope to inspire a more in-depth research agenda to eventually develop a joint conception of culturally sensitive design in the future.

ZUSAMMENFASSUNG

Das Forschungsfeld Mensch-Computer Interaktion hat sich seit seinem Ursprung gegen Ende des vergangenen Jahrhunderts stark verändert. Wohingegen zunächst Rechenleistung und Funktionalität die bestimmenden Faktoren in der technologischen Produktentwicklung waren, traten insbesondere mit der Verbreitung grafischer Benutzerschnittstellen verstärkt Fragestellungen aus dem Bereich der Bedienbarkeit, der sogenannten *Usability*, in den Vordergrund. Seit der Jahrhundertwende führten die stets steigenden Kundenanforderungen einer zunehmenden Erlebnisgesellschaft schließlich zur Entstehung des Forschungsfeldes *User Experience* (UX). Heutzutage beschäftigen sich viele verschiedene Disziplinen mit der Analyse des Nutzungserlebnisses - der UX - von technologischen Produkten, um der Komplexität von Erlebnissen und resultierenden Emotionen und Gefühlen gerecht zu werden. Das Forschungsfeld UX hat sich dadurch zu einem multidisziplinären Spannungsfeld entwickelt, dessen Eckpfeiler folglich aus verschiedenen wissenschaftlichen Auffassungen und dadurch unterschiedlicher Meinungen bezüglich relevanter Begrifflichkeiten und methodischen Vorgehensweisen basieren. Forscher aus der Wissenschaft und der Industrie sind sich unabhängig davon darüber einig, dass viele Aspekte bei einer ganzheitlichen Betrachtung von UX berücksichtigt werden müssen, z.B. Alter, Geschlecht, Persönlichkeit und Kultur. Vor dem Hintergrund einer weiterhin zunehmenden Globalisierung wird Letzterem vor allem von global agierenden Unternehmen mittlerweile mehr und mehr Bedeutung zugesprochen. Insgesamt steht das Forschungsfeld UX aktuell vor den zwei Herausforderungen zu verstehen, wie kulturelle Charakteristiken von Endnutzern das methodische Vorgehen im Bereich UX Forschung & Evaluation beeinflussen und wie funktionsübergreifende UX Teams effizienter zusammenarbeiten können. Beide Herausforderungen werden im Zuge dieser Dissertation adressiert. Doch wie kann UX am besten gemessen bzw. beurteilt werden? Wie können kulturelle Unterschiede und Gemeinsamkeiten bestmöglich im UX Prozess berücksichtigt werden? Und wie können funktionsübergreifende UX Teams effizient zusammenarbeiten, um Nutzungserlebnisse allumfassend zu analysieren? Das Ziel dieser Dissertations ist es, sich diesen Fragestellung zu widmen. Der wissenschaftliche Beitrag teilt sich dazu in drei Teile auf:

Im ersten Teil präsentiere ich eine umfangreiche Analyse des aktuellen Stands der Forschung im Bereich der UX Evaluationsmethoden. Dieser Überblick stellt die theoretische und methodische Grundlage der vorliegenden Arbeit dar. Die Analyse basiert auf der Auswertung von 100 Publikationen von acht renommierten Konferenzbänden bzw. Fachzeitschriften. Bei der Untersuchung wurde besonderer Wert auf die Erforschung von Methodekombinationen bzw. Methodentriangulierung gelegt. Es hat sich gezeigt, dass der Großteil wissenschaftlicher empirischer Studien auf lediglich ein paar wenigen Methoden beruht, insbesondere Fragebögen, Interviews und der Analyse von Logdaten. Die Studie dient hierbei als Überblick über den aktuell etablierten Methodeneinsatz, vernachlässigten Forschungsfragen im Bereich UX Methoden sowie als Anstoß und Inspiration zukünftiger Forschungsarbeiten, insbesondere hinsichtlich der Methodentriangulierung.

Im zweiten Teil stelle ich den Kern meiner Dissertation vor. Dieser ergibt sich aus drei Fallstudien, die eine geeignete Berücksichtigung von kulturellen Charakteristiken im UX Prozess diskutieren. Jede Fallstudie verfolgt dabei einen anderen Ansatz: Zunächst wird ein Messansatz basierend auf Logdaten mit dem Ziel, kulturelle Unterschiede "zu messen" beschrieben. Im Anschluss wird durch Anwendung von ethnographischen Interviews ein empathiefokussierter Ansatz dargestellt. Schließlich wird ein pragmatischer Ansatz verfolgt. Dieser beruht auf der Analyse und der Übersetzung der Kulturtheorie, genauer gesagt etablierten Kulturdimensionen. Die Aufteilung in drei verschiedene Ansätze wurde aus der Literatur abgeleitet und ermöglicht die Analyse von kulturellen Charakteristi-

ken aus verschiedenen Blickwinkeln. Das Ziel dieser Aufteilung ist es, zu verstehen, welcher Ansatz in bestimmten Situationen geeignet ist. Aus der Synthese dieser Fallstudien wird in dieser Dissertation der Begriff *kultursensitives Design* etabliert. Der Begriff betont, dass die Berücksichtigung von kulturellen Unterschieden und Gemeinsamkeiten nicht durch eine blinde Anwendung von statischen Designrichtlinien befriedigt werden kann, sondern kulturelle Besonderheiten vielmehr “bewusst gemacht” und Produkte - falls sinnvoll - entsprechend angepasst werden müssen. Zusätzlich wird in dieser Dissertation ein Framework erarbeitet, das die Anwendbarkeit der drei methodischen Vorgehensweisen in einen kulturellen Kontext setzt.

Im dritten Teil beschreibe ich zwei digitale UX Evaluationsapplikationen, welche funktionsübergreifende UX Teams dabei unterstützen UX ganzheitlich zu messen sowie zu kommunizieren bzw. empathische Beziehungen mit Nutzern aufrecht zu halten. Da sich UX Teams heutzutage in der Regel interdisziplinär zusammensetzen um der Komplexität von Nutzungserlebnissen gerecht zu werden, verändern sich die Anforderungen an bisher etablierte Evaluationsmethoden. Die Evaluation des ersten Tools, welches darauf abzielt UX ganzheitlich zu messen, hat gezeigt, dass der Ansatz hilfreich ist um einen Überblick zu schaffen und einen guten Vergleichsindex für Produktvergleiche darstellt. Das zweite Tool, das auf empathische Beziehungen mit Nutzern abzielt, wurde insbesondere für langfristige Studien als hilfreich angesehen. In dem noch wenig untersuchten Bereich der funktionsübergreifenden UX Methoden stellen diese beiden Konzepte einen ersten Schritt in Richtung einer allumfassenden und interdisziplinären UX Evaluation dar.

Mit dieser Dissertation möchte ich das Forschungsfeld UX mit einem besonderen Fokus auf UX Methoden und unter Berücksichtigung kultureller Fragestellungen weiterentwickeln. Der wissenschaftliche Beitrag gliedert sich hierbei in die bestehenden Diskussionen an der Schnittstelle eines gemeinsamen methodischen Verständnisses von UX sowie kultureller Unterschiede und Gemeinsamkeiten ein. Dieser Beitrag besteht insbesondere aus einer aktuellen Meta-Übersicht über die Anwendung von UX Methoden in der Wissenschaft sowie der Konzeptionierung zweier Software-Artefakte zur Erforschung von UX für funktionsübergreifende Teams. Zusätzlich werden die empirischen Ergebnisse aus den Untersuchungen des Nutzungsverhalten von französischen, deutschen und italienischen Nutzern eines Website Plug-ins, bezüglich Präferenzen im Webseitendesign in Vietnam und Deutschland sowie hinsichtlich der Wahrnehmung von Personalisierung in Mumbai, Indien, vorgestellt. Basierend auf den empirischen Ergebnissen wird schließlich der Begriff *kultursensitives Design* entwickelt, welcher über bisherige Ansätze an der Schnittstelle Kultur und UX hinausgeht. Aktuelle Vorgehensweisen vergleichen bisher nicht im Detail die Vor- und Nachteile verschiedene methodischer UX Forschungsansätze. Ich hoffe daher, mit dieser Arbeit weitere Forschungsarbeiten und eine aktuelle Forschungsagenda anzustoßen, um schließlich eine einheitliche Auffassung von kultursensitivem Design zu entwickeln.

ACKNOWLEDGEMENT

The completion of this dissertation has been a long journey that many people have accompanied and shaped along the way. To begin with, I want to thank Andreas Butz for supporting personal research interest at our department as well as any of my research ideas, even if some of them seemed to be a dead end at the beginning. Thank you for the perfect balance between guidance and room for exploration as well as between academic supervision and personal advice from the first until the last step of this journey. My thanks also go to my two reviewers Sarah Diefenbach and Coye Cheshire. Thank you for all your time, effort, and commitment to review my thesis.

Also, I would like to thank all supporters, friends, and companions from the Center for Digital Technology and Management (CDTM). Thank you Laura Bechthold, Gesa Biermann, Patrick Bilic, Patrick Christ, Michael Chromik, Aaron Defort, Maximilian Engelken, Philipp Hulm, Claudius Jablonka, Florian Korte, Till Kröger, Julia Mecheels, Kilian Moser, Philipp Nägelein, Stefan Nothelfer, Tom Schelo, and Stefanie Weniger for making the CDTM such a wonderful, cheerful, and inspiring place. Every one of you pushed me forward and helped me in so many situations, ranging from course onboarding and project partner acquisition over co-authoring papers and reviewing this thesis to overcoming personal challenges in general. Thank you for all your personal feedback, all the incredible kickoff and course completion parties, and for your support to train world class table soccer skills. As part of the CDTM community, I would also like to thank Uta Weber for her dedication to build up a home for our ideas. My thanks also go to Maximilian Wühr, Michael Fröhlich, Oskar Hargedahl, Konstantin Huneke, Sascha Ballweg, Felix Wolf, and Manuel Grossmann for the incredible time during my research stay abroad at the University of California, Berkeley. And not to forget: thanks to every single CDTM who crossed my path and taught me serenity while pushing my limits.

Further, I want to say thank you to my friends, co-authors, colleagues, professors, and companions from the LMU Media Informatics and Human-Computer Interaction Groups. Thank you Hanna Schneider, Malin Eiband, Daniel Buschek, Christian Mai, Mohamed Khamis, Sarah Theres Völkel, Ceenu George, Nađa Terzimehić, Renate Häuslschmid, Maria Fysarak, Sarah Aragon Bartsch, Christina Schneegaß, Axel Hösl, Daniel Ullrich, Bastian Pflöging, Michael Braun, Tobias Seitz, Henri Palleis, Emanuel von Zezschwitz, Sarah Tausch, Gesa Wiegand, Beat Rossmly, Yong Ma, Lukas Mecke, Sarah Prange, Sylvia Rothe, Heiko Drewes, Bernhard Slawik, Alexander Wiethoff, Florian Alt, Heinrich Hußmann, Rainer Fink, and Franziska Schwamb for welcoming me so warmly at our lab and helping me with all my bureaucratic, technical, and research struggles. I feel honoured for sharing my academic journey with such an ambitious yet positively thinking group. In this context, I also want to thank all my bachelor and master thesis as well as practical course students. I truly enjoyed every single meeting with all of you and learned a lot from every supervision. Thank you Florian Fincke, Markus Sterner, Klara-Maria Dersche, Bettina Eska, Ou Changkun, Mai-Anh-Ngyuen, Alexander Schenkel, Felix Koschmidder, Lisa Simon, Fabian Ramelsberger, Lakmal Padmakumara, Victoria Müller, Daniela Pauelscu, Maxim Karl, Regina Hasholzner, Danqing Liu, Mengyi Zhang, Matthias Müller, Charlotte Anlauf, Tim Ordenewitz, Kevin Probst, Jonas Minkler, and Sabine Kaupp.

Outside of the CDTM and LMU environment, I want to say thank you to you, Katharina Frison, Ingrid Pettersson, and Andreas Riener for all the skype calls, emails, and personal meetings to understand the real core of UX evaluation. In that regard, I also want to thank you, Michael Maretzke, for your advice and guidance to shape my UX research path.

Finally, I want to express my heartfelt thanks to my family. Thank you, Sebastian Lachner, for being the best brother, reviewer, and mentor in personal union that I can imagine. Thank you, Maria and Alois Lachner, for being the most supportive parents I could wish for. Without your caring, support, and motivation I could not have taken this long journey. I feel blessed and I am truly grateful for having you.

Last but not least, I want to thank you, Daisy Paiva, for being a living example for resilience, courage, and happiness in good as well as in bad times. Thank you for every laughter on this journey, thank you for your never-ending smile, and thank you for making my life complete. Thank you, my love, for always being there for me. *Muito obrigado por tudo, eu te amo!*

TABLE OF CONTENTS

1 Introduction	1
1.1 The Concept of User Experience	2
1.2 The Role of Culture in Human-Computer Interaction	4
1.3 User Experience in Cross-Functional Teams	6
1.4 Contributing Publications	8
1.5 Methodology	9
2 Publications	11
2.1 A Review of Method Application in UX Evaluation	12
2.2 Case Studies for Culturally Sensitive UX Design	13
2.2.1 Measuring Approach	13
2.2.2 Empathic Approach	14
2.2.3 Pragmatist Approach	15
2.2.4 Constraints and Limitations	16
2.3 Tools for Cross-Functional UX Teams	17
3 Conclusion	19
3.1 Reflection	20
3.2 Future Work	22
List of Figures	25
List of Tables	27
References	29
Appendix: Original Contributing Publications	A 1

Introduction

From now on, leading-edge companies - whether they sell to consumers or businesses - will find that the next competitive battleground lies in staging experiences.

J. Pine & J. G. Browne, Welcome to the Experience Economy, 1998

User Experience (UX), nowadays, is a deep-rooted concept in product development processes [41]. However, researchers and practitioners often still use the term as a buzzword without a deep understanding of its notion and scope [7, 32]. As a consequence, diverse disciplines study and describe various influences on as well as methods to evaluate users' experiences, ranging from demographic backgrounds to design and product features. In this context, this thesis investigates the status quo of UX methods, the role of culture in Human-Computer Interaction (HCI), and the resulting cross-functional collaboration of UX teams. As a main contribution, this thesis provides a comprehensive analysis on how to consider culture during the UX process. I start by mapping out my overall motivation and derived research questions before presenting and discussing the results of all contributing publications of this cumulative dissertation in the Chapters 2 and 3. Figure 1.1 illustrates the structure and relation between the sections in this thesis.

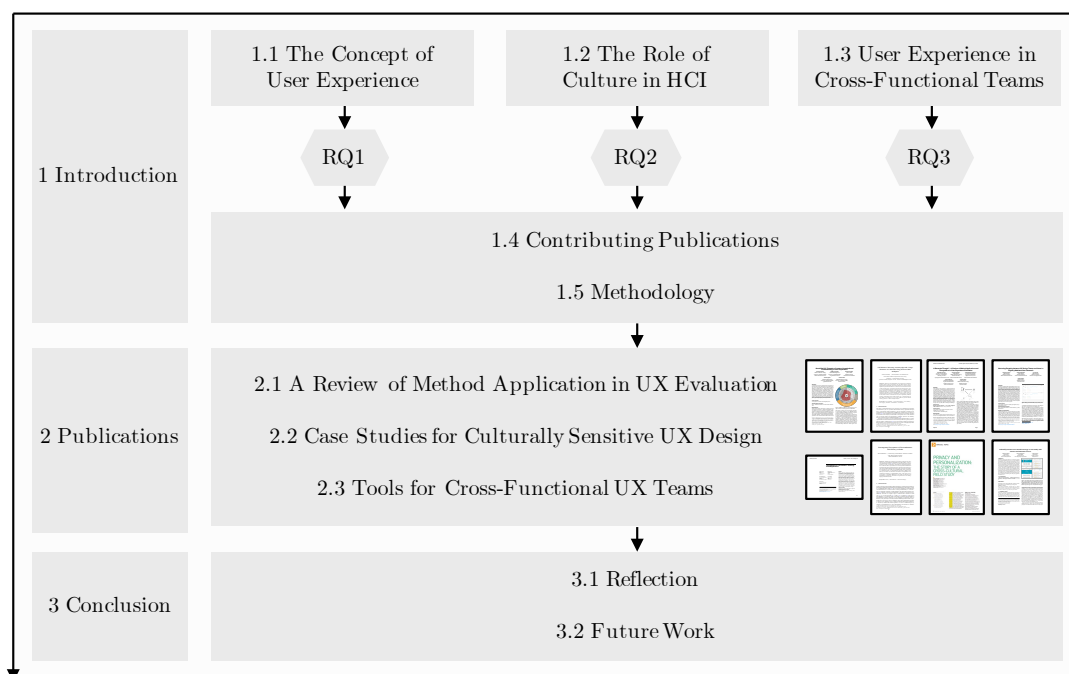


Figure 1.1: Overview and relation of the chapters in this thesis - including the relation of the respective research questions (RQ).

1.1 The Concept of User Experience

“Knowing your users” has been accepted as a guiding principle for the field of Human-Computer Interaction (HCI) since its early days in the 1990s [97]. Initial work in this context mainly focused on an established usability paradigm with task- and feature-related product considerations at the center of interest [36]. At this time, prevalent epistemologies and methodologies have shaped a HCI mindset that was predominantly based on the way of thinking “to measure is to know” [53]. However, around the turn of the millenium, research in the field of HCI has witnessed a substantial shift in the general perception of its theoretical principles and methodologies to “understand users” [51, 53]. The needs of the post-materialistic society in those days shifted from work- to experience-related factors [31, 50]. As a consequence, researchers and practitioners alike generally agreed that for a holistic and successful product development process both pragmatic and hedonic design characteristics beyond usability need to be considered to ensure commercial success in an “experience economy” [33, 51, 69]. The resulting concept of User Experience (UX), nowadays, still represents a core aspect of both academic and industrial HCI research [55, 57, 62, 67, 76].

Until today, however, researchers from academia and industry did not yet agree on a joint definition of UX and hence the term itself is often still used as a buzzword [2, 78]. In addition, since human experiences related to technological products and associated emotions are complex in nature [8, 17, 54, 80], the field of UX requires an interdisciplinary research and design approach leading to terminological and methodological influences from diverse disciplines (e.g., computer science, psychology, ergonomics, human factors, or engineering) [2, 14, 37]. Nowadays, the scope of UX continues to additionally consider further aspects and disciplines, such as branding and marketing in recent HCI literature [82, 85]. Current definitions of UX range from, e.g., “a person’s perceptions and responses that result from the use or anticipated use of a product, system or service” [1], “a momentary, primarily evaluative feeling (good-bad) while interacting with a product or service” [33], to “all aspects of the end-user’s interaction with the company, its services, and its products” [66].

The essential difference between traditional usability considerations and UX is the recognition that emotions, feelings, and thoughts and with that a user’s product appraisal can be formed and shaped before, during, and after an interaction with a product or service (see Figure 1.2) [3, 49, 52, 78]. In this context, Karapanos et al. [47] highlight that the development of UX follows the distinct temporal pattern of anticipation, orientation, incorporation, and identification. Along this cognitive process, researchers need to take diverse contextual aspects into account, ranging from stimulation

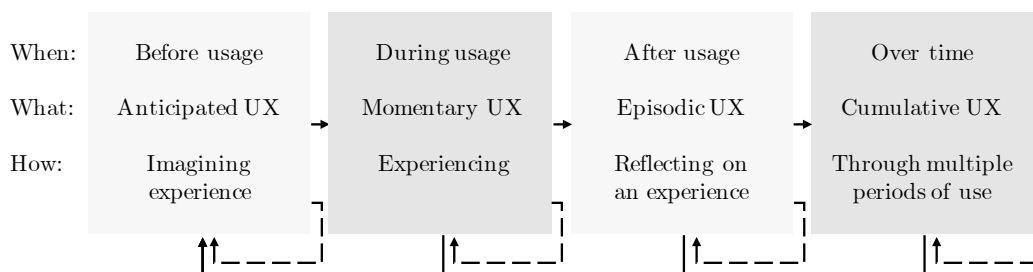


Figure 1.2: Temporal dimensions of UX illustrating the origin of experiences at different stages as well as the internal evaluation process and associated dependencies (adapted from Roto et al. [78]).

and learnability over usefulness and usability to personal reflection and social interaction [10, 24, 47, 58]. As a consequence, the field of UX yielded diverse methods and tools to consider and measure such diverse aspects and eventually established UX evaluation as vital part for human-centered design approaches [9]. However, there is still an ongoing debate about how to best describe and evaluate users' experiences [2].

One of the main discussions in UX evaluation that emerged over the last two decades is based on varying appreciations of qualitative and quantitative evaluation approaches [51, 67]. This discussion is led by divergent understandings of UX as a psychological needs-driven concept or a system-oriented concept focusing on task-fulfillment [9, 21, 83, 92]. Whereas psychological needs (see, e.g., Sheldon et al. [89]) and associated hedonic "be-goals" (i.e., why users want a product) focus on the creation of general product concepts and features from a macro perspective, a task-oriented concept based on "do-goals" (i.e., how users use a product) is rather suitable to identify product improvements, e.g., on a visceral level (see Norman [65]), and UX details from a micro perspective [33, 34, 35, 86]. As a consequence, different UX approaches are suitable for different contexts. Beyond the mere focus on evaluation, Battarbee & Koskinen [11] describe three general UX research and design approaches:

- **Measuring approaches:** Approaches that primarily focus on measurable aspects of UX based on suitable monitoring and feedback mechanisms.
- **Empathic approaches:** Approaches that are mainly based on considering emotional aspects of UX through interacting with users to "observe and feel the user".
- **Pragmatist approaches:** Approaches that are inspired by pragmatist philosophy and generally rely on insights from theoretical models.

In conclusion, the differing perceptions of UX evaluation approaches still represent a major challenge for current UX researchers and practitioners. Whereas Alves et al. [2] recently presented a study of UX evaluation practice, the in-depth analysis of the academic status quo of empirical UX studies dates back to the turn of the last decade. In that sense, the first objective of this dissertation is to shed light on the status quo of UX evaluation based on the following research question RQ1:

RQ1: *"How is user experience evaluated in empirical HCI studies?"*

Definition of UX within this thesis. It is important to note that the understanding of UX in this thesis is based on the concept of Hassenzahl and colleagues (see [33, 35, 36]). Although hedonic aspects can be primarily seen as a driver or motivator for experiences, pragmatic and hedonic aspects jointly form and influence both a user's purpose or need for and the actual interaction with a product [34]. Consequently, I set out to study both pragmatic and hedonic product aspects in the context of this thesis. In the theoretical analysis of the current state of research to address RQ1, however, no papers were excluded due to differing understandings or definitions of UX. In fact, the respective contributing publication [P6] considers the individual perspectives of the respective authors within the analyzed papers and hence all papers that mention a focus on UX have been considered in the analysis.

Contribution. This thesis contributes to an extensive state of research via a literature review on UX evaluation method application based on the analysis of empirical UX studies in 100 papers from eight renowned HCI venues. Due to the multi-dimensionality of UX, the review puts a particular emphasis on the investigation of method triangulation and the consideration of the different temporal stages for UX evaluation. The combined application of UX methods has not been addressed in previous reviews published around the turn of the decade, yet this is an important next step for a better methodological understanding of UX [9, 92]. As a consequence, the review serves as an extension of previous UX method overviews in academic UX studies to derive current gaps and future directions. In addition, the analysis represents a theoretical grounding for further contributing publications in this thesis, particularly with regards to the discussion about quantitative vs. qualitative UX approaches, one of the core questions of this dissertation related to cultural (see Section 1.2) and cross-functional considerations (see Section 1.3) as illustrated in Section 1.5.

1.2 The Role of Culture in Human-Computer Interaction

The role of culture has been a topic of interest for the field of HCI since a long time. With an ongoing globalization process and technological maturity of the web, the majorities of users in many countries have access to products and services from any country within just a few clicks. As a consequence, industries have started to adapt their products for different markets while researchers have investigated the theoretical background of cultural differences in HCI [71, 94]. The scope of culture in HCI, however, has changed over the course of the last two decades. Initially, when considering unique characteristics of international user groups, HCI researchers and practitioners primarily focused on the adaptation of visible product aspects, e.g., font, language, and further aspects for the development of adapted user interfaces (UI) [64]. In this context, Barber & Badre [6] derived the concept of culturability to emphasize cultural influences on traditional usability considerations. With the shift from a task-related usability perspective in HCI to a needs-driven UX perspective, cultural considerations became one of the key influencing factors on a product's UX [78].

Until now, the landscape of cultural considerations in HCI still lacks a common theoretical grounding, hence is characterized by controversial opinions and discussions [40, 73]. Established general definitions of culture range from "race and ethnicity [...] manifested in customary behaviours, [...] values, patterns of thinking and communication style" [18] over "a system of meaning that underlies routine and behaviour in everyday working life" [15] to "the collective programming of the mind that distinguishes the members of one group or category of people from another" [38]. The common ground of these perspectives in relation to HCI is the presence of both visible or conscious (e.g., language) and invisible or unconscious differences between cultures (e.g., values). Various meta models illustrate this breakdown of conscious and unconscious cultural aspects (e.g., the iceberg model from Hoft [40], see Figure 1.3, or the onion and pyramid model from Hofstede et al. [39]). In addition, Hofstede et al. [39] highlight that culture can be studied at different levels, e.g., at a national level (i.e., country of origin), regional or ethnic level, gender or generation level, social class, or at an organizational and corporate level.

Due to the lack of a common descriptive definition of culture in HCI or a generally accepted theoretical framework, two divergent streams of literature have emerged in the field of HCI within the recent years [95]. On the one hand, a cultural theory focused perspective primarily based on cultural dimension (see, e.g., Ford & Kotzé [23], Aaron & Gould [56], or Reinecke et al. [72]). On the other

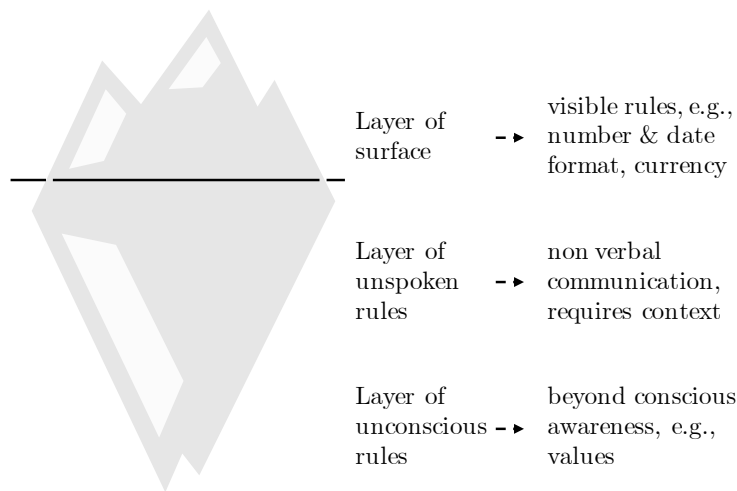


Figure 1.3: Different metaphorical layers of culture according to the iceberg model whereof a large part of cultural characteristics are not directly observable (adapted from Hoft [40] and Galletta & Zhang [26]).

hand, a perspective that aims to give users a greater agency during the research process through a distinct focus on an interaction at eye level between users and researchers in the age of “postcolonial computing”, particularly with regards to emerging countries (see Irani et al. [43] and Irani & Dourish [42]). In contrast to the latter, the first, i.e., cultural theory focused perspective, aims to derive cultural insights or justification directly from cultural dimensions. In general, cultural dimensions are personal traits that characterize a culture’s general preference regarding common societal problems [38]. Geert Hofstede and colleagues [39], Shalom Schwartz [88], Trompenaars & Hampden-Turner [90], Edward Hall [29, 27, 28], Kluckhohn & Strodtbeck [48], and David Victor [93] present currently established cultural dimensions.

Until today, research in the field of UX in relation to cultural considerations still rather explores suitable approaches than establishing common cross-cultural research patterns. As a consequence, the investigation and comparison of different UX approaches in cross-cultural contexts represents a core part in the context of this dissertation. Thus, the second part of this thesis aims to address RQ2:

RQ2: “How can we consider cultural characteristics in UX research?”

Definition of culture within this thesis. The overall understanding of culture in this thesis is based on the iceberg model of Hoft [40] as illustrated in Figure 1.3 and the concept of national cultures. Narrowing down the theoretical framework of culture in this thesis to the iceberg metamodel and not a textual description was helpful to consider both pragmatic and hedonic aspects of UX and to allow for an independent application of all three case study approaches, i.e., a measuring, an empathic, and a pragmatist approach at a suitable theoretical level. Consequently, it was possible to deal with established cultural theory as well as the opposing concept of postcolonial computing. However, I acknowledge that the focus on the concept of national culture comes along with inherent limitations and challenges to potentially neglect further influences, e.g., from sub-cultures or preferences of different generations. The decision to study national cultures, however, emerged from a detailed analysis of current HCI work in relation to cross-cultural consideration and the recognition that country borders represent a relevant topic of interest in both academic UX research and industrial practice.

Contribution. Over the course of my doctoral research, I conducted three case studies to investigate the applicability of different UX approaches in cross-cultural contexts. Based on the results of addressing RQ1, the case studies broaden the scope of this dissertation and juxtapose different methods for UX research. Thus, it is possible to consider specific study setting requirements of the applied methods to investigate suitable cultural layers. Each case study applies a different UX research approach yet pursues the same goal, i.e., identify cultural differences as part of a UX-focused user research process, to eventually draw a conclusion on benefits and drawbacks of different strategies. I base my methodological foundation, in this context, on the work of Battarbee & Koskinen [11] that distinguish between a *measuring approach*, an *empathic approach*, and a *pragmatist approach* as overall concepts in the field of UX research and design. This breakdown aligns currently established research streams in the field of cross-cultural UX and the insights from [P6] that addresses RQ1 .

The first case study in [P1] follows a measuring approach and investigates the development of UX metrics based on the analysis of log data from an e-commerce website plug-in. Whereas Battarbee & Koskinen [11] primarily mentioned the measurement of sensations, e.g., through psychophysiological user feedback, this study is based on log data as the use of large-scale behavioral data represents an up-to-date topic in quantitative UX evaluation (see, e.g., Rodden et al. [77]). The second case study in [P7] and [P8] focuses on an empathic approach and presents an ethnographic interview study in Mumbai, India based on the idea that “designers must both observe and feel for the users” [11, 46]. The third case study in [P3] investigates how cultural theory, in particular cultural dimensions, can be applied to consider cultural characteristics for UX research. As part of this study, I substantiate the concept of *culturally sensitive design* (see [P3], page 2). Similar to the idea of value-sensitive design according to Friedman [25], I establish the term culturally sensitive design to emphasize that the core idea of my research is to design with culture in mind and that the consideration of cultural characteristics requires different research approaches in different contexts. The overarching contribution of the three case studies is twofold. First, each case study presents distinct user insights derived from the study itself. The insights can be used as starting points for further domain specific case studies in similar contexts. Second, the different UX approaches allow me to juxtapose different methodologies to eventually derive the concept of culturally sensitive design as part of this thesis in Chapter 3.

1.3 User Experience in Cross-Functional Teams

The historical evolution of the field HCI is commonly seen as a gradual process that was pushed forward in three waves. The first wave describes the initial effort of experts from the fields of cognitive science and ergonomics in the second half of the 20th century to rather define product characteristics based on user behavior and not the other way round. With the advent of personal computing, the user (or human) perspective became increasingly important. The second wave emerged from the realization that technical systems actually represent a collaborative system of several users. Bannon [4] describes this step between the 1990s and 2000s as the move from “human factors to human actors”. Since the beginning of this century, the third wave eventually embraced experiences as well as meaning-making and hence contextual factors became more important [16, 30, 63]. Today, more and more voices raise the question if we are currently experiencing the development of a fourth wave forged ahead by more and more disciplines entering the stage of HCI as well as a more distinct focus on ethical values, creativity, well-being, and self-realization [5, 13, 16, 74].

The shift from human-centered design question as an isolated responsibility to an interdisciplinary discourse can also be seen in practice. Whereas the conventional product development process used to be a stage-gate process of all involved departments, the disciplinary borders of UX teams have blurred steadily to address the complexity of current design questions in the 21st century. Traditionally, human factor engineers or user researchers identified user requirements and derived design objectives. Next, computer scientists and designers created prototypes and conducted user tests to develop an optimized version of the product. Finally, marketing experts translated the branding strategy and product characteristics into an appealing sales slogan [45, 91]. Nowadays, companies see product development and particularly the consideration of UX more as a company-wide topic of interest. At the same time, research increasingly investigates the relation between an established design-focused UX perspective and an originally marketing-focused brand experience perspective [79, 81, 84]. As a consequence, current UX teams are typically staffed with experts from different disciplines [44]. To properly address this change, the field of HCI should not only be seen as an individual research area but as an interdisciplinary research field or an “inter-discipline” [12, 19, 70].

Nevertheless, established UX methods and tools rarely cope with the interdisciplinary and diverse needs of all stakeholders within the UX process [2]. Whereas design-focused stakeholders often ask for qualitative insights and methods based on a strong end-user involvement to derive concrete product decisions, business-focused stakeholders generally require quantified feedback to validate business decisions [53, 91]. However, although design education, nowadays, is already seen as an interdisciplinary endeavour [13], the field of HCI and associated methods still reside in a “chaos of multiplicity” [16]. Yet, as an effective and efficient cross-functional collaboration between all UX stakeholders is necessary in the context of cross-cultural considerations to ensure customer loyalty and global product success, the third contribution of this dissertation results from the analysis of RQ3:

RQ3: *“How can digital tools support cross-functional UX teams?”*

Contribution. With the guidance of RQ3, this thesis presents two concepts for digital UX evaluation applications for cross-functional teams. Both concepts have been developed along the methodological journey of this dissertation as illustrated in Section 1.5. To begin with, together with my co-authors of the respective publication [P2], I developed a UX measurement tool that supports cross-functional UX teams to measure, visualize, and eventually communicate UX feedback. In general, the tool aims to break down the concept of UX to a quantifiable level that we called Quantified UX (QUX). This tool was developed at the beginning of this dissertation with a general focus on measuring approaches in UX. Next, based on further insights and results in the meantime, the second tool from [P4] focuses on establishing and maintaining digitally-mediated long-term relationships between UX teams and users. The concept of this tool is based on the idea of empathic UX approaches that strongly influenced the second phase of this dissertation. The overall contribution of both associated publications [P2] and [P4] can be summarized at a methodological and at an artifact level (see also Chapter 3). First, both tools can be integrated in the established method toolkit of cross-functional UX teams. On the one hand, a tool that serves as starting point and guidance for further in-depth evaluations as well as a quantified communication tool and, on the other hand, a tool that supports empathic relationships between UX teams and users. Second, the respective publications summarize the development of the concepts including a requirement analysis and definition phase, a prototyping and concept phase, and eventually a test and feedback phase. In addition to the final results also the interim results can be used for future work and help to develop further suitable UX methods.

1.4 Contributing Publications

The results of this cumulative dissertation have been published in individual publications before. This thesis, therefore, serves as a summary of all projects. The underlying publications are listed in alphabetical order in the separate reference list below. Citations of these publications are marked with a “P” (e.g., [P3]). The original publications are fully attached in *Appendix: Original Publications*. [P6] received an *Honourable Mention Award* at the respective conference (CHI ’18).

- [P1] Florian Lachner, Florian Fincke, and Andreas Butz. “UX Metrics: Deriving Country-Specific Usage Patterns of a Website Plug-In from Web Analytics”. In: *16th IFIP TC 13 International Conference on Human-Computer Interaction*. INTERACT 2017. Springer-Verlag, 2017, pp. 142–159. DOI: 10.1007/978-3-319-67687-6_11 (cited on pp. 6, 9, 11, 13, 15, 16, 19–22, A 1).
- [P2] Florian Lachner, Philipp Naegelein, Robert Kowalski, Martin Spann, and Andreas Butz. “Quantified UX: Towards a Common Organizational Understanding of User Experience”. In: *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. NordiCHI ’16. ACM, 2016, 56:1–56:10. DOI: 10.1145/2971485.2971501 (cited on pp. 7, 9, 11, 17, 19, 22, 23, A 1).
- [P3] Florian Lachner, Mai-Anh Nguyen, and Andreas Butz. “Culturally Sensitive User Interface Design: A Case Study with German and Vietnamese Users”. In: *Proceedings of the 2nd African Conference on Human Computer Interaction*. AfriCHI ’18. ACM, 2018. DOI: 10.1145/3283458.3283459 (cited on pp. 6, 8, 9, 11, 13, 15–17, 19, 21, 22, A 1).
- [P4] Florian Lachner, Hanna Schneider, Lisa Simon, and Andreas Butz. “Nurturing Empathy between UX Design Teams and Users in Digitally-Mediated User Research”. In: *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*. NordiCHI ’18. ACM, 2018. DOI: 10.1145/3240167.3240182 (cited on pp. 7, 9, 11, 17, 19, 22, A 1).
- [P5] Ingrid Pettersson, Anna-Katharina Frison, Florian Lachner, Andreas Riener, and Jesper Nohage. “Triangulation in UX Studies: Learning from Experience”. In: *Proceedings of the 2017 ACM Conference Companion Publication on Designing Interactive Systems*. DIS ’17 Companion. ACM, 2017, pp. 341–344. DOI: 10.1145/3064857.3064858 (cited on pp. 9, 11, 12, 19, A 1).
- [P6] Ingrid Pettersson, Florian Lachner, Anna-Katharina Frison, Andreas Riener, and Andreas Butz. “A Bermuda Triangle? - A Review of Method Application and Triangulation in User Experience Evaluation”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. ACM, 2018, 461:1–461:16. DOI: 10.1145/3173574.3174035 (cited on pp. 3, 6, 8, 9, 11, 12, 19, 20, A 1).
- [P7] Hanna Schneider, Ceenu George, Malin Eiband, and Florian Lachner. “Investigating Perceptions of Personalization and Privacy in India”. In: *16th IFIP TC 13 International Conference on Human-Computer Interaction*. INTERACT 2017. Springer-Verlag, 2017, pp. 488–491. DOI: 10.1007/978-3-319-68059-0_57 (cited on pp. 6, 11, 13, 14, 16, 19–22, A 1).
- [P8] Hanna Schneider, Florian Lachner, Malin Eiband, Ceenu George, Purvish Shah, Chinmay Parab, Anjali Kukreja, Heinrich Hussmann, and Andreas Butz. “Privacy and Personalization: The Story of a Cross-cultural Field Study”. In: *Interactions* 25.3 (2018), pp. 52–55. DOI: 10.1145/3197571 (cited on pp. 6, 9, 11, 13, 14, 16, 19–22, A 1).

1.5 Methodology

The objective of this thesis is to study the applicability of different UX methods in the context of cross-cultural design. To better understand the relation and origin of the different contributing publications (see Section 1.4) as well as corresponding methods, I proceed with an overview of the applied methods and the underlying structural framework.

Applied Research Methods

To study the applicability of different UX methods this cumulative dissertation represents a methodological journey itself. In addition to the reflection on method application in Section 2.2.4, I clarify the reasonings for all applied methods in the originally published publications below.

- **Literature analyses** were generally applied in order to understand relevant theory as well as UX method requirements, e.g., in [P2]. In addition, [P6] represents an extensive structured literature review adapted from the QUOROM statement (see Moher et al. [60]) to address RQ1.
- **Semi-structured interviews** served as method in early design stages for design exploration in [P8] and concept definition in [P2], or for expert feedback in [P2] and group evaluation in [P4].
- Together with my co-authors, I developed a **questionnaire** based UX measurement tool in [P2] as questionnaires represent one of the currently most applied UX methods [9].
- [P1] presents a **log analysis** that was used to study the applicability of quantitative UX metrics based on large-scale behavioral data.
- **Workshops** were organized to discuss initial study results with further experts in [P5] and to gather and rate design options for the UX tool in [P4].
- To better understand and compare the mental model of people from a collectivist and an individualistic country, we defined and used a **drawing task** in [P8], asking our study participants to sketch the flow of information for their online behavior.
- **Observation** was applied in [P8] to immerse in the foreign environment and to identify contextual factors related to the respective research question.
- The **think-aloud technique** was used in [P3] to understand emotions and feelings while the study participants interacted with the prototypes in the respective study.
- In [P4], my co-authors and I conducted a **focus group** study to investigate benefits and drawbacks of the collaborative UX tool concept.

Structure of the Thesis

This dissertation is the result of a comprehensive analysis of UX methods and cross-cultural considerations in UX design. Preliminary results, obstacles, and industry partner collaborations in different stages of this overall work influenced the origin and completion of the publications. As a

consequence, the structural order of this thesis differs from the chronological order of the individual publications. In general, **Chapter 1** introduces the theoretical background related to the fields of UX and cultural research as well as cross-functional collaboration in UX. In addition, it presents the research questions that this dissertation addresses. **Chapter 2** summarizes the publications included in this thesis, highlights their individual contributions, and clarifies how the authors of each publication contributed to the respective final manuscript. Figure 1.4 illustrates the structural framework of the publications that jointly form this cumulative thesis. **Chapter 3** concludes this thesis and presents an outlook for future work.

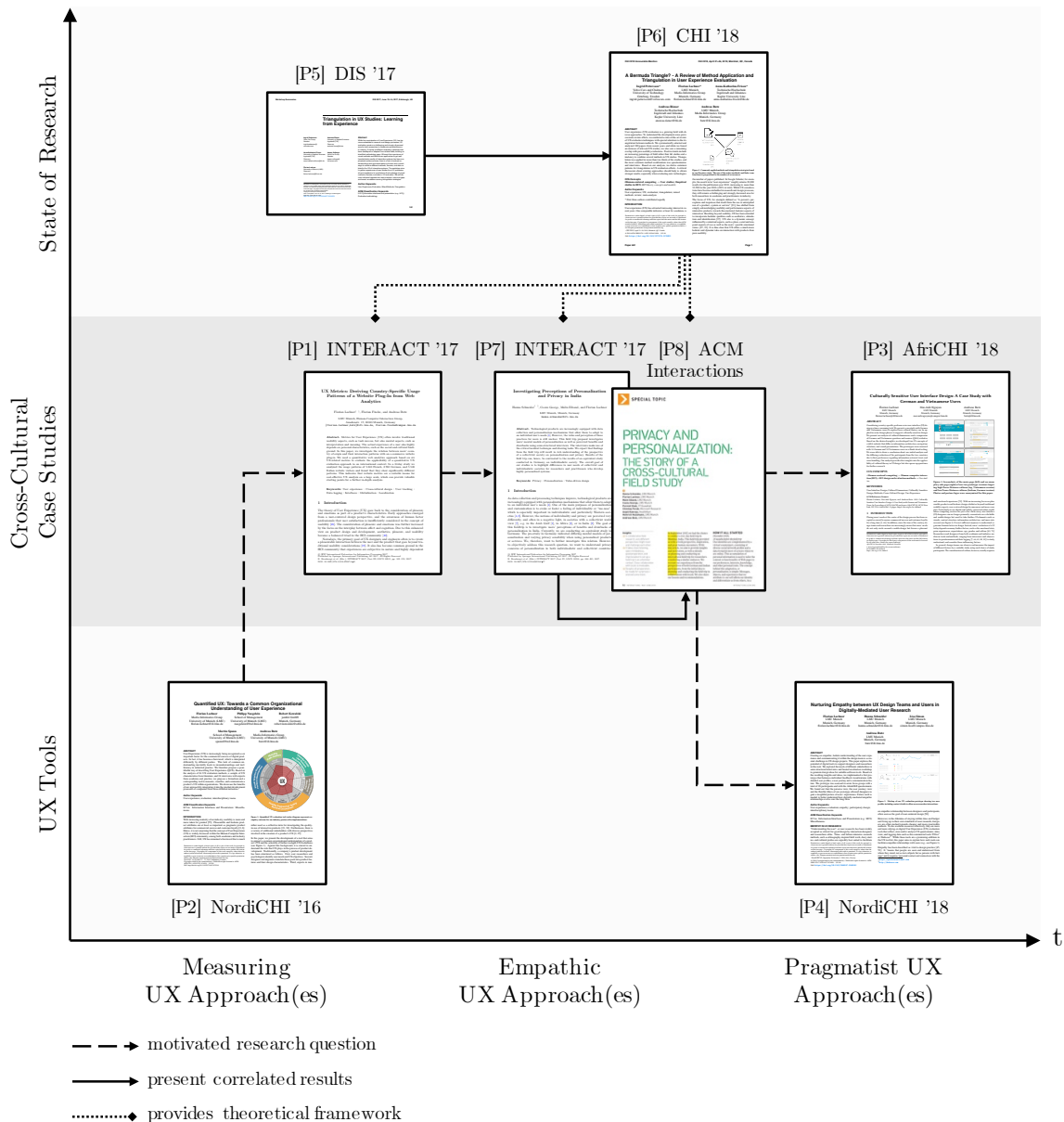


Figure 1.4: Chronological correlation (x-axis) and thematic coherence (y-axis) of this dissertation including the structural relationships of the original contributing publications.

2

Publications

Increasing technological sophistication in many countries of the world and the resulting larger world trade implies greater need to pay attention to international aspects.

J. Nielsen, Designing User Interfaces for International Use, 1990

After introducing the research questions, theoretical basis, and structure of this thesis in the previous Chapter, I proceed by outlining the individual results and contributions of all publications of this dissertation from section 1.4. Therefore, accompanied by a preview of the first page, I summarize all publications and explain the particular contributions beyond the abstracts of the original papers. As all publications were the result of collaborative projects with colleagues, students and/or my supervisor, I refer to the scientific “We” throughout the sections of this chapter. In addition, I juxtapose my personal contribution to the publication of the original manuscripts and the contributions of others. Table 2.1 presents an overview and summarizes the outcomes of all publications to provide the reader with a better guidance for further in-depth studies.

RQ	Section	Publication Title and Publishing Venue	Type	Method(s)	Primary Outcome
RQ1	2.1	[P6] “A Bermuda Triangle? - A Review of Method Application and Triangulation in User Experience Evaluation” in <i>CHI '18</i>	Full Paper (16 pages)	Literature Review (N=100)	Analysis of the current state of research of UX evaluation methods
		[P5] “Triangulation in UX Studies: Learning from Experience” in <i>DIS '17</i>	Workshop (4 pages)	Theoretical Analysis	Workshop procedure, methods, and goals
RQ2	2.2	[P1] “UX Metrics: Deriving Country-Specific Usage Patterns of a Website Plug-In from Web Analytics” in <i>INTERACT '17</i>	Full Paper (18 pages)	34 days Log Analysis (N=14,151)	Development of UX metrics for large-scaled website tracking
		[P7] “Investigating Perceptions of Personalization and Privacy in India” in <i>INTERACT '17</i>	Field Trip (4 pages)	Theoretical Analysis	Field Trip procedure, methods, and logistics
		[P8] “Privacy and Personalization: The Story of a Cross-Cultural Field Study” in <i>ACM INTERACTIONS</i>	Magazine Article (4 pages)	Interviews, Drawing Task, Observation (N=16)	Learnings for ethnographic field research in India
		[P3] “Culturally Sensitive User Interface Design: A Case Study with German and Vietnamese Users” in <i>AfriCHI '18</i>	Full Paper (12 pages)	Prototype Design and Evaluation (N=28)	Insights into the applicability of cultural theory in UI design
RQ3	2.3	[P2] “Quantified UX: Towards a Common Organizational Understanding of User Experience” in <i>NordiCHI '16</i>	Full Paper (10 pages)	Interviews (N=24), Concept Evaluation (N=3)	Development of QUX, a UX measurement tool for interdisciplinary teams
		[P4] “Nurturing Empathy between UX Design Teams and Users in Digitally-Mediated User Research” in <i>NordiCHI '18</i>	Full Paper (13 pages)	Prototype Design and Evaluation (N=26)	Design opportunities for empathic relationships with users through digital tools

Table 2.1: Overview of the publications as well as their methods and outcomes.

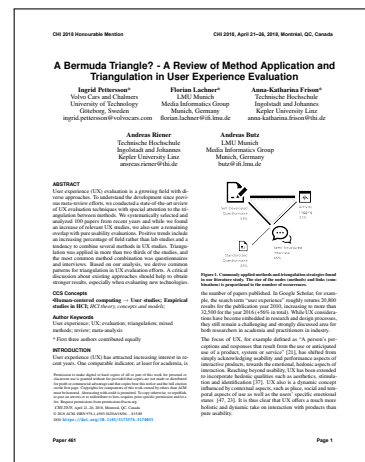
2.1 A Review of Method Application in UX Evaluation

[P6] and [P5] represent the first publications of this dissertation and serve as an analysis of related work regarding UX methods. Whereas [P6] introduces a full overview of established UX evaluation methods, [P5] reports an associated workshop where initial results of [P6] have been jointly discussed with further researchers for the final analysis. Altogether, [P5] and [P6] address RQ1:

RQ1: “How is user experience evaluated in empirical HCI studies?”

[P6] A Bermuda Triangle? - A Review of Method Application and Triangulation in User Experience Evaluation

Summary: This paper presents a state-of-the-art analysis of UX evaluation methods in academic publications. We conducted a systematic literature review with 100 full papers from well-known HCI conferences and journals. In this analysis, we specifically focused on the analysis of UX method combination, i.e., method triangulation. The motivation for this paper emerged from a reflection of previous work in the field of UX methods that studies the application of UX methods in general, but does not discuss the reasonings for explicit and implicit method triangulation. The qualitative review revealed eight patterns for UX method triangulation. As the idea of and the collaboration for this paper emerged around the half-way point of this dissertation, the chronologically first papers in this document do not refer to the results of this paper. However, this work guided the overall topic of this dissertation based on the discussion about quantitative and qualitative data in UX.

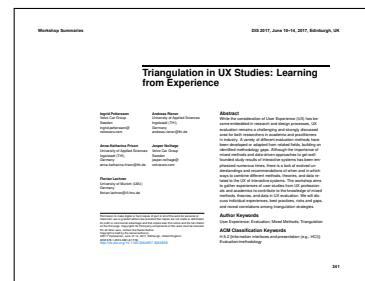


Author contributions: Ingrid Pettersson, Anna-Katharina Frison and I contributed equally to this publication. We came up with the research strategy and jointly conducted the whole literature review. I mainly contributed to the identification of the triangulation patterns. Andreas Riener and Andreas Butz supervised the project and revised the paper for conciseness and readability.

[P5] Triangulation in UX Studies: Learning from Experience

Summary: This publication represents a workshop summary that was developed to discuss initial results of [P6] with experts from both academia and industry. In addition, we wanted to get further insights about the motivation for mixed methods and data-driven approaches in UX from various UX professionals.

Author contributions: I contributed to the concept development and the realization of the workshop. Ingrid Pettersson and Anna-Katharina Frison came up with the idea and were the leading authors of the publication. Jesper Nollhage and Andreas Riener revised the paper for conciseness and readability.



2.2 Case Studies for Culturally Sensitive UX Design

To triangulate different culturally sensitive UX approaches for user research, this dissertation is based on the differentiation by Battarbee & Koskinen [11]. In particular, the next sections present three case studies following a *measuring approach* based on [P1] (Section 2.2.1), an *empathic approach* based on [P7] and [P8] (Section 2.2.2), and a *pragmatist approach* based on [P3] (Section 2.2.3). All publications [P1, P7, P8, P3] jointly address RQ2:

RQ2: “How can we consider cultural characteristics in UX research?”

As part of this dissertation, the focus of each case study lies on the applied method(s). As a consequence, the fields of application differ for each case study. In addition, due to inherent constraints and challenges of cultural research, each case study focuses on the analysis of user feedback from different cultural regions (see also Section 2.2.4). Against the background of the results from RQ1 and current efforts in cross-cultural UX, the overall goal is to derive a better understanding how different approaches can be best applied for UX research in cross-cultural contexts.

2.2.1 Measuring Approach

Originally, Battarbee & Koskinen [11] describe measuring approaches in UX as methodologies that focus on quantifying emotional responses. For this thesis, I want to expand this concept using large-scale behavioral data to measure UX. [P1] demonstrates how metrics based on large-scale behavioral data of a website plug-in can support a culturally sensitive design process.

[P1] UX Metrics: Deriving Country-Specific Usage Patterns of a Website Plug-In from Web Analytics

Summary: In this paper, we pursued a quantitative web analytics approach to investigate how pre-defined UX metrics based on log data can help to derive and communicate country-specific usage patterns. The project was motivated as existing literature argues that quantitative behavioral data represents a cost-effective strategy for UX evaluation in addition to established qualitative UX approaches. In collaboration with a partner company, we implemented a plug-in in an e-commerce website that allowed us to track click data from users from different countries. During a 34 day study, we collected and analyzed usage data of 5.843 French, 2.760 German, and 5.548 Italian users. In our study, we realized that our UX metrics were helpful to derive marginal yet significantly different usage patterns of users from the investigated countries. We hope to see more insights in the field of quantitative UX evaluation in the future since computing power will allow for an increasingly better analysis of large-scale behavioral data.

Author contributions: I came up with the idea of UX metrics and supervised the technical implementation and data collection. Further, I was responsible for the data analysis and was the leading

UX Metrics: Deriving Country-Specific Usage Patterns of a Website Plug-In from Web Analytics

Florian Luchner¹, Florian Finckle, and Andreas Böttz

LMU Munich, Human-Computer Interaction Group,
Amalienstr. 17, 80333 Munich, Germany
{florian.luchner,bottz}@ifi.lmu.de, florian.finckle@campus.lmu.de

Abstract. Metrics for User Experience (UX) often involve traditional usability aspects, such as task success, but also mental aspects, such as interpretation and meaning. The actual experience of a user also highly depends on personal characteristics, such as the social and cultural background. In this paper, we investigate the relation between users' country of origin and their interaction patterns with an e-commerce website plug-in. We used a quantitative web analytics approach based on six UX-related metrics to evaluate the applicability of a quantitative UX evaluation approach in an international context. In a 34 day study, we analyzed the usage patterns of 5.843 French, 2.760 German, and 5.548 Italian website visitors and found that they show significantly different patterns. This indicates that website metrics are a suitable means for cost-effective UX analysis on a large scale, which can provide valuable starting points for a further in-depth analysis.

Keywords: User experience · Cross-cultural design · User tracking · Data logging · Interfaces · Globalization · Localization

1 Introduction

The theory of User Experience (UX) goes back to the consideration of pleasure and emotions as part of a product's characteristics. Early approaches emerged from a user-centered design perspective, and the emergence of human factors professionals that user satisfaction is insufficiently considered in the concept of usability [9]. The consideration of pleasure and emotions was further enhanced by the focus on the interplay between affect and cognition. Due to this enhanced view on product design and development, aesthetics, pleasure, and usability became a balanced triad in the HCI community [16].

Nonetheless, the primary goal of UX designers and engineers often is to create a pleasurable interaction between the user and the product that goes beyond traditional usability considerations [19]. It also has become common ground in the HCI community that experiences are subjective in nature and highly dependent

© IFIP International Federation for Information Processing 2017.
Published by Springer International Publishing AG 2017. All Rights Reserved.
B. Böttz et al. (Eds.), INTERACT 2017, Part III, LNCS 10116, pp. 134–144, 2017.
DOI: 10.1007/978-3-319-64674-2_11

author of the publication. Florian Fincke acquired the industry partner, implemented the website plug-in and was responsible for the data collection and data preparation. Andreas Butz supervised the project and revised the paper for conciseness and readability.

2.2.2 Empathic Approach

Two publications belonging together pursue an empathic approach and primarily focus on qualitative ethnographic methods. [P7] describes the concept and setup of a field trip in Mumbai, India, to study a distinct sub-concept of UX, i.e., personalization and privacy. [P8] presents the results and insights of the conducted field trip as well as a general reflection. This approach comes closest to the aforementioned concept of postcolonial computing according to Irani et al. [43] in Chapter 1.2.

[P7] Investigating Perceptions of Personalization and Privacy in India

Summary: With increasing possibilities to collect and analyze user data while interacting with a digital product, researchers and practitioners from academia and industry need to investigate both benefits and drawbacks of potential product optimizations that such continuous data tracking enables. Whereas, on the one hand, the analysis of usage data allows to develop individual personalization mechanisms to ultimately increase a product’s UX, constant data tracking mechanisms, on the other hand, can raise critical privacy issues. In this paper, we outline the motivation and project plan including logistics and methods, namely semi-structured interviews, critical incident technique, and a drawing task, for a field trip in Mumbai, India, to investigate the perception of privacy and personalization in a collectivist country. The field trip proposal was motivated through a previously conducted similar study in Germany, an individualistic country, and eventually conducted as part of the conference INTERACT 2017. Field trips represented a new opportunity as part of the venue’s call for participation. The conference organizers supported all field trip hosts in logistics and the acquisition of participants for the interview study. We saw this as a great opportunity to conduct ethnographic on-site research in an unknown environment yet ensuring ethical conduct.

Author contributions: I contributed to the theoretical background related to cultural theory and implemented suggested changes from the editors for the final version. In addition, I contributed to the advertisement and call for participation of the field trip. Hanna Schneider came up with the research idea and led the authoring of the publication. Ceenu George and Malin Eiband provided feedback and supported writing the manuscript.

[P8] Privacy and Personalization: The Story of a Cross-cultural Field Study

Summary: This article summarizes the results and experiences of the previously described field trip in Mumbai, India, in September 2017. The overall goal of this publication is to share our insights and experiences about the implementation of the field trip and our general results. We pursued a storytelling approach to better explain the different stages of the field trip planning and implementation.



In total, ten researchers and practitioners from Germany, India, the USA, and the UK joined the two-day field trip. We have split the group in smaller teams and conducted 16 interviews altogether. We found out that the Indian participants had fewer concerns to provide personal data than German participants in an earlier study. Besides insights about the perception of benefits and drawbacks related to privacy and personalization in India, we made enriching experiences outside of our comfort zone and well-known research environment with a culturally diverse research team.

Author contributions: Together with Hanna Schneider, I planned, organized, and conducted the field trip. In addition, I contributed to the theoretical background of the publication. Hanna Schneider contributed significantly to the implementation of the field trip and was the leading author of the article. All field trip participants supported the data collection, the remaining authors contributed their perspectives to the resulting publication.

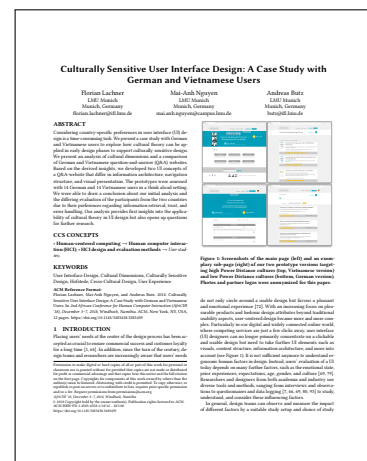
2.2.3 Pragmatist Approach

One publication follows a pragmatist approach, deriving cultural design characteristics primarily from cultural theory. [P3] explores how cultural theory can be used to identify elements that are critical for cross-cultural research. In this chronologically last publication, the term *culturally sensitive design* was put in concrete terms (see [P3], page 2), although I initially used it in [P1] on page 14.

[P3] Culturally Sensitive User Interface Design: A Case Study with German and Vietnamese Users

Summary: Considering cultural differences in design is a time consuming task. As a consequence, more and more researchers rely on cultural theory to anticipate and explain cultural differences in design. In this paper, we present a case study with German and Vietnamese participants that provided feedback for two question-and-answer website prototypes. The differing design elements were derived from an analysis of cultural dimensions according to Hofstede [39]. Based on our analysis of cultural dimensions we were able to anticipate differing preferences of German and Vietnamese participants regarding information retrieval, trust, and error handling. Finally, we reflect about our approach and discuss general critique and challenges of applying cultural dimensions. Our overall goal was to reveal insights into the applicability of cultural dimensions in UX design.

Author contributions: I came up with the research idea, acquired the collaboration partner, supervised the technical implementation, contributed to the user study and was the leading author of the publication. Mai-Anh Ngyuen significantly contributed to the user study and implemented the prototypes. Andreas Butz supervised the project and revised the paper for conciseness and readability.



2.2.4 Constraints and Limitations

At the core of this dissertation, my co-authors and I showcased the applicability of different approaches for culturally sensitive design in the previous sections. We took several decisions to properly address RQ2 from diverse perspectives, yet keep the study complexity across the different studies at a manageable level. However, every decision comes at the cost of neglecting differing conditions for the moment and possibly further research insights.

Methods: For each case study, we have selected a distinct set of methods, namely log analysis for the measuring approach [P1], semi-structured interviews, a drawing task, and observation for the empathic approach [P7, P8], and the analysis of cultural dimensions for the pragmatist approach [P3]. We acknowledge that different methods might have yielded varying results, hence influenced the overall assessment. At the same time, we ensured to apply the most common research methods for each approach based on a comprehensive analysis of the state of research in each project.

Participants: We recruited all study participants with the help of our collaboration partners and our personal network, i.e., mailing lists. As a consequence, participants with different cultural and different personal backgrounds joined our user studies. Further, it was not possible to receive feedback from the same participants for all three case study approaches. As other study participants might have provided differing feedback, our general country-specific design insights cannot be blindly generalized or transferred to other countries without further analysis. For this dissertation, however, the main interest lies in evaluating the applicability of the three distinct design approaches and methodologies beyond the mere comparison of country-specific design insights.

Application areas: To embed each case study in an authentic research setting and to gather real world data, we collaborated with different partners in every research stage. Applying different approaches for the same product and application area might be best suitable to juxtapose benefits and drawbacks of different methods, however, time, budget, and project scope constraints did not allow to work together with the same partner for all projects. We see great potential in this context to identify more insights for UX methods in cross-cultural contexts for industrial practice.

Study settings: In our case studies, we did not investigate the influence of further aspects, such as environmental conditions, hardware/device setups, or varying durations of the study, that might have had an impact on the participants evaluations and on the implementation of the methods. Such investigations will help to further study the suitability of the different research and design approaches.

Researchers: The insights of the case study are the results of collaborations with different contributors. Besides not being able to jointly discuss general differences in the application of the different approaches, i.e., measuring (see [P1]), empathic (see [P7] and [P8]), and pragmatist approach (see [P3]), all contributors naturally analyzed the results based on their personal cultural background. In future studies, the influence of the researchers' cultural backgrounds on method application needs to be studied in more detail.

2.3 Tools for Cross-Functional UX Teams

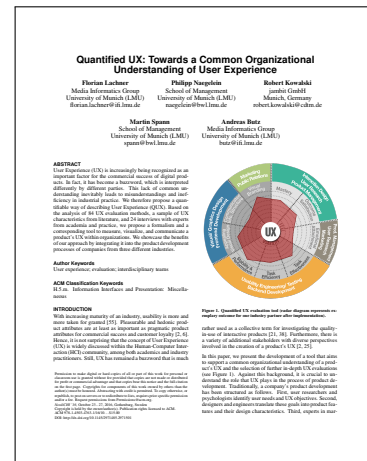
[P2] and [P3] round off this dissertation with the presentation of two concepts for digital UX tools that support cross-functional UX teams to measure, visualize, and communicate UX feedback. The publications collectively address RQ3:

RQ3: “How can digital tools support cross-functional UX teams?”

[P2] Quantified UX: Towards a Common Organizational Understanding of UX

Summary: Stakeholders in traditional design teams traditionally worked separated from each other to identify user needs, develop a product concept, and eventually market a product. In contrast, cross-functional teams, nowadays, work more and more together along different development stages. However, established UX methods lack the capability to satisfactorily communicate a product’s overall UX in interdisciplinary design teams. In this paper, we describe the conceptualization of the UX measurement tool “QUX” (i.e., Quantified UX) that aims to help organizations to measure and communicate a product’s UX across different disciplines based on 24 expert interviews. Finally, we test the concept with partner companies from three different industries. We saw that QUX supports prioritization of user research tasks, allows for benchmarking, and facilitates communication in teams.

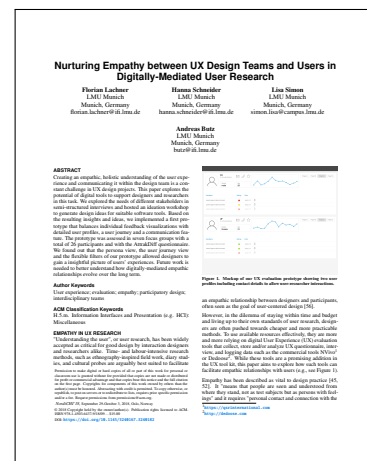
Author contributions: I was the leading author of this publication. Philipp Nägelein contributed significantly and led the concept evaluation, Robert Kowalski designed the figure on the first page. Martin Spann and Andreas Butz revised the paper for conciseness and readability.



[P4] Nurturing Empathy between UX Design Teams and Users in Digitally-Mediated User Research

Summary: In this paper, we investigate design opportunities for digitally-mediated empathic relationships with users. Based on a collaboration with a design agency, we derive current challenges for empathic relationships. Next, we develop a prototype of a digital UX tool that allows interdisciplinary teams to access detailed user profiles, individual feedback visualizations, and a communication feature. Finally, we test the digital tool with 26 study participants in seven focus groups to derive general design insights. The study revealed that our tool helps to establish relationships with users and is useful for long-term studies at later stages.

Author contributions: Lisa Simon developed the prototype and significantly contributed to the user study. Hanna Schneider developed the idea and supervised the development. Andreas Butz supervised the project and revised the paper. I supervised the user study and was the leading author of the final publication.



Conclusion

Obviously, there is much that can be counted, measured, and submitted to statistical analysis. But all human act is impregnated with meaning, and meaning is hard to measure, though it can often be grasped, even if only fleetingly and ambiguously.

V. Turner & E. Bruner, The Anthropology of Experience, 1986

The goal of this dissertation was to advance the field of UX methods, compare different approaches that support the consideration of cultural characteristics during the UX design process, and eventually support cross-functional UX teams. Consequently, this thesis offers a manifold contribution for both UX academia and industry. Initially, I started with an analysis of the status quo of UX methods. According to Wobbrock & Kientz [96] that highlight seven types of research contributions in HCI, the respective publications [P5] and [P6] can be seen as a survey contribution, i.e., a meta-analysis to identify trends and gaps, for UX methods (see Table 3.1). Next, I presented three collaborative case studies in [P1, P7, P8, P3] based on a measuring, an empathic, and a pragmatist approach (according to Battarbee & Koskinen [11]) to investigate how culture can be considered during the UX research process. The studies in the respective publications provide an empirical contribution for cross-cultural UX design as well as a methodological contribution based on the reflection and evaluation of the applied methods. Finally, this dissertation contributes two artifacts for digital UX tools in [P2] and [P4] to enhance the method toolkit of cross-functional UX teams. As part of this chapter, I discuss the overall theoretical contribution related to culturally sensitive design. This term emerged from the comparison of the different approaches in the contributing publications [P1, P7, P8, P3]. Hence, in the next sections, I reflect on my general research approach and contribution, elaborate on my understanding of culturally sensitive design, and outline starting points for future work.

Contribution Type	Contribution of this Dissertation	Related RQ
Survey Contribution	The full paper [P6] and the associated workshop paper [P5] jointly present the results of a thorough meta-review of UX evaluation methods in empirical studies.	RQ1
Empirical Contribution	The results of the cross-cultural case studies in [P1, P7, P8, P3] offer insights in culturally differing user behaviors and preferences.	RQ2
Methodological Contribution	The main contribution of this dissertation resides in the investigation of methods for cross-cultural considerations during the UX process and methods for cross-functional collaboration of UX teams.	RQ2 & RQ3
Artifact Contribution	[P2] and [P4] showcase two artifacts for digital UX tools.	RQ3
Theoretical Contribution	I derive the concept of culturally sensitive design and embed the applied UX approaches of [P1, P7, P8, P3] into the cultural model of Hofstede [40] in this chapter.	RQ1 & RQ2

Table 3.1: Survey, empirical, methodological, artifact, and theoretical contribution of this dissertation according to Wobbrock & Kientz [96].

3.1 Reflection

The variety of UX methods and the applicability of distinct methods in different contexts have been widely discussed in HCI literature (see, e.g., Robinson et al. [75], Bargas-Avila & Hornbæk [9], Vermeeren et al. [92], Obrist et al. [67], or Roto et al. [83]). Consequently, one could raise the question why it was necessary to conduct another meta-review on UX methods as part of the analysis of UX in cross-cultural contexts? In fact, neither the field of UX evaluation nor cross-cultural considerations in HCI currently benefit from a common understanding, established guidelines or frameworks [53]. Hence, the review of UX methods in [P6] provides a valuable theoretical basis for the cross-cultural case studies in this thesis. In particular, the review revealed that quantitative methods (e.g., questionnaires or log analysis) still represent two of the three most frequently applied methods although recent studies emphasize the importance of qualitative data in UX (see Law et al. [53]). As a result, I started the investigation of cultural differences or similarities in UX with a measuring-focused log analysis in [P1]. This measuring approach showcased the applicability of metrics based on quantitative data, more precisely large-scaled behavioral data. Based on the analysis of log data from a website plug-in, we saw that there was a relationship between the country of origin (France, Germany, or Italy) and the interaction with the plug-in. In our case study, Italian users relied more on the plug-in's suggestion for a shoe size at the respective e-commerce website compared to French and German users. In addition, the case study revealed divergent temporal usage patterns as well as country-specific adoption and dropout rates. However, we could not understand the reasonings for varying behaviors of different users due to the lack of qualitative insights. In relation to the iceberg model, the measuring approach was rather suitable to study the visible layer of culture (see Figure 3.1).

Since [P6] showed that interviews represent the second most commonly used UX method, [P7] and [P8] applied an empathic approach based on qualitative ethnographic interviews. This project represented a turning point of this dissertation since the initial work - as previously outlined - focused on "quantifying UX" inspired by the increasing relevance of UX metrics (see et. Rodden et al. [77] and Law et al. [53]). The on-site study as presented in [P8], in contrast, evinced that qualitative and empathy-focused UX approaches allow to collect rich insights at various levels, from visible aspects to contextual information and eventually unconscious values, as Irani et al. [43] similarly highlight in their work on postcolonial computing. More specifically, the interviews illustrated that our participants in India had fewer concerns regarding data privacy and even different mental models of data usage compared to participants from a similar study in Germany that my co-authors from [P7] conducted earlier. When we talked about social media in our study, one participant, for example, mentioned that "all [his] data information is with Mark Zuckerberg, and the government is paying Facebook for each account a citizen created" (see [P8], page 4). A measuring approach would have most certainly not provided such an in-depth insight. The empathic approach, in contrast, was helpful to study deeper cultural layer, i.e., unspoken and unconscious rules (see Figure 3.1). Nevertheless, such on-site research is generally time-consuming and costly, whereby the remote feedback of measuring approaches can offer continuous insights at lower costs. Consequently, as emphasized in [P6] and postulated from Vermeeren and colleagues [92], suitable approaches and frameworks for effective method triangulation are required to better leverage project resources, particularly for further cross-cultural considerations in the future. In this context, my co-authors from [P6] and I identified eight UX method triangulation patterns, e.g., deeper insights often motivates method triangulation, pre- and post-evaluation is infrequently studied, and the combination of interviews and questionnaires as a common post-use evaluation strategy, that need to be studied in more detail.

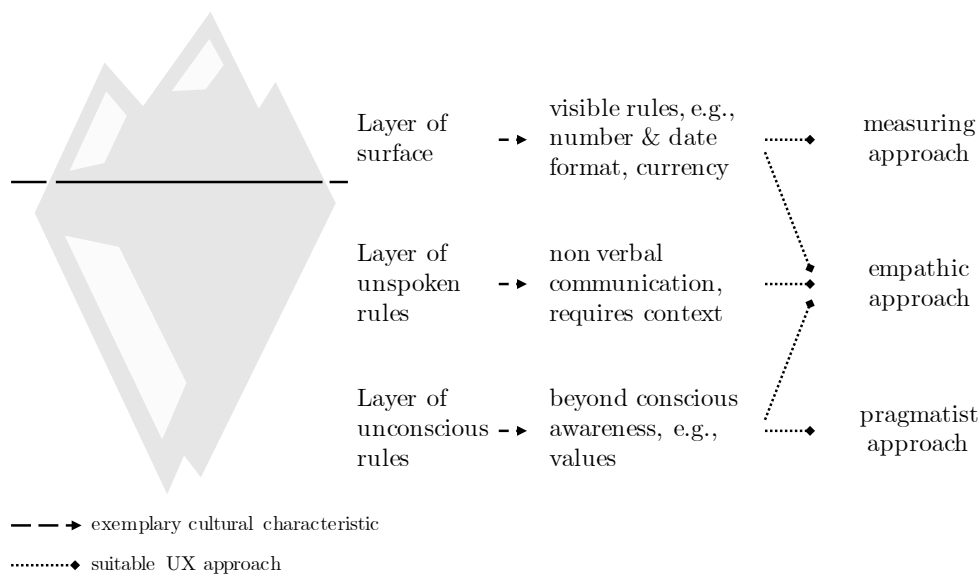


Figure 3.1: Suitable UX approaches (own elaboration) for culturally sensitive research and design (adapted from the general UX approaches from Battarbee & Koskinen [11] based on the case studies [P1, P7, P8, P3]) to investigate different characteristics of culture according to the iceberg model of Hofstede [40].

The pragmatist approach in [P3] aimed to enrich our mental model of culturally sensitive design elements based on cultural theory (i.e., cultural dimensions) at low costs but still with a focus on deeper (i.e., unconscious) cultural layers. The goal was to identify and better understand culturally diverse preferences as part of a goal-directed design process (see Cooper et al. [20] or Hassenzahl [32]). This work showcased that cultural dimensions are helpful to identify culturally sensitive design aspects in general. My co-authors from [P3] and I, for example, derived the impact of hierarchy in a society on website design from the cultural dimensions according to Hofstede et al. [39]. In our case study, the Vietnamese study participants (generally high acceptance of hierarchy) highly valued the photos of individuals on the website (i.e., the founders of the platform that we analyzed) whereas the German study participants (generally low acceptance of hierarchy) mentioned that such photos are not necessary or even too dominant (see [P3], page 9). At the same time, the study found inconsistent insights compared to guidelines that we have derived from the theoretical analysis. As a consequence, cultural dimensions, in this context, have been rather helpful as an inspiration to identify unconscious cultural aspects or values that may influence the design of a product, yet further investigations were necessary to validate respective design decisions (see Figure 3.1). In the pragmatist case study in [P3], for example, a benchmark analysis of established websites in both target countries complemented the theoretical analysis and provided further insights on country specific design elements.

In sum, the different approaches were useful for each case study due to differing contextual factors. The measuring approach in [P1] was helpful on a meta-level to identify visible cultural differences and similarities on a design-level. The empathic approach in [P7, P8] (i.e., ethnography) was experienced as beneficial in various contexts for both conscious design aspects and unconscious cultural aspects. The pragmatist approach in [P3] (i.e., cultural theory) was helpful as an inspiration for unconscious values whose influence on design should be investigated in additional in-depth studies.

Conclusion

Figure 3.1 illustrates the suitability of different approaches for different aspects according to the iceberg model of Hoft [40] based on my reflection within this dissertation. In an ideal world, I suggest to combine the different research approaches, e.g., develop awareness for culturally sensitive design elements based on cultural theory (pragmatist approach), conduct on-site research to talk to and to observe users (empathic approach), and finally measure and track user behavior (measuring approach) to optimize a product or service. Such a holistic culturally sensitive design process based on suitable method triangulation is, to the best of my knowledge and belief, not yet applied in the field of UX in cross-cultural contexts.

The digital UX tools for cross-functional collaboration of UX teams in [P2] and [P4] resulted from similar motivations and hence serve similar needs. First, a measuring-focused survey tool to identify starting points for in-depth studies (see [P2]) and, second, an empathy-focused communication platform to foster relationships with users (see [P4]). In general, the applicability of UX methods in a multidisciplinary setting is seen as an important requirement for UX methods [91]. Thus, I see these tools as fruitful starting points for further work at the intersection of cross-functional collaboration and cultural considerations in the field of UX.

As part of this thesis, a UX process that suitably adjusts and combines appropriate methods to consider cultural peculiarities to eventually enrich the designers' and researchers' mental models, i.e., minimize biases due to ones own cultural background, throughout the design process is described as a *culturally sensitive design* process. The term emerged from the methodological reflection as described above, was initially used in [P1] and put in concrete terms in [P3]. Based on this understanding of the term, I want to initiate further research efforts to generate additional insights, guidelines, and knowledge about how we can best design with culture in mind across the whole UX process in collaboration with all involved disciplines.

Summary: Culturally Sensitive Design

Culturally sensitive design describes a design process that combines suitable research methods to consider cultural peculiarities throughout the whole design, starting from the analysis of relevant cultural theory, reflection and validation of theoretical insights through active interaction with and observation of people in their local context, and measuring user behavior in the long run.

3.2 Future Work

The contribution of this dissertation as subdivided and described in Table 3.1 represents another step towards a better understanding of cultural considerations in UX research and design. At a time where the fourth (i.e., the multidisciplinary) design paradigm is gaining in importance this understanding becomes increasingly important. However, further work and in-depth analyses are required to eventually draw the big picture of UX in cross-cultural contexts. The three case studies in [P1, P7, P8, P3] provided an initial insight how different methods can be applied in distinct cross-cultural contexts. Nevertheless, the respective design and study setting decisions of the individually contributing publications that kept time and cost effort at a manageable level leave room for additional research questions beyond the inherent constraints and limitations summarized in Section 2.2.4.

First, regardless of the significance of cross-cultural considerations (as evidenced in multiple HCI studies) it is necessary to study and compare further design strategies in addition to national cultural boundaries, e.g., personality (see, e.g., Schneider et al. [87]), age (see, e.g., Morris & Venkatesh [61]) or further influencing factors in this context. Thus, we will better understand the role of culture in relation to other influences to eventually derive a more in-depth theoretical framework for culturally sensitive design.

Second, more countries and domains need to be studied in detail to challenge the applicability of the framework in Figure 3.1 on a global scale. Furthermore, not every cultural customization might be suitable for all target countries (see Reinecke & Bernstein [72]). Thus, future investigations are necessary to understand how to best balance globally defined and locally adjusted design aspects, e.g., for distinct (cultural) regions. In this context, especially the applicability of cultural theory (i.e., cultural dimensions) as described in this thesis requires future work. Until now, researchers did not yet find a common ground on the application of existing cultural theory in HCI. Yaaqoubi & Reinecke [98], e.g., argue in favor of the use of cultural dimensions, Winschiers [95] presents an argumentation against narrowing down cultural aspects to two-dimensional scales, particularly when it comes to developing countries [59]. Future work may include the study of cultural dimensions from other researchers (see Section 1.2) or an alternative, potentially newly defined cultural theory for HCI since most cultural dimensions have been defined before the age of ubiquitous computing.

Third, future work needs to extend the culturally sensitive design framework in Section 3.1 with a more distinct analysis of additional methods and eventually the indication of particular methods for specific use cases and objectives. In this regard, current work related to automated UX analyses or user classification based on machine learning models represent a promising and up-and-coming research area (see, e.g., Dove et al. [22] or Yang et al. [99]). In addition, more work is required to understand how different disciplines can contribute to different aspects in a culturally sensitive design process at different cultural levels, similar to the insights from [P2] for UX in general.

Finally, research related to UX in cross-cultural contexts will remain an ongoing task due to the dynamic nature of the cultural concept [68]. However, interdisciplinary collaboration will allow us to define and establish a suitable terminology as well as methods and tools. Such methods and tools will need to take cultural changes and dynamics into account as, in the words of Don Norman [65], “with the passage of time, the psychology of people stays the same, but the tools and objects in the world change. Cultures change.”

The primary challenge of the first two decades of interaction design practice was to invent the process, tools, roles, and methods needed to succeed. Each of these best practices is now evolving as we integrate our skills more deeply into our teams. Specifically, we need to work more effectively with business people and developers.

A. Cooper, R. Reimann, D. Cronin & C. Nessel, About Face, 2014

Conclusion

LIST OF FIGURES

1.1	Overview and relation of the chapters in this thesis - including the relation of the respective research questions (RQ).	1
1.2	Temporal dimensions of UX illustrating the origin of experiences at different stages as well as the internal evaluation process and associated dependencies (adapted from Roto et al. [78]).	2
1.3	Different metaphorical layers of culture according to the iceberg model whereof a large part of cultural characteristics are not directly observable (adapted from Hoft [40] and Galletta & Zhang [26]).	5
1.4	Chronological correlation (x-axis) and thematic coherence (y-axis) of this dissertation including the structural relationships of the original contributing publications.	10
3.1	Suitable UX approaches (own elaboration) for culturally sensitive research and design (adapted from the general UX approaches from Battarbee & Koskinen [11] based on the case studies [P1, P7, P8, P3]) to investigate different characteristics of culture according to the iceberg model of Hoft [40].	21

LIST OF TABLES

2.1	Overview of the publications as well as their methods and outcomes.	11
3.1	Survey, empirical, methodological, artifact, and theoretical contribution of this dissertation according to Wobbrock & Kientz [96].	19

REFERENCES

- [1] ISO 9241-210:2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. International Standardization Organization (ISO), 2010 (cited on p. 2).
- [2] Rui Alves, Pedro Valente, and Nuno Jardim Nunes. “The State of User Experience Evaluation Practice”. In: *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. NordiCHI '14. ACM, 2014, pp. 93–102. DOI: 10.1145/2639189.2641208 (cited on pp. 2, 3, 7).
- [3] Martina Balestra, Lior Zalmanson, Coye Cheshire, Ofer Arazy, and Oded Nov. “It Was Fun, but Did It Last?: The Dynamic Interplay Between Fun Motives and Contributors’ Activity in Peer Production”. In: *Proc. ACM Hum.-Comput. Interact.* 1.CSCW (2017), 21:1–21:13. DOI: 10.1145/3134656 (cited on p. 2).
- [4] Liam J. Bannon. “From human factors to human actors: The role of psychology and human-computer interaction studies in system design”. In: *Readings in Human-Computer Interaction*. Elsevier, 1995, pp. 205–214 (cited on p. 6).
- [5] Liam J. Bannon. “Reimagining HCI: Toward a More Human-centered Perspective”. In: *interactions* 18.4 (2011), pp. 50–57. DOI: 10.1145/1978822.1978833 (cited on p. 6).
- [6] Wendy Barber and Albert Badre. “Culturability: The merging of culture and usability”. In: *Proceedings of the 4th Conference on Human Factors and the Web*. Vol. 7. 4. 1998, pp. 1–10 (cited on p. 4).
- [7] Javier A. Bargas-Avila and Kasper Hornbæk. “Foci and Blind Spots in User Experience Research”. In: *Interactions* 19.6 (2012), pp. 24–27. DOI: 10.1145/2377783.2377790 (cited on p. 1).
- [8] Javier A. Bargas-Avila and Kasper Hornbæk. “Foci and Blind Spots in User Experience Research”. In: *interactions* 19.6 (2012), pp. 24–27. DOI: 10.1145/2377783.2377790 (cited on p. 2).
- [9] Javier A. Bargas-Avila and Kasper Hornbæk. “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*. CHI '11. ACM, 2011, pp. 2689–2698. DOI: 10.1145/1978942.1979336 (cited on pp. 3, 4, 9, 20).
- [10] Katja Battarbee. “Defining Co-experience”. In: *Proceedings of the 2003 International Conference on Designing Pleasurable Products and Interfaces*. DPPI '03. ACM, 2003, pp. 109–113. DOI: 10.1145/782896.782923 (cited on p. 3).

- [11] Katja Battarbee and Ilpo Koskinen. “Co-experience: user experience as interaction”. In: *CoDesign* 1.1 (2005), pp. 5–18 (cited on pp. 3, 6, 13, 19, 21).
- [12] Alan F. Blackwell. “HCI As an Inter-Discipline”. In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '15. ACM, 2015, pp. 503–516. DOI: 10.1145/2702613.2732505 (cited on p. 7).
- [13] Eli Blevis, Ilpo K Koskinen, Kun-Pyo Lee, Susanne Bødker, Lin-Lin Chen, Youn-kyung Lim, Huaxin Wei, and Ron Wakkary. “Transdisciplinary interaction design in design education”. In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM. 2015, pp. 833–838 (cited on pp. 6, 7).
- [14] Mark Blythe, Josephine Reid, Peter Wright, and Erik Geelhoed. “Interdisciplinary criticism: analysing the experience of riot! a location-sensitive digital narrative”. In: *Behaviour & Information Technology* 25.2 (2006), pp. 127–139 (cited on p. 2).
- [15] Keld Bødker and Jesper Strandgaard Pedersen. “Workplace cultures: Looking at artifacts, symbols and practices”. In: *Design at work*. Lawrence Erlbaum Associates, Incorporated, 1991, pp. 121–136 (cited on p. 4).
- [16] Susanne Bødker. “Third-wave HCI, 10 Years Later - Participation and Sharing”. In: *interactions* 22.5 (2015), pp. 24–31. DOI: 10.1145/2804405 (cited on pp. 6, 7).
- [17] Kirsten Boehner, Rogério DePaula, Paul Dourish, and Phoebe Sengers. “How emotion is made and measured”. In: *International Journal of Human-Computer Studies* 65.4 (2007), pp. 275–291 (cited on p. 2).
- [18] Christine L Borgman. “The user’s mental model of an information retrieval system: an experiment on a prototype online catalog”. In: *International Journal of man-machine studies* 24.1 (1986), pp. 47–64 (cited on p. 4).
- [19] Daniel G. Cabrero, Arminda Guerra Lopes, and Barbara Rita Barricelli. “HCI Within Cross-Cultural Discourses of Globally Situated Rhetorical and Etymological Interactions”. In: *International Conference on Cross-Cultural Design*. Springer. 2016, pp. 16–25 (cited on p. 7).
- [20] Alan Cooper, Robert Reimann, David Cronin, and Christopher Noessel. *About face: the essentials of interaction design*. John Wiley & Sons, 2014 (cited on p. 21).
- [21] Sarah Diefenbach and Marc Hassenzahl. “The Dilemma of the Hedonic - Appreciated, but Hard to Justify”. In: *Interacting with Computers* 23.5 (2011), pp. 461–472. DOI: 10.1016/j.intcom.2011.07.002 (cited on p. 3).
- [22] Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. “UX Design Innovation: Challenges for Working with Machine Learning as a Design Material”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM. 2017, pp. 278–288 (cited on p. 23).
- [23] Gabrielle Ford and Helene Gelderblom. “The effects of culture on performance achieved through the use of human computer interaction”. In: *Proceedings of the 2003 annual research conference of the South African institute of computer scientists and information technologists on Enablement through technology*. South African Institute for Computer Scientists and Information Technologists. 2003, pp. 218–230 (cited on p. 4).
- [24] Jodi Forlizzi and Katja Battarbee. “Understanding experience in interactive systems”. In: *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*. ACM. 2004, pp. 261–268 (cited on p. 3).

- [25] Batya Friedman. “Value-sensitive Design”. In: *interactions* 3.6 (1996), pp. 16–23. DOI: 10 . 1145/242485 . 242493 (cited on p. 6).
- [26] Dennis Galletta and Ping Zhang. Human-computer interaction and management information systems: Applications. Vol. 6. ME Sharpe, 2006 (cited on p. 5).
- [27] Edward Twitchell Hall. Beyond culture. Anchor, 1989 (cited on p. 5).
- [28] Edward Twitchell Hall. The dance of life: The other dimension of time. Anchor, 1989 (cited on p. 5).
- [29] Edward Twitchell Hall. The hidden dimension. Doubleday & Co, 1966 (cited on p. 5).
- [30] Steve Harrison, Deborah Tatar, and Phoebe Sengers. “The three paradigms of HCI”. In: *Alt. Chi. Session at the SIGCHI Conference on Human Factors in Computing Systems San Jose, California, USA*. 2007, pp. 1–18 (cited on p. 6).
- [31] Marc Hassenzahl. “Experiences Before Things: A Primer for the (Yet) Unconvinced”. In: *CHI '13 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '13. ACM, 2013, pp. 2059–2068. DOI: 10 . 1145/2468356 . 2468724 (cited on p. 2).
- [32] Marc Hassenzahl. “The Thing and I: Understanding the Relationship Between User and Product”. In: *Funology 2: From Usability to Enjoyment*. Ed. by Mark Blythe and Andrew Monk. Springer International Publishing, 2018, pp. 301–313. DOI: 10 . 1007/978-3-319-68213-6_19 (cited on pp. 1, 21).
- [33] Marc Hassenzahl. “User Experience (UX): Towards an Experiential Perspective on Product Quality”. In: *Proceedings of the 20th Conference on L'Interaction Homme-Machine*. IHM '08. ACM, 2008, pp. 11–15. DOI: 10 . 1145/1512714 . 1512717 (cited on pp. 2, 3).
- [34] Marc Hassenzahl, Sarah Diefenbach, and Anja Göritz. “Needs, affect, and interactive products—Facets of user experience”. In: *Interacting with computers* 22.5 (2010), pp. 353–362 (cited on p. 3).
- [35] Marc Hassenzahl, Kai Eckoldt, Sarah Diefenbach, Matthias Laschke, Eva Len, and Joonhwan Kim. “Designing moments of meaning and pleasure. Experience design and happiness”. In: *International Journal of Design* 7.3 (2013) (cited on p. 3).
- [36] Marc Hassenzahl and Noam Tractinsky. “User experience—a research agenda”. In: *Behaviour & information technology* 25.2 (2006), pp. 91–97 (cited on pp. 2, 3).
- [37] Mats Hellman and Kari Rönkkö. “Is User Experience supported effectively in existing software development processes?” In: *Proc. of COST294-MAUSE Workshop on Valid Useful User Experience Measurement (VUUM)*. Reykjavik, Island. 2008, pp. 32–37 (cited on p. 2).
- [38] Geert Hofstede. Culture’s consequences: Comparing values, behaviors, institutions and organizations across nations. Sage publications, 2001 (cited on pp. 4, 5).
- [39] Geert Hofstede, Gert Jan Hofstede, and Michael Minkov, eds. Cultures and Organizations. Software of the Mind. Intercultural Cooperation and Its Importance for Survival. McGraw-Hill, 2010 (cited on pp. 4, 5, 15, 21).
- [40] Nancy L. Hoft. “International Users Interface”. In: ed. by Elisa M. del Galdo and Jakob Nielsen. John Wiley & Sons, Inc., 1996. Chap. Developing a Cultural Model, pp. 41–73 (cited on pp. 4, 5, 19, 21, 22).

- [41] Kasper Hornbæk and Morten Hertzum. “Technology Acceptance and User Experience: A Review of the Experiential Component in HCI”. In: *ACM Trans. Comput.-Hum. Interact.* 24.5 (2017), 33:1–33:30. DOI: 10.1145/3127358 (cited on p. 1).
- [42] Lilly C. Irani and Paul Dourish. “Postcolonial Interculturality”. In: *Proceedings of the 2009 International Workshop on Intercultural Collaboration. IWIC '09.* ACM, 2009, pp. 249–252. DOI: 10.1145/1499224.1499268 (cited on p. 5).
- [43] Lilly C. Irani, Janet Vertesi, Paul Dourish, Kavita Philip, and Rebecca E. Grinter. “Postcolonial Computing: A Lens on Design and Development”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '10. ACM, 2010, pp. 1311–1320. DOI: 10.1145/1753326.1753522 (cited on pp. 5, 14, 20).
- [44] Christian Jetter and Jens Gerken. “A simplified model of user experience for practical application”. In: *2nd COST294-MAUSE.* 2007, pp. 106–111 (cited on p. 7).
- [45] Patrick W. Jordan. “Human factors for pleasure in product use”. In: *Applied ergonomics* 29.1 (1998), pp. 25–33 (cited on p. 7).
- [46] Anu Kankainen. Thinking model and tools for understanding user experience related to information appliance product concepts. Helsinki University of Technology, 2002 (cited on p. 6).
- [47] Evangelos Karapanos, John Zimmerman, Jodi Forlizzi, and Jean-Bernard Martens. “User Experience over Time: An Initial Framework”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '09. ACM, 2009, pp. 729–738. DOI: 10.1145/1518701.1518814 (cited on pp. 2, 3).
- [48] Florence R. Kluckhohn and Fred L. Strodtbeck. Variations in Value Orientation. Row, Peterson, 1961 (cited on p. 5).
- [49] Joke Kort, Arnold Vermeeren, and Jenneke Fokker. “Conceptualizing and measuring user experience”. In: *Proc. of COST294-MAUSE Workshop on Valid Useful User Experience Measurement (VUUM).* Reykjavik, Island. 2008, pp. 57–64 (cited on p. 2).
- [50] Carine Lallemand, Vincent Koenig, and Guillaume Gronier. “How Relevant is an Expert Evaluation of User Experience Based on a Psychological Needs-driven Approach?” In: *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational.* NordiCHI '14. ACM, 2014, pp. 11–20. DOI: 10.1145/2639189.2639214 (cited on p. 2).
- [51] Effie Lai-Chong Law. “The Measurability and Predictability of User Experience”. In: *Proceedings of the 3rd ACM SIGCHI Symposium on Engineering Interactive Computing Systems.* EICS '11. ACM, 2011, pp. 1–10. DOI: 10.1145/1996461.1996485 (cited on pp. 2, 3).
- [52] Effie Lai-Chong Law, Virpi Roto, Marc Hassenzahl, Arnold POS Vermeeren, and Joke Kort. “Understanding, scoping and defining user experience: a survey approach”. In: *Proceedings of the SIGCHI conference on human factors in computing systems.* ACM. 2009, pp. 719–728 (cited on p. 2).
- [53] Effie Lai-Chong Law, Paul van Schaik, and Virpi Roto. “Attitudes towards user experience (UX) measurement”. In: *International Journal of Human-Computer Studies* 72.6 (2014), pp. 526–541 (cited on pp. 2, 7, 20).
- [54] Effie Lai-Chong Law, Arnold P. O. S. Vermeeren, Marc Hassenzahl, and Mark Blythe. “Towards a UX Manifesto”. In: *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...But Not As We Know It - Volume 2.* BCS-HCI '07. BCS Learning & Development Ltd., 2007, pp. 205–206 (cited on p. 2).

- [55] Effie Law, Virpi Roto, Arnold P.O.S. Vermeeren, Joke Kort, and Marc Hassenzahl. “Towards a Shared Definition of User Experience”. In: *CHI '08 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '08. ACM, 2008, pp. 2395–2398. DOI: 10.1145/1358628.1358693 (cited on p. 2).
- [56] Aaron Marcus and Emilie West Gould. “Crosscurrents: cultural dimensions and global Web user-interface design”. In: *interactions* 7.4 (2000), pp. 32–46 (cited on p. 4).
- [57] John McCarthy and Peter Wright. “Technology as experience”. In: *interactions* 11.5 (2004), pp. 42–43 (cited on p. 2).
- [58] Kimra McPherson, Kai Huotari, F. Yo-Shang Cheng, David Humphrey, Coye Cheshire, and Andrew L. Brooks. “Glitter: A Mixed-methods Study of Twitter Use During Glee Broadcasts”. In: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion*. CSCW '12. ACM, 2012, pp. 167–170. DOI: 10.1145/2141512.2141569 (cited on p. 3).
- [59] Samantha Merritt and Shaowen Bardzell. “Postcolonial Language and Culture Theory for HCI4D”. In: *CHI '11 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '11. ACM, 2011, pp. 1675–1680. DOI: 10.1145/1979742.1979827 (cited on p. 23).
- [60] David Moher, Deborah J. Cook, Susan Eastwood, Ingram Olkin, Drummond Rennie, and Donna F Stroup. “Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement”. In: *Oncology Research and Treatment* 23.6 (2000), pp. 597–602 (cited on p. 9).
- [61] Michael G. Morris and Viswanath Venkatesh. “Age differences in technology adoption decisions: Implications for a changing work force”. In: *Personnel psychology* 53.2 (2000), pp. 375–403 (cited on p. 23).
- [62] Livia J. Müller, Elisa D. Mekler, and Klaus Opwis. “Facets In HCI: Towards Understanding Eudaimonic UX – Preliminary Findings”. In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '15. ACM, 2015, pp. 2283–2288. DOI: 10.1145/2702613.2732836 (cited on p. 2).
- [63] Brad A. Myers. “A Brief History of Human-computer Interaction Technology”. In: *interactions* 5.2 (1998), pp. 44–54. DOI: 10.1145/274430.274436 (cited on p. 6).
- [64] Jakob Nielsen. *Designing user interfaces for international use*. Elsevier Science Publishers Ltd., 1990 (cited on p. 4).
- [65] Don Norman. *The design of everyday things: Revised and expanded edition*. Basic Books, 2013 (cited on pp. 3, 23).
- [66] Don Norman and Jakob Nielsen. *The Definition of User Experience (UX)*. 2018. URL: <https://www.nngroup.com/articles/definition-user-experience/> (cited on p. 2).
- [67] Marianna Obrist, Virpi Roto, and Kaisa Väänänen-Vainio-Mattila. “User Experience Evaluation: Do You Know Which Method to Use?” In: *CHI '09 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '09. ACM, 2009, pp. 2763–2766. DOI: 10.1145/1520340.1520401 (cited on pp. 2, 3, 20).
- [68] Kavita Philip, Lilly C. Irani, and Paul Dourish. “Postcolonial computing: A tactical survey”. In: *Science, Technology, & Human Values* 37.1 (2012), pp. 3–29 (cited on p. 23).

- [69] B. Joseph Pine and James H. Gilmore. “Welcome to the experience economy”. In: *Harvard business review* 76 (1998), pp. 97–105 (cited on p. 2).
- [70] Stuart Reeves. “Human-computer interaction as science”. In: *Proceedings of The Fifth Decennial Aarhus Conference on Critical Alternatives*. Aarhus University Press. 2015, pp. 73–84 (cited on p. 7).
- [71] Katharina Reinecke and Abraham Bernstein. “Improving Performance, Perceived Usability, and Aesthetics with Culturally Adaptive User Interfaces”. In: *ACM Trans. Comput.-Hum. Interact.* 18.2 (2011), 8:1–8:29. DOI: 10.1145/1970378.1970382 (cited on p. 4).
- [72] Katharina Reinecke and Abraham Bernstein. “Knowing What a User Likes: A Design Science Approach to Interfaces that Automatically Adapt to Culture.” In: *Mis Quarterly* 37.2 (2013) (cited on pp. 4, 23).
- [73] Katharina Reinecke, Minh Khoa Nguyen, Abraham Bernstein, Michael Näf, and Krzysztof Z. Gajos. “Doodle around the world: Online scheduling behavior reflects cultural differences in time perception and group decision-making”. In: *Proceedings of the 2013 conference on Computer supported cooperative work-CSCW’13*. ACM Press. 2013 (cited on p. 4).
- [74] Xiangshi Ren. “Rethinking the Relationship between Humans and Computers.” In: *IEEE Computer* 49.8 (2016), pp. 104–108 (cited on p. 6).
- [75] Joy Robinson, Candice Lanius, and Ryan Weber. “The Past, Present, and Future of UX Empirical Research”. In: *Commun. Des. Q. Rev* 5.3 (2018), pp. 10–23. DOI: 10.1145/3188173.3188175 (cited on p. 20).
- [76] Simon Robinson and Matt Jones. “Mobile UX: Breaking the Glass to Richer User Experiences”. In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’16. ACM, 2016, pp. 969–972. DOI: 10.1145/2851581.2856677 (cited on p. 2).
- [77] Kerry Rodden, Hilary Hutchinson, and Xin Fu. “Measuring the user experience on a large scale: user-centered metrics for web applications”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2010, pp. 2395–2398 (cited on pp. 6, 20).
- [78] Virpi Roto, Effie Law, Arnold Vermeeren, and Jettie Hoonhout. UX White Paper. Bringing clarity to the concept of user experience. Result from Dagstuhl Seminar on Demarcating User Experience, 2011 (cited on pp. 2, 4).
- [79] Virpi Roto, Jung-Joo Lee, Tuuli Mattelmäki, and John Zimmerman. “Experience Design Meets Service Design: Method Clash or Marriage?” In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA ’18. ACM, 2018, W26:1–W26:6. DOI: 10.1145/3170427.3170626 (cited on p. 7).
- [80] Virpi Roto, Ming Lee, Kari Pihkala, Brenda Castro, Arnold Vermeeren, Effie Law, Kaisa Väänänen-Vainio-Mattila, Jettie Hoonhout, and Marianna and Obrist. All about UX. Information for user experience professionals. 2018. URL: <http://www.allaboutux.org/ux-definitions> (cited on p. 2).
- [81] Virpi Roto, Yichen Lu, Harri Nieminen, and Erdem Tural. “Designing for User and Brand Experience via Company-wide Experience Goals”. In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’15. ACM, 2015, pp. 2277–2282. DOI: 10.1145/2702613.2732818 (cited on p. 7).

- [82] Virpi Roto, Yichen Lu, Harri Nieminen, and Erdem Tural. “Designing for user and brand experience via company-wide experience goals”. In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM. 2015, pp. 2277–2282 (cited on p. 2).
- [83] Virpi Roto, Kaisa Väänänen-Vainio-Mattila, Effie Law, and Arnold Vermeeren. “User experience evaluation methods in product development (UXEM’09)”. In: *IFIP Conference on Human-Computer Interaction*. Springer. 2009, pp. 981–982 (cited on pp. 3, 20).
- [84] Virpi Roto, Heli Väättäjä, Effie Law, and Rachel Powers. “Experience Design for Multiple Customer Touchpoints”. In: *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. NordiCHI ’16. ACM, 2016, 146:1–146:3. DOI: 10.1145/2971485.2987685 (cited on p. 7).
- [85] Virpi Roto, Heli Väättäjä, Effie Law, and Rachel Powers. “Experience design for multiple customer touchpoints”. In: *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. ACM. 2016, p. 146 (cited on p. 2).
- [86] Constantin von Saucken, Ioanna Michailidou, and Udo Lindemann. “How to design experiences: macro UX versus micro UX approach”. In: *International Conference of Design, User Experience, and Usability*. Springer. 2013, pp. 130–139 (cited on p. 3).
- [87] Hanna Schneider, Katrin Schauer, Clemens Stachl, and Andreas Butz. “Your Data, Your Vis: Personalizing Personal Data Visualizations”. In: *IFIP Conference on Human-Computer Interaction*. Springer. 2017, pp. 374–392 (cited on p. 23).
- [88] Shalom H. Schwartz. “A theory of cultural values and some implications for work”. In: *Applied psychology* 48.1 (1999), pp. 23–47 (cited on p. 5).
- [89] Kennon M. Sheldon, Andrew J Elliot, Youngmee Kim, and Tim Kasser. “What is satisfying about satisfying events? Testing 10 candidate psychological needs.” In: *Journal of personality and social psychology* 80.2 (2001), p. 325 (cited on p. 3).
- [90] Fons Trompenaars and Charles Hampden-Turner. *Riding the waves of culture: Understanding diversity in global business*. Nicholas Brealey Publishing, 2011 (cited on p. 5).
- [91] Kaisa Väänänen-Vainio-Mattila, Virpi Roto, and Marc Hassenzahl. “Towards practical user experience evaluation methods”. In: *Meaningful measures: Valid useful user experience measurement (VUUM)* (2008), pp. 19–22 (cited on pp. 7, 22).
- [92] Arnold P. O. S. Vermeeren, Effie Lai-Chong Law, Virpi Roto, Marianna Obrist, Jettie Hoonhout, and Kaisa Väänänen-Vainio-Mattila. “User Experience Evaluation Methods: Current State and Development Needs”. In: *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*. NordiCHI ’10. ACM, 2010, pp. 521–530. DOI: 10.1145/1868914.1868973 (cited on pp. 3, 4, 20).
- [93] David A Victor. *International Business Communication*. HarperCollins, 1992 (cited on p. 5).
- [94] Tanja Walsh, Pii Nurkka, and Rod Walsh. “Cultural differences in smartphone user experience evaluation”. In: *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*. ACM. 2010, p. 24 (cited on p. 4).
- [95] Heike Winschiers. “The Challenges of Participatory Design in a Intercultural Context: Designing for Usability in Namibia”. In: *PDC*. 2006, pp. 73–76 (cited on pp. 4, 23).

- [96] Jacob O. Wobbrock and Julie A. Kientz. “Research Contributions in Human-computer Interaction”. In: *Interactions* 23.3 (2016), pp. 38–44. DOI: 10.1145/2907069 (cited on p. 19).
- [97] Peter Wright and John McCarthy. “Empathy and Experience in HCI”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. ACM, 2008, pp. 637–646. DOI: 10.1145/1357054.1357156 (cited on p. 2).
- [98] Judith Yaaqoubi and Katharina Reinecke. “The Use and Usefulness of Cultural Dimensions in Product Development”. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, CS18 (cited on p. 23).
- [99] Qian Yang, Nikola Banovic, and John Zimmerman. “Mapping Machine Learning Advances from HCI Research to Reveal Starting Places for Design Innovation”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, p. 130 (cited on p. 23).

APPENDIX: ORIGINAL PUBLICATIONS

This appendix includes all contributing publications of this thesis in their original format without any modifications (except page numeration) in chronological order.

Quantified UX: Towards a Common Organizational Understanding of User Experience [P2]	A 3
Triangulation in UX Studies: Learning from Experience [P5]	A 13
UX Metrics: Deriving Country-Specific Usage Patterns of a Website Plug-In from Web Analytics [P1]	A 17
Investigating Perceptions of Personalization and Privacy in India [P7]	A 35
A Bermuda Triangle? - A Review of Method Application and Triangulation in User Experience Evaluation [P6]	A 39
Privacy and Personalization: The Story of a Cross-Cultural Field Study [P8]	A 55
Nurturing Empathy between UX Design Teams and Users in Digitally-Mediated User Research [P4]	A 59
Culturally Sensitive User Interface Design: A Case Study with German and Vietnamese Users [P3]	A 73

Quantified UX: Towards a Common Organizational Understanding of User Experience

Florian Lachner
Media Informatics Group
University of Munich (LMU)
florian.lachner@ifi.lmu.de

Philipp Naegelein
School of Management
University of Munich (LMU)
naegelein@bwl.lmu.de

Robert Kowalski
jambit GmbH
Munich, Germany
robert.kowalski@cdtm.de

Martin Spann
School of Management
University of Munich (LMU)
spann@bwl.lmu.de

Andreas Butz
Media Informatics Group
University of Munich (LMU)
butz@ifi.lmu.de

ABSTRACT

User Experience (UX) is increasingly being recognized as an important factor for the commercial success of digital products. In fact, it has become a buzzword, which is interpreted differently by different parties. This lack of common understanding inevitably leads to misunderstandings and inefficiency in industrial practice. We therefore propose a quantifiable way of describing User Experience (QUX). Based on the analysis of 84 UX evaluation methods, a sample of UX characteristics from literature, and 24 interviews with experts from academia and practice, we propose a formalism and a corresponding tool to measure, visualize, and communicate a product's UX within organizations. We showcase the benefits of our approach by integrating it into the product development processes of companies from three different industries.

Author Keywords

User experience; evaluation; interdisciplinary teams

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation: Miscellaneous

INTRODUCTION

With increasing maturity of an industry, usability is more and more taken for granted [55]. Pleasurable and hedonic product attributes are at least as important as pragmatic product attributes for commercial success and customer loyalty [2, 6]. Hence, it is not surprising that the concept of User Experience (UX) is widely discussed within the Human-Computer Interaction (HCI) community, among both academics and industry practitioners. Still, UX has remained a buzzword that is much

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
NordiCHI '16, October 23 - 27, 2016, Gothenburg, Sweden
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-4763-1/16/10...\$15.00
DOI: <http://dx.doi.org/10.1145/2971485.2971501>

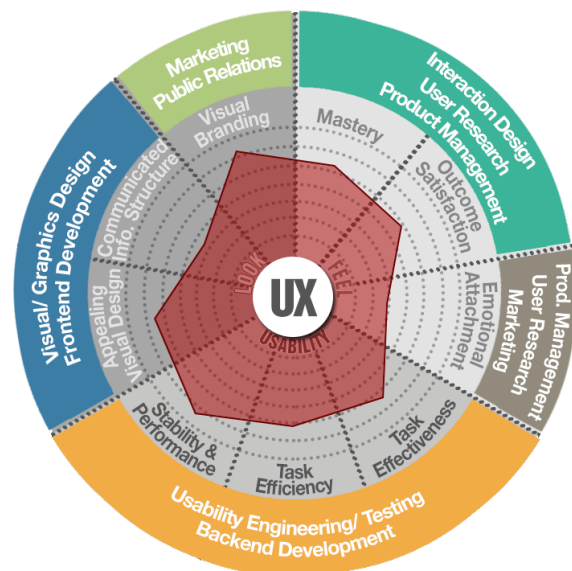


Figure 1. Quantified UX evaluation tool (radar diagram represents exemplary outcome for one industry partner after implementation).

rather used as a collective term for investigating the quality-in-use of interactive products [21, 38]. Furthermore, there is a variety of additional stakeholders with diverse perspectives involved in the creation of a product's UX [2, 25].

In this paper, we present the development of a tool that aims to support a common organizational understanding of a product's UX and the selection of further in-depth UX evaluations (see Figure 1). Against this background, it is crucial to understand the role that UX plays in the process of product development. Traditionally, a company's product development has been structured as follows. First, user researchers and psychologists identify user needs and UX objectives. Second, designers and engineers translate these goals into product features and their design characteristics. Third, experts in mar-

keting and branding define advertising messages to convey the respective experience [29, 66]. Finally, product managers incorporate the UX goals in the business context. Ideally, these steps are not separated from one another but strongly interlinked to ensure a holistic and consistent UX [25]. To create a certain UX, a systematic approach and an associated description of UX are needed to consider and measure the intended experience. Within this context, professionals demand a UX description that contains relevant criteria to support a transfer of UX into industrial practice. Existing development and design methods, however, rarely cope with the required degree of interdisciplinarity to reflect the different angles of e.g. engineering, design, marketing, or psychology [11, 38].

In the following sections, we will discuss the roles relevant for and the disciplines involved in the creation and improvement of a product's UX as part of design processes. Our goal is to address the following research question:

How can we help organizations to measure, visualize, and communicate a product's UX within interdisciplinary teams?

This paper offers two main contributions: First, we propose a specific, quantifiable way of describing user experience, which we call quantified UX (QUX). Second, we develop a graphical tool that connects these UX characteristics with associated disciplines in a visually appealing way to support the compact communication of UX goals within an organization.

UX THEORY

After several years of UX research, scholars seem to have reached consensus with regard to experience-oriented concepts that exceed traditional functionality and usability considerations [23]. UX evaluation ranges from the analysis of psychological needs to task-oriented user goals or guidelines [2]. The satisfaction of human needs is seen as a driver of experiences [60]. However, the consideration of such psychological needs is rather suitable for a *macro perspective*, i.e., the product's overall purpose. For the evaluation of a product on the market, a rather focused *micro perspective* on specific product characteristics, i.e., visceral characteristics, should be analyzed in detail [21, 50]. We argue that it is inevitable for a practically oriented UX evaluation and communication process in interdisciplinary teams to narrow down the broad scope of UX to a quantifiable level. Therefore, we base our research on the concept of product-oriented user goals and define UX as the result of enjoyable interactions and/or anticipated interactions with a product.

Different perceptions of UX are not limited to academia. Many newcomers to the field of UX, and a large number of UX practitioners, struggle with the complexity and vague definition of UX as well [18]. Furthermore, industry practitioners are presented with another challenge: to cope with the inability to talk to users directly while they interact with their product, as (prototype) workshops or laboratory experiments are often cost-intensive and time-consuming. Interdisciplinary project settings may increase the level of complexity even further.

To achieve the intended UX, a large variety of different UX tools and methods are used along the distinct phases of prod-

uct development processes [66]. In general, organizations are thereby particularly interested in long-term UX as they want to foster a positive overall experience rather than focus on temporary emotions [38]. Most academic researchers concentrate on investigating UX from a theoretical perspective. Industry practitioners, in contrast, need tools and methods that make UX assessable and manageable. As a consequence, it has remained a challenge to close this gap between theory and practice [66].

EXISTING UX EVALUATION METHODS

Traditionally, research and development (R&D) departments focused their user research and product testing on usability requirements and quantitative methods, whereas marketing and advertising departments were responsible for communicating a certain experience [66]. However, along with a shift from a usability-focused to an experience-oriented perspective on product interactions, a shift within evaluation methodologies seems to have taken place [5].

The aforementioned gap between academic and practical interpretations of UX leads to substantial differences in the question of how UX should be evaluated [66]. First, user researchers typically disentangle evaluation processes from metric-based methods and focus on qualitative data in order to evaluate UX. However, the practicability of such methods is comparably low since the analysis of associated data may be hard and time-consuming. Thus, organizations and UX practitioners need evaluation tools which are quick to use and provide validated UX measures [25, 69]. Second, since UX evaluation is usually considered costly, UX research often addresses evaluation methodologies for early product stages to identify requirements as early as possible. In industrial practice, however, UX evaluation is mainly pursued to improve and refine existing products [2, 8, 66].

Against this background, we analysed 84 UX evaluation tools from <http://www.allaboutux.org/> [1], a collection of tools of a holistic study of UX measuring methods used in academia and industry [69]. In general, the landscape of UX evaluation offers a wide variety of tools and methods. From the viewpoint of an organization and its interdisciplinary product teams however, we conjecture that it is still hard to measure, visualize, and communicate a product's intended UX. In order to deduce requirements that meet the needs of interdisciplinary teams we examined the focus of the 84 UX evaluation tools from [1]. Thus, we were able to identify requirements for an interdisciplinary QUX approach based on five different evaluation clusters:

1. Measuring Sensation

A range of methods, such as *Emocards* or *Emofaces* [15] as well as *PrEmo* [14], overcome the intangibility of measuring emotions by substituting verbal measurement dimensions with cartoons. Evaluators describe their experiences of using a product by choosing one out of a number of predefined cartoons. Furthermore, *FaceReader* [13] is a tool that automatically tracks facial expressions of users or evaluators. With a focus on feelings and sensation, pragmatic characteristics

recede in the background of UX evaluation. For QUX, however, we want to focus on both hedonic and pragmatic product characteristics.

2. Specific Use Case

Further methods focus on a specific use case, e.g., a specific product or feature: The *Aesthetics scale* [37] helps to evaluate websites, whereas the *Perceived Comfort Assessment* [24] is a method of measuring the comfort level of, e.g., car seats. In contrast, we want to ensure the applicability of our QUX approach for various types of products.

3. Extensive Analysis

The *Experience Sampling Method (ESM)* [59] asks participants at certain times during the day to take notes about their current experiences. The *Outdoor Play Observation Scheme* [3] integrates video recording to analyze childrens' experiences with outdoor games. Both methods indicate the time-consuming analysis of UX evaluations. However, fast-paced industry projects generally require cost-efficient evaluation methods [65].

4. Qualitative Evaluation

The *Day Reconstruction Method (DRM)* [33] is a self-report method where participants note experiences in form of a diary. The *UX Curve* [36] respectively *iScale* [32] measures the quality of an experience over time. Thereby, researchers understand when and how an experience changed but cannot easily analyze why a certain experience was formed or triggered.

5. Questionnaire-based Methods

Questionnaires are widely used in the field of UX evaluation [5]. The *Product Attachment Scale* [48], for example, represents a questionnaire-based evaluation tool to measure the hedonic emotional bonding of a user to a product. On the contrary, *AttrakDiff* [22] analyzes hedonic and pragmatic product attributes via semantic differentials. The summative visualization then again makes it difficult to deduce concrete plans for action in interdisciplinary development projects. For a holistic QUX approach we want to ensure the communication of objective UX goals by incorporating a concrete set of UX characteristics as well as a formative visualization of UX measurements into a visual tool.

METHODOLOGY

The main goal of this paper is to create a tool that helps interdisciplinary development teams to measure, visualize, and communicate a product's UX. We, therefore, aim to reduce the gap between academia and industrial practice by following a systematic methodological approach (see Figure 2).

To start with, we pursue an elaboration of UX characteristics based on the analysis of published work in the field of UX. This literature analysis represents a two-phase process with the goal of identifying relevant published work and extracting prevailing UX characteristics that serve as a basis for discussion in the third phase. The third phase of our analysis process consists of expert interviews with practitioners and researchers in the field of UX. Based on that approach,

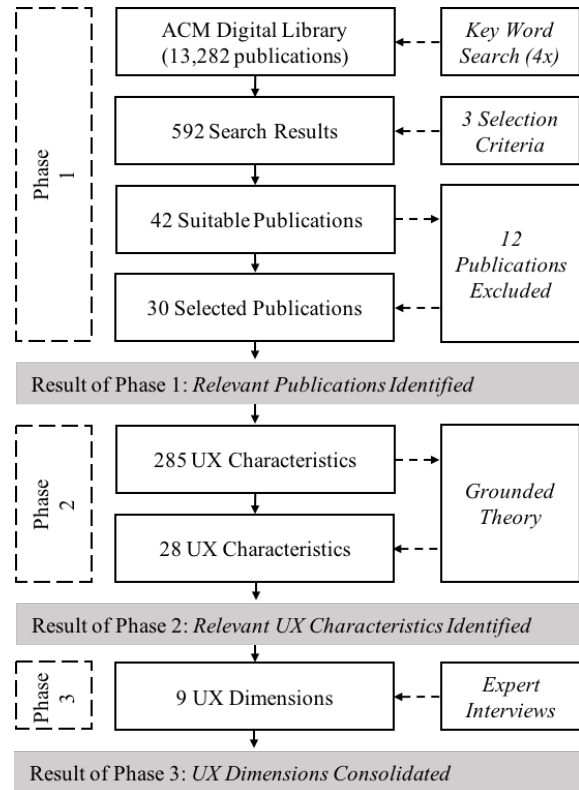


Figure 2. Three-phase methodological approach.

we were able to review a diverse spectrum of UX perspectives and consolidate the extracted characteristics into nine substantial components of UX, which we refer to as "UX dimensions". All in all, the literature analysis does not claim collective exhaustiveness of all possibly existing UX specifications but represents an elaborate foundation to support the subsequent interview process in phase 3.

Phase 1: Relevant Publications Identified

Source selection. We conducted a selective literature analysis inspired by the methodology of [5] with the ACM Digital Library (DL) as a research database in order to develop a guideline for our expert interviews (phase 3). Within the ACM DL, we selected five conferences and one journal for our source research: The Conference on Human Factors in Computing Systems (CHI), the Conference on Computer Supported Cooperative Work (CSCW), the Symposium on User Interface Software and Technology (UIST), Human Computer Interaction with Mobile Devices and Services (MobileHCI), Transactions on Computer-Human Interaction (TOCHI), and the International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp).

We identified the mentioned venues as sources for top HCI publications based on the h5-index indicated by Google Scholar, and the relevance for the underlying research ques-

tion. Besides the top three mentions CHI, CSCW, and UIST on Google Scholar we integrated UbiComp and MobileHCI in our analysis as they focus on HCI topics that we considered as highly relevant for our research (i.e., mobile/handheld devices). Furthermore, TOCHI complements insights from conference proceedings with findings from an established HCI journal.

Search procedure. We used a set of four combinations for every venue using the ACM DL input mask. The first combination consisted of the key words “*user experience*” (all of this text) plus “*communicate, measure, measuring, visualize, framework, tool, guideline, emotions, usability, evaluate, evaluating, evaluation, satisfaction*” (any of this text). The second combination consisted of the key words “*measure, emotions*” (all of this text), the third of “*measure, usability*” (all of this text), and the fourth of “*user, satisfaction*” (all of this text) plus “*measure, framework, visualize*” (any of this text). The combinations were used to search the publications’ abstract in order to focus on publications that are highly relevant for the underlying research question. Furthermore, the particular combinations allowed us to focus on a precise selection of publications to develop a suitable interview guideline for phase 3. We did not limit our search procedure to a specific time span.

Venue	Database (ACM DL)	Results per Combination				Selected Publications
		1	2	3	4	
CHI	7,080	192	4	21	70	11
CSCW	3,061	39	3	1	12	2
UIST	1,214	28	0	2	5	1
MobileHCI	991	131	0	6	27	10
TOCHI	518	24	1	5	6	4
UbiComp	418	13	1	0	1	2
SUM	13,282	592				30

Table 1. Number of identified publications in phase 1 (per venue).

Search results. At the time the search was conducted, the ACM DL provided 13,282 publications for the six venues in total. The search process resulted in 592 relevant papers, articles and works-in-progress as illustrated in Table 1.

Selection process. Before we selected relevant publications from the overall number of search results for phase 2, the authors of this paper (three with an HCI background and two with a business background) jointly defined the following three criteria for a structured selection process: Select publications that (1) describe a UX-related framework, such as in [45], (2) analyze UX characteristics of a specific product, such as in [62], and (3) directly discuss UX-related characteristics, such as in [71]. However, we excluded 12 publications (e.g., [67]) that met one of these criteria but had a focus on UX metrics that have already been covered in our analysis by other identified publications to limit double results. The selection was conducted by the first author, who has extensive knowledge in the field of UX research. Thus, we were able to narrow down the number of relevant publications to 30: [19, 26, 27, 35, 39, 45, 53, 54, 56, 61, 16] from CHI, [9, 30] from CSCW, [47] from UIST, [31, 40, 41, 43, 46, 51, 58, 64, 68,

71] from MobileHCI, [20, 34, 52, 63] from TOCHI, and [62, 70] from UbiComp.

Phase 2: Relevant UX Characteristics Identified

Screening. In this phase, our goal was to detect general UX characteristics within the 30 identified publications. For this purpose, we listed all characteristics that (1) represented UX elements within a theoretical framework, (2) were used to describe a product or service specific UX and (3) were directly mentioned as UX characteristics in any of the 30 publications. In total, we identified 285 UX characteristics.

Consolidation. To reduce the list to a usable number of UX characteristics for our interviews and to derive valuable UX dimensions in phase 3, we followed the interpretive grounded theory research approach by [17] and [28]. Grounded theory is based on a constant comparison of existing data throughout the analysis process and allows researchers to identify recurring key aspects of qualitative data [17]. We were able to recognize seven clusters as well as associated sub-clusters. The outcome of this methodological step is indicated in Table 2. Based on prior experience with the analysis of qualitative UX data, this step was carried on by the first author.

Some authors used rather general *emotions*, such as fun or satisfaction (see [19, 26]) as UX characteristics. Further clusters were based on *design-related*, e.g., color or aesthetics (see [19, 45]), *content-related*, e.g., information quality (see [61]), *technology-related*, e.g., controllable (see [16]), or *outcome-related* characteristics, e.g., error-free (see [64]). With the clusters *further disciplines* and *environment* we were able to evaluate UX characteristics such as status, brand, or context of use (see [20, 51, 68]).

Exclusion. We iteratively compared the clusters to narrow down the number of UX characteristics. To ensure a structured analysis process we jointly defined the following exclusion criteria: Exclude UX characteristics that (1) are specific for a particular product, such as network speed (see [40]), (2) overlap with other UX characteristics, such as social connectivity and social recommendation (see [16, 54]), and (3) are similarly used, such as efficiency of use and efficiency (see [20, 52, 43]). Thus, we ended up with a list of 28 UX characteristics, with all identified clusters represented in our shortlist (see Table 2). However, we realized that these characteristics had not yet offered a clear comprehension of an interdisciplinary QUX approach. To gain a better understanding about practically oriented UX dimensions for interdisciplinary design processes, we used this shortlist as a basis for the interviews in phase 3.

Phase 3: UX Dimensions Consolidated

Participants. Over the course of one week, we conducted expert interviews with 11 UX researchers and 13 UX practitioners to reflect their respective views. The listing below provides an overview of affiliations (less than 24 values due to companies who asked not be mentioned). With regard to expert status, all our academic interviewees are (or were) researchers at institutes with a significant track record of publications at leading HCI conferences. As for practice, our

Cluster	Sub-Cluster	Exemplary characteristics from the selected publications	28 consolidated UX characteristics
Emotions	-	e.g., fun [19], pleasure producing [35], happiness [56], ...	Satisfaction, Pleasure
Design	Form General	e.g., colour [19], clear [41], interface quality [61], ... e.g., natural [64], design [20], visual appearance [16], ...	Interface, Aesthetics
Content	Information General	e.g., information quality [61], information accessed [26], ... e.g., effectiveness [20, 43], usefulness [46], ...	Information, Effectiveness
Technology	Productivity Controllability Progression	e.g., efficiency [16, 35, 52], efficiency of use [20], ... e.g., data security [70], control [19], safety [43], ... e.g., easy to learn [64], usability problems [26], ...	Efficiency, Functionality, Ease of Use, Performance Usability, Utility, Security, Control, Learnability
Result	Outcome Expectation	e.g., task success [56], quality [45], product success [35], ... e.g., completeness [26], low error frequency [20], ...	Quality of Outcome, Error-free
Further Disciplines	Business User Social	e.g., money [63], brand [45], communication process [46], ... e.g., personalization [62], personification [39], ... e.g., social context [68], popularity [58], recommend [19], ...	Brand History, Advertisement, Price Expectation, Customization, Self-realization Group Affiliation, Social Connectivity
Environment	Temporal Context	e.g., time [19, 68], memorability [20], use frequency [26], ... e.g., device context [68], implicit interaction [30], ...	Memorability, Time Context, Location Context

Table 2. Clusters, sub-clusters, and consolidation of identified UX characteristics from the 30 selected publications.

sample reflects the perspectives of UX professionals from established firms as well as from emerging, digital startups in the fields of e.g. education, sports, finance, or smart home.

- *University Affiliations:* Aalborg University, University of Bristol, University of Lugano, New Jersey Institute of Technology, University of Oulu (2x), Queensland University of Technology (2x), University of Stuttgart, Tampere University of Technology (2x).
- *Industry Affiliations:* AirBnB, Allianz, GoCardless, Google, IICM, Nokia, Number26, Stylight, Tado, Talentry, Twitter.

Procedure. The first part of our interviews consisted of open questions about disciplines and departments involved in the product development process. In the second part, our experts were presented with our shortlist of the 28 UX characteristics and respective definitions. Participants were asked to complete this table by indicating the most relevant characteristics, reviewing our definitions, and linking them to responsible disciplines. On the basis of the interviews, we were able to narrow our list down to 9 relevant UX dimensions.

RESULTS

Below, we structure our findings into two interrelated sections. First, we propose a formalism to quantify UX based on our literature analysis and interviews. Second, we develop a corresponding tool to visualize QUX and to enhance communications within interdisciplinary teams.

Part 1: Quantifying UX (QUX)

We analyzed our expert interviews using a qualitative content analysis as proposed by [44], with a high inter-rater agreement (Cohen's Kappa $\kappa = .84$, see [10]). To start with, we presented our participants with a list of 28 UX characteristics and asked them to select the 10 they regarded as most important. This procedure allowed us to reduce the number

of relevant characteristics to 15, which are reported in Table 3. Next, to add more structure and balance, we decided to cluster our dimensions into the categories of *Look*, *Feel*, and *Usability* similar to [50].

Look	<i>n</i>
Aesthetics / Design	14
Interface	7
Brand History / Brand Name	5
Information Value	5
Advertisement / Brand	3
Feel	<i>n</i>
Control	13
Ease of Use	13
Learnability	12
Pleasure	12
Satisfaction	12
Usability	<i>n</i>
Efficiency	11
Utility	11
Effectiveness	10
Functionality	9
System Performance	8

Table 3. Top 5 UX characteristics per category.

Based on recommendations by our interviewed experts, we then merged some dimensions that were close to each other and/or partly overlapping. This way, we ended up with a total of nine relevant UX dimensions (three per category) as reported in Table 4. In a last step, we developed three corresponding items/questions per dimension (based on existing, pre-tested scales from [7], who provide multi-item measures for consumer insight research) to quantify a product's UX via answers on 7-point Likert scales.

Area	Dimension	Scales	ID	Related Work
Look	<i>Appealing Visual Design</i>	How balanced and harmonic do you find the product?	<i>d1</i>	[11, 20, 37, 61]
		Do you like the design, colors, fonts used in this product?		
		Do you find the text:image ratio appropriate?		
	<i>Communicated Information Structure</i>	Does the product provide clear navigation and orientation?	<i>d2</i>	
		How consistently is the content and information organized?		
		Do you find the provided information understandable?		
<i>Visual Branding</i>	Do you trust this brand?	<i>d3</i>		
	Do you think this is an honest brand?			
	Do you feel the brand is safe?			
Feel	<i>Mastery</i>	Do you find this product easy to use?	<i>d4</i>	[29, 36, 35, 50]
		Do you find it easy to learn (and to remember) how to use the product?		
		Do you feel you have full control over the product?		
	<i>Outcome Satisfaction</i>	How satisfied are you with the outcome?	<i>d5</i>	
		To what extent are you feeling successful with the outcome?		
		How happy are you with the outcome?		
<i>Emotional Attachment</i>	How pleasurable do you find using the product?	<i>d6</i>		
	Does the process of using the product provide you with gratification?			
	Do you feel excited when you are using the product?			
Usability	<i>Task Effectiveness</i>	Do you think the product does what it is supposed to do?	<i>d7</i>	[26, 29, 49, 57]
		Do you find the product effective?		
		Does the product help you fulfill your task?		
	<i>Task Efficiency</i>	Is the product the fastest way to achieve your goal?	<i>d8</i>	
		Is the product the most convenient way to achieve your goal?		
		Does using the product fit with your schedule?		
<i>Stability and Performance</i>	Does the system run smoothly?	<i>d9</i>		
	Are errors handled well?			
	Does the product work fast and responsively?			

Table 4. Interdisciplinary UX dimensions with corresponding questionnaire items and related work for in-depth, follow-up analyses.

Part 2: Visualizing and Communicating QUX

In addition, we asked all interview participants which disciplines should be involved in UX design processes. We identified the most relevant disciplines for each dimension using inductive category formation (see [44]). Besides HCI-related disciplines, such as Backend Development or Interaction Design, practitioners and researchers alike considered further disciplines, such as Marketing and Product Management, as highly relevant for the UX design process. Table 5 provides an overview of the top ten disciplines involved in the UX design process according to our interviewees. Furthermore, we asked all participants to link the respective disciplines to our list of 28 UX characteristics. Thus, we were able to assign responsibilities (i.e., disciplines) to our nine consolidated UX dimensions (see Figure 1).

Next, we were interested in practices and tools currently used to communicate goals and objectives in UX design processes. The majority of participants named meetings as the most important forum for discussing UX goals. Specific tools or visualizations are rarely used, whereas prototypes often serve as a basis to illustrate specific UX objectives. However, several UX practitioners described a kind of uncertainty when it comes to communicating UX within teams.

In sum, we were able to derive the following needs for our QUX approach from our interviews: (1) Combine measurement scales with a suitable visualization to enhance communication of concrete UX goals, (2) realize an easy-to-use application to support practitioners with different levels of expertise, and (3) consider the perspectives of different stakeholders as UX is multidisciplinary by nature. These needs are consistent with the findings of [65].

Top 10 disciplines involved in UX design process	<i>n</i>	Percentage
Backend Development	20	83%
Visual/Graphics Design	18	75%
Marketing	18	75%
Interaction Design	12	50%
Product Management	12	50%
User Research	10	42%
Usability Engineering/Testing	5	21%
UI/Frontend Development	5	21%
General Management	5	21%
Public Relations	3	13%

Table 5. Top 10 disciplines involved in the UX design process, sorted by number of occurrence in expert interviews (multiple responses possible).

Based on our identified needs we developed a graphical tool to measure, visualize, and communicate UX goals within interdisciplinary teams. The visual foundation of our QUX tool is a radar diagram with the categories *Look, Feel, and Usability* as focus areas. Next, we included the nine UX dimensions as well as the associated disciplines in accordance with the respective category.

The outer circle of the radar diagram connects our dimensions with the respective disciplines and illustrates the need for an interdisciplinary exchange. We designed our tool following a goal-oriented approach (see [21]). The UX-related disciplines are therefore centered around the nine UX dimensions (i.e., the UX goals) which represent the core of QUX.

As a final step, the 7-point Likert scale that is used to evaluate a product's UX based on the nine UX dimensions is illustrated as dotted circles. We use the questionnaire as indicated in Table 4 to quantify a product's UX and calculate average scores for each dimension (based on the associated scale). To illustrate a product's quantified UX, the scores for every dimension (be it as mean, median, or confidence interval) can be inserted in the radar diagram, linked, and visualized as a spanned plane (see Figure 1). Thus, development teams can easily detect weak spots in a product's UX and communicate further required actions, from product management over usability engineering to marketing. The basic idea of this visualization is similar to the UX wheel (see [42]).

EVALUATION

The goal of our evaluation is to judge the practical applicability of QUX within organizations. This is why we integrated our tool in the design processes of our industry partners and asked for their feedback. We chose to work with partners in the fields of sports, event ticketing, and food delivery to cover a certain range of B2C consumer applications in fast growing industries that increasingly focus on mobile apps. To bring the tool to life (i.e., discuss with professionals over real QUX scores rather than theoretical ideas), we asked our partners to collect exemplary survey data from their users via a Google form containing our 27 questions.

	Sports	Ticketing	Delivery
<i>n</i>	616	18	21
<i>Gender</i>	(m) 24% (f) 76%	(m) 67% (f) 33%	(m) 67% (f) 33%
<i>Age Range</i>	9 - 56 yrs	n/a	n/a
<i>Average</i>	29 yrs	n/a	n/a

Table 6. Demographic Data.

Sample Description

Our partner firms collected one large sample (n=616, by providing a lottery of high-end workout equipment as an incentive) and two smaller samples (n=18 and n=21, with no further incentive). Table 6 summarizes the demographic data. For all three samples, we computed Cronbach's α for each UX dimension. As shown in Table 7, alpha values range from .74 to .96, indicating consistently high construct reliabilities [12]. This indicates that each three items/questions we

developed from [7] seem to reliably measure the respective UX dimension derived from our methodology. We see this as a promising foundation for subsequent user-driven scale development and empirical studies, as discussed below in our section on future work.

UX Dimension	Sports	Ticketing	Delivery
<i>d1</i>	.87	.95	.75
<i>d2</i>	.89	.95	.81
<i>d3</i>	.90	.95	.92
<i>d4</i>	.88	.96	.86
<i>d5</i>	.91	.89	.93
<i>d6</i>	.93	.95	.86
<i>d7</i>	.88	.93	.85
<i>d8</i>	.84	.94	.92
<i>d9</i>	.88	.95	.74

Table 7. Cronbach's α by sample and dimension.

Exemplary QUX Analysis

We analyzed the data retrieved from survey respondents and visualized it using our QUX tool. Figure 1 shows an exemplary outcome for an application, which suffers from a rather poorly communicated information structure (users have problems with understanding the product's navigation and structure) and a lack of emotional attachment (users do not identify with the product, do not have any positive memories about the last use, etc.). Beyond this first diagnosis, our tool offers subsequent suggestions which departments or disciplines should be involved when conducting further in-depth UX evaluation. In this case, the *Visual/Graphics* department or *Frontend Developer* could initiate additional A/B-Testing to work towards a better information structure. Furthermore, *Product Management* might meet with *Marketing* to think about ways to improve emotional attachment of users (e.g., include animations or information that motivates recurring usage).

Qualitative Feedback from UX Professionals

We presented our findings at our partner companies to those responsible for UX (#1: a CTO, #2: a Vice President of Product and Design, and #3: a Senior Product Manager). Across all companies, our QUX tool received consistently positive feedback which falls into the following three categories.

Provides Overview and Helps to Prioritize

All our partners emphasized that the QUX tool provides a good starting point for thinking about UX: "The tool provides a useful overview of different aspects of UX. I must admit that I haven't had all of them on my radar yet."(#2) Furthermore, it "helps to identify strengths and weaknesses, which in turn helps us to prioritize our next steps in development."(#1) While the high-level overview was greatly appreciated, one product manager added that "it would be really helpful if you could provide us with some additional, qualitative tools to analyze our weak spots in more detail."(#3)

Allows for Benchmarking

Another key property of QUX seems to be its suitability for benchmarking: "For us, it would be highly interesting to conduct the same kind of analysis with our competitors' products to understand where we stand relative to them."(#3) One of

the partners can even imagine “using the tool to track user experience over time, so that we can track progress in our product development efforts.”(#1) To better judge the significance of the results, he proposed that in a revised version of QUX, we should also think about visualizing standard deviations/variances for each dimension.

Facilitates Communication in Teams

Our partners also emphasized the benefits of our visualization: “The radar diagram is a smart way of illustration. It helps to bring across the most relevant aspects at first glance. This visualization of UX provides a solid basis for deriving concrete actions.”(#3) Another partner expressed that he finds it helpful because he doesn’t “have to waste time and resources to prepare and visualize the data. So it really makes sense to agree on one single method, and stick to it.”(#2)

LIMITATIONS, FUTURE WORK, AND CONCLUSION

We see our approach as a first step towards achieving a more common view of UX within and beyond organizations. However, a unified measurement approach comes at the cost of potentially neglecting highly specific product details. While we acknowledge that every product is unique, we are convinced that QUX can be an important first step to obtain an overview and common understanding of a product’s UX. In this regard, QUX can be thought of as representing a manual for a toolbox rather than a tool itself. In future work, it might prove useful to not only link QUX findings to the associated disciplines, but to also use them for suggesting evaluation tools and methods for further in-depth analyses.

In our expert interviews, we learned once more that product development processes and respective UX paradigms are still dramatically different from one company to another. Yet, we believe that our approach can be valuable in similarly different ways. For example, early-stage startups might share a much more holistic view on their product and UX, but lack structured processes. Here, QUX can provide a meaningful guideline. With companies increasing in size and industry sectors maturing, the need for departmentalization and number of involved stakeholders is rising steadily. Here, QUX can facilitate efficient communications.

To showcase how QUX works in practice, we integrated it into the product development process of firms from three different industries. We found it encouraging that we received positive feedback across industry sectors. Still, a much broader sample drawn from a variety of products, services and sectors might hold many exciting insights. Our primary goal was to design a tool for measuring UX which is both building on and intended for industrial practice. This is why we deducted UX dimensions from expert interviews and then designed and qualitatively evaluated a tool by integrating it into the workflows of our industry partners.

We acknowledge that an alternative approach would have been a user-driven scale development through factor analysis-based dimensional reduction of large-scale surveys (as in, e.g., [4]). Our work aims at deriving needs from UX experts and practitioners, which we see as a foundation for complementary empirical work that focuses on the end-user side.

Overall, we see QUX as complementary to the landscape of existing UX evaluation methods and as a solid foundation for future work towards a common organizational understanding of UX.

ACKNOWLEDGMENTS

We would like to thank our interview partners and firms involved in the evaluation process, as well as all publications support and staff, who wrote and provided helpful comments on previous versions of this document. The authors gratefully acknowledge the grant provided by the German Federal Ministry for Education and Research (#01IS12057).

REFERENCES

1. All About UX. Information for user experience professionals, 2015. <http://www.allaboutux.org/>.
2. Alves, R., Valente, P., and Nunes, N. J. The State of User Experience Evaluation Practice. In *Proc. NordiCHI 2014*, ACM Press (2014), 93–102.
3. Bakker, S., Markopoulos, P., and de Kort, Y. OPOS: an Observation Scheme for Evaluating Head-Up Play. In *Proc. NordiCHI 2008*, ACM Press (2008), 18–22.
4. Bargas-Avila, J. A., and Brühlmann, F. Measuring user rated language quality: Development and validation of the user interface language quality survey (lqs). *International Journal of Human-Computer Studies* 86 (2016), 1–10.
5. Bargas-Avila, J. A., and Hornbæk, K. Old Wine in New Bottles or Novel Challenges? A Critical Analysis of Empirical Studies of User Experience. In *Proc. CHI 2011*, ACM Press (2011), 2689–2698.
6. Battarbee, K., and Koskinen, I. Co-experience: user experience as interaction. *CoDesign* 1, 1 (2005), 5–18.
7. Bruner, G. C., Hensel, P. J., and James, K. E. Marketing scales handbook, vol. 6. American Marketing Association (2012).
8. Burmester, M., Mast, M., Jäger, K., and Homans, H. Valence Method for Formative Evaluation of User Experience. In *Proc. DIS 2010* (2010), 364–367.
9. Cheng, S. “Third Eye”: Designing Eye Gaze Visualizations For Online Shopping Social Recommendations. In *Proc. CSCW 2013*, ACM Press (2013), 125–128.
10. Cohen, J. A coefficient of agreement for nominal scales. *Educational and psychological measurement* 20, 1 (1960), 37–46.
11. Cooper, A., Reimann, R., Cronin, D., and Noessel, C. *About Face. The Essentials of Interaction Design*, 4th ed. John Wiley & Sons, 2014.
12. Cronbach, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 3 (1951), 297–334.
13. Den Uyl, M. J., and van Kuilenburg, H. The FaceReader: Online facial expression recognition. In *Proc. Measuring Behavior 2005* (2005), 589–590.

14. Desmet, P. Measuring Emotion: Development and Application of an Instrument to Measure Emotional Responses to Products. In *Human-Computer Interaction Series Volume 3*. Kluwer Academic, 2004, ch. 9, 111–123.
15. Desmet, P., Overbeeke, K., and Tax, S. Designing products with added emotional value: development and application of an approach for research through design. *The Design Journal* 4, 1 (2001), 32–47.
16. Frandsen-Thorlacius, O., Hornbæk, K., Hertzum, M., and Clemmensen, T. Non-Universal Usability? A Survey of How Usability is Understood by Chinese and Danish Users. In *Proc. CHI 2009*, ACM (2009), 41–50.
17. Glaser, B. G., and Strauss, A. L. *The Discovery of Grounded Theory: Strategies for Qualitative Research*, vol. 36. Aldine Transaction, 1971.
18. Gray, C. M. Evolution of Design Competence in UX Practice. In *Proc. CHI 2014*, ACM Press (2014), 1645–1654.
19. Hart, J., Sutcliffe, a. G., and De Angeli, A. Love it or Hate it! Interactivity and User Types. In *Ext. Abstracts CHI 2013*, ACM Press (2013), 2059–2068.
20. Hartmann, J., Sutcliffe, A., and Angeli, A. D. Towards a theory of user judgment of aesthetics and user interface quality. *ACM Trans. Comput.-Hum. Interact.* 15, 4 (2008), 1–30.
21. Hassenzahl, M. User experience (UX): Towards an experiential perspective on product quality. In *Proc. IHM 2008*, ACM Press (2008), 11–15.
22. Hassenzahl, M., Burmester, M., and Koller, F. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Mensch & Computer 2003. Interaktion in Bewegung*, J. Ziegler and G. Szwillus, Eds. Teubner, 2003, 187–196.
23. Hassenzahl, M., and Tractinsky, N. User experience-a research agenda. *Behaviour & Information Technology* 25, 2 (2006), 91–97.
24. Helander, M. G., and Zhang, L. Field studies of comfort and discomfort in sitting. *Ergonomics* 40, 9 (1997), 895–915.
25. Hellman, M., and Rönkkö, K. Is User Experience Supported Effectively in Existing Software Development Processes? In *Proc. VUUM 2008*, Institute of Research in Informatics of Toulouse (2008), 32–37.
26. Hornbæk, K., and Law, E. L.-C. Meta-Analysis of Correlations Among Usability Measures. In *Ext. Abstracts CHI 2007*, ACM Press (2007), 617–626.
27. Iachello, G., Truong, K. N., Abowd, G. D., Hayes, G. R., and Stevens, M. Prototyping and Sampling Experience to Evaluate Ubiquitous Computing Privacy in the Real World. In *Proc. CHI 2006*, ACM Press (2006), 1009–1018.
28. Isabella, L. A. Evolving Interpretations as a Change Unfolds: How Managers Construe Key Organizational Events. *Academy of Management Journal* 33, 1 (1990), 7–41.
29. Jordan, P. W. Human factors for pleasure in product use. *Applied Ergonomics* 29, 1 (1998), 25–33.
30. Ju, W., Lee, B. A., and Klemmer, S. R. Range: Exploring Implicit Interaction through Electronic Whiteboard Design. In *Proc. CSCW 2008*, ACM Press (2008), 17–26.
31. Jung, Y., Anttila, A., and Blom, J. Designing for the Evolution of Mobile Contacts Application. In *Proc. MobileHCI 2008*, ACM Press (2008), 449–452.
32. Karapanos, E., Martens, J.-B., and Hassenzahl, M. On the Retrospective Assessment of Users' Experiences Over Time: Memory or Actuality? In *Ext. Abstracts CHI 2010*, ACM Press (2010), 4075–4080.
33. Karapanos, E., Zimmerman, J., Forlizzi, J., and Martens, J.-b. User Experience Over Time: An Initial Framework. In *Proc. CHI 2009*, ACM Press (2009), 729–738.
34. Karran, A. J., Fairclough, S. H., and Gilleade, K. A Framework for Psychophysiological Classification within a Cultural Heritage Context Using Interest. *ACM Trans. Comput.-Hum. Interact.* 21, 6 (2015).
35. Kujala, S., and Miron-Shatz, T. Emotions, experiences and usability in real-life mobile phone use. In *Proc. CHI 2013*, ACM Press (2013), 1061–1070.
36. Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E., and Sinnelä, A. UX Curve: A method for evaluating long-term user experience. *Interacting with Computers* 23, 5 (2011), 473–483.
37. Lavie, T., and Tractinsky, N. Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human Computer Studies* 60, 3 (2004), 269–298.
38. Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P., and Kort, J. Understanding, scoping and defining user experience. In *Proc. CHI 2009*, ACM Press (2009), 719–728.
39. Lee, J. M., Jeong, S. Y., and Ju, D. Y. Emotional Interaction and Notification of Flexible Handheld Devices. In *Ext. Abstracts CHI 2015* (2015), 2025–2030.
40. Liu, N., Liu, Y., and Wang, X. Data Logging plus E-diary: Towards an Online Evaluation Approach of Mobile Service Field Trial. In *Proc. MobileHCI 2010*, ACM Press (2010), 287–290.
41. Löchtefeld, M., Lautemann, N., Gehring, S., and Krüger, A. ambiPad Enriching Mobile Digital Media With Ambient Feedback. In *Proc. MobileHCI 2014*, ACM Press (2014), 295–298.
42. Magnus Revang, The User Experience Wheel, 2007. <http://userexperienceproject.blogspot.de/>.

43. Markova, M., Aula, A., Vainio, T., Wigelius, H., and Kulju, M. MoBiS-Q: a tool for evaluating the success of mobile business services. In *Proc. MobileHCI 2007*, ACM (2007), 238–245.
44. Mayring, P., and Fenzl, T. *Qualitative Content Analysis*. Springer, 2014.
45. Meschtscherjakov, A., Wilfinger, D., and Tscheligi, M. Mobile Attachment - Causes and Consequences for Emotional Bonding with Mobile Phones. In *Proc. CHI 2014*, ACM Press (2014), 2317–2326.
46. Milic-Frayling, N., Hicks, M., Jones, R., and Costello, J. On the Design and Evaluation of Web Augmented Mobile Applications. In *Proc. MobileHCI 2007* (2007), 226–233.
47. Mueller, F. F., Vetere, F., Gibbs, M., Edge, D., Agamanolis, S., and Sheridan, J. Jogging over a Distance between Europe and Australia. In *Proc. UIST 2014*, ACM Press (2010), 189–198.
48. Mugge, R., Schifferstein, H. N. J., and Schoormans, J. P. L. A Longitudinal Study of Product Attachment and its Determinants. *European Advances in Consumer Research* 7 (2006), 641–647.
49. Nielsen, J. Enhancing the Explanatory Power of Usability Heuristics. In *Proc. CHI 1994*, ACM Press (1994), 152–158.
50. Norman, D. A. *The Design of Everyday Things*. Basic Books, 2013.
51. Obrist, M., Beck, E., Wurhofer, D., and Tscheligi, M. Experience Characters: A Design Tool for Communicating Mobile Phone Experiences to Designers. In *Proc. MobileHCI 2011* (2011), 385–394.
52. Oliveira, R. D., Cherubini, M., and Oliver, N. Influence of Personality on Satisfaction with Mobile Phone Services. *ACM Trans. Comput.-Hum. Interact.* 20, 2 (2013), 1–23.
53. Olsson, T. Layers of User Expectations of Future Technologies: an Early Framework. In *Ext. Abstracts CHI 2014*, ACM Press (2014), 1957–1962.
54. Olsson, T., and Salo, M. Narratives of satisfying and unsatisfying experiences of current mobile augmented reality applications. In *Ext. Abstracts CHI 2012*, ACM Press (2012), 2779.
55. Pine, J., and Gilmore, J. H. Welcome to the Experience Economy. *Harvard Business Review* 76, 4 (1998), 97–105.
56. Rodden, K., Hutchinson, H., and Fu, X. Measuring the User Experience on a Large Scale: User-Centered Metrics for Web Applications. In *Proc. CHI 2010*, ACM Press (2010), 2395–2398.
57. Sauro, J., and Lewis, J. R. Correlations among Prototypical Usability Metrics: Evidence for the Construct of Usability. In *Proc. CHI 2009*, ACM Press (2009), 1609–1618.
58. Schulze, K., and Kroemker, H. Extracting User Experience Centered Product Requirements for Mobile Social Media Applications. In *Proc. MobileHCI 2012*, ACM Press (2012), 143–148.
59. Scollon, C. N., Kim-Prieto, C., and Diener, E. Experience Sampling: Promises and Pitfalls, Strengths and Weaknesses. *Journal of Happiness Studies* 4, 1 (2003), 5–34.
60. Sheldon, K. M., Elliot, A. J., Kim, Y., and Kasser, T. What Is Satisfying About Satisfying Events? *Journal of Personality and Social Psychology* 80, 2 (2001), 325–339.
61. Sonderegger, A., Uebelbacher, A., Pugliese, M., and Sauer, J. The influence of aesthetics in usability testing. In *Proc. CHI 2014*, ACM (2014), 21–30.
62. Srinivasan, V., Moghaddam, S., Mukherji, A., Rachuri, K. K., Xu, C., and Tapia, E. M. MobileMiner: Mining Your Frequent Patterns on Your Phone. In *Proc. UbiComp 2014*, ACM Press (2014), 389–400.
63. Tuch, A. N., and Hornbæk, K. Does Herzbergs Notion of Hygienes and Motivators Apply to User Experience? *ACM Trans. Comput.-Hum. Interact.* 22, 4 (2015).
64. Turunen, M., Melto, A., Hella, J., Heimonen, T., Hakulinen, J., Mäkinen, E., Laivo, T., and Soronen, H. User Expectations and User Experience with Different Modalities in a Mobile Phone Controlled Home Entertainment System. In *Proc. MobileHCI 2009*, ACM Press (2009), 1–4.
65. Väänänen-Vainio-Mattila, K., Roto, V., and Hassenzahl, M. Now Let's Do It in Practice: User Experience Evaluation Methods in Product Development. In *Ext. Abstracts CHI 2008*, ACM Press (2008), 3961–3964.
66. Väänänen-Vainio-Mattila, K., Roto, V., and Hassenzahl, M. Towards Practical User Experience Evaluation Methods. In *Proc. VUUM 2008*, Institute of Research in Informatics of Toulouse (2008), 19–22.
67. van Schaik, P., Hassenzahl, M., and Ling, J. User-Experience from an Inference Perspective. *ACM Trans. Comput.-Hum. Interact.* 19, 2 (2012), 1–25.
68. Vandenbroucke, K., Ferreira, D., Goncalves, J., Kostakos, V., and De Moor, K. Mobile Cloud Storage: A Contextual Experience. In *Proc. MobileHCI 2014*, ACM Press (Toronto, ON, Canada, 2014), 101–110.
69. Vermeeren, A. P., Law, E. L.-C., Roto, V., Obrist, M., Hoonhout, J., and Väänänen-Vainio-Mattila, K. User experience evaluation methods: current state and development needs. In *Proc. NordiCHI 2010*, ACM Press (2010), 521–530.
70. Vescovi, M., Bruno, L., Perentis, C., Moiso, C., and Leonardi, C. My Data Store: Toward User Awareness and Control on Personal Data. In *Proc. UbiComp 2014*, ACM Press (2014), 179–182.
71. Vuolle, M., Tiainen, M., Kallio, T., Vainio, T., Kulju, M., and Wigelius, H. Developing a Questionnaire for Measuring Mobile Business Service Experience. In *Proc. MobileHCI 2008*, ACM Press (2008), 53–62.

Triangulation in UX Studies: Learning from Experience

Ingrid Pettersson

Volvo Car Group
Sweden
ingrid.pettersson@
volvocars.com

Andreas Riemer

University of Applied Sciences
Ingolstadt (THI),
Germany
andreas.riemer@thi.de

Anna-Katharina Frison

University of Applied Sciences
Ingolstadt (THI),
Germany
anna-katharina.frison@thi.de

Jesper Nohage

Volvo Car Group
Sweden
jesper.nohage@
volvocars.com

Florian Lachner

University of Munich (LMU)
Germany
florian.lachner@ifi.lmu.de

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
Copyright is held by the owner/author(s).
DIS'17 Companion, June 10-14, 2017, Edinburgh, United Kingdom
ACM 978-1-4503-4991-8/17/06.
<http://dx.doi.org/10.1145/3064857.3064858>

Abstract

While the consideration of User Experience (UX) has become embedded in research and design processes, UX evaluation remains a challenging and strongly discussed area for both researchers in academia and practitioners in industry. A variety of different evaluation methods have been developed or adapted from related fields, building on identified methodology gaps. Although the importance of mixed methods and data-driven approaches to get well-founded study results of interactive systems has been emphasized numerous times, there is a lack of evolved understandings and recommendations of when and in which ways to combine different methods, theories, and data related to the UX of interactive systems. The workshop aims to gather experiences of user studies from UX professionals and academics to contribute to the knowledge of mixed methods, theories, and data in UX evaluation. We will discuss individual experiences, best practices, risks and gaps, and reveal correlations among triangulation strategies.

Author Keywords

User Experience; Evaluation; Mixed Methods; Triangulation

ACM Classification Keywords

H.5.2 [Information interfaces and presentation (e.g., HCI)]: Evaluation/methodology

Introduction

As an academic discipline, the field of User Experience (UX) research has a multi-disciplinary heritage, involving a variety of different perspectives that focused on studying human experiences with products, systems, and services. This led to a wide spectrum of methods that are used for studying users' experiences. Traditional Human-computer interaction (HCI) theory has passed on methodological approaches akin to those used in usability evaluation studies. Other disciplines that have significantly influenced UX research are those of social sciences, ethnography, and philosophy.

There have been great efforts in academia to create new methods for effectively evaluating UX, aimed at both academic and industrial application [1]. Our proposition in this workshop is, however, that we often do not need to develop new methods but rather use existing tools and approaches from the wide flora of UX evaluation more efficiently. UX evaluation is no longer an unknown territory and we want to encourage reflection on established approaches as well as lessons learned along the way. We want to explore the existing know-how of UX professionals, from academia and industry, in combining different UX evaluation methods (e.g., qualitative and quantitative methods) within so called mixed methods approaches and triangulation strategies.

Background & Motivation

Past workshops in the ACM community have already explored UX methods from different perspectives [4, 6, 3, 5]. However, a focus on triangulation, also called mixed methods, or multi-method approaches, is still missing. To combine different ways of research to get a more holistic view on UX is nowadays one of the key areas for further UX research [1, 4, 8]. Within a SIG session Roto et al. 2009 [4] analyzed UX evaluation methods in the industrial and

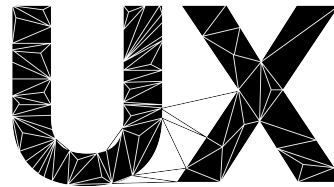


Figure 1: How can holistic User Experience (UX) evaluation be optimized by triangulation?

academic context. They revealed that rich data can be collected by applying mixed methods e.g., through the combination of system logging with subjective user statements from questionnaires and interviews. The authors conclude that mixing methods allows to understand the reasoning behind the concept of UX. Van Turnhout et al. [7] investigated common mixed research approaches of the NordiCHI proceedings 2012 to lay a foundation for further research and a more thoughtful application of multi-methods. However, best practices for using such multi-method perspectives inspired by the needs of academia and industry are not yet explored in depth.

Employing a mix of methods and theories to study a subject has been claimed to contribute to more reliable, holistic and well-motivated understandings of a phenomenon [2]. Furthermore, a mixed methods approach can uncover unexpected results, generate important and unforeseen research questions while at the same time providing answers to those new questions. This is particularly important for complex topics, such as the concept of UX. We argue that investigating UX design and evaluation from different angles will lead to a well-founded understanding of UX.

Workshop Theme & Goal

Researchers and practitioners have developed their own best practices over decades based on experiences, reflection, theoretic background, or intuition. We want to bring this wide-spread knowledge together and learn from each other by uncovering basic challenges, aims, and strategies related to UX work.

It will be an opportunity to share experiences with different UX evaluation methods, collect empirical data of practices, and a way to jointly suggest ways of improving the learning process from user studies. Finally, we want to support a more holistic understanding of the quality of a certain experience, which should be applicable for research projects in academia and industry. Specifically, we want to answer following questions:

- What are the motivations and the outcomes of different UX research and evaluation methods?
- How do we best draw conclusions from multiple and different sources, such as qualitative and quantitative or attitudinal and behavioral data?
- Can combinations of contrasting theories that exist in UX be better exploited, and if so how?
- How can we define best practices and where are gaps or development needs in mixed method approaches in the field of UX?

Duration

The presented theme and questions shall be discussed and edited in one full-day workshop.

Intended Outcome & Future Work

Our ambition is that the workshop will evolve and spread knowledge as well as awareness of how to get more out of UX studies. Consequently, participants will be able to apply particular methods more efficiently and effectively. A cooperatively developed mixed method map will summarize the outcomes. In combination with an already ongoing literature review on documented UX studies, the outcomes of the workshop will unfold the state of the art of using mixed method approaches in UX research. Further future work can be identified during the day and within the networking session.

REFERENCES

1. Javier A Bargas-Avila and Kasper Hornbæk. 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*. ACM, 2689–2698.
2. R Burke Johnson, Anthony J Onwuegbuzie, and Lisa A Turner. 2007. Toward a definition of mixed methods research. *Journal of mixed methods research* 1, 2 (2007), 112–133.
3. Marianna Obrist, Virpi Roto, and Kaisa Väänänen-Vainio-Mattila. 2009. User experience evaluation: do you know which method to use?. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2763–2766.
4. Virpi Roto, Marianna Obrist, and Kaisa Väänänen-Vainio-Mattila. 2009a. User experience evaluation methods in academic and industrial contexts. In *Proceedings of the Workshop UXEM*, Vol. 9. Citeseer.
5. Virpi Roto, Kaisa Väänänen-Vainio-Mattila, Effie Law, and Arnold Vermeeren. 2009b. User experience

- evaluation methods in product development (UXEM'09). In *IFIP Conference on Human-Computer Interaction*. Springer, 981–982.
6. Kaisa Väänänen-Vainio-Mattila, Virpi Roto, and Marc Hassenzahl. 2008. Now let's do it in practice: user experience evaluation methods in product development. In *CHI'08 extended abstracts on Human factors in computing systems*. ACM, 3961–3964.
 7. Koen van Turnhout, Arthur Bennis, Sabine Craenmehr, Robert Holwerda, Marjolein Jacobs, Ralph Niels, Lambert Zaad, Stijn Hoppenbrouwers, Dick Lenior, and René Bakker. 2014. Design patterns for mixed-method research in HCI. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. ACM, 361–370.
 8. Arnold POS Vermeeren, Effie Lai-Chong Law, Virpi Roto, Marianna Obrist, Jettie Hoonhout, and Kaisa Väänänen-Vainio-Mattila. 2010. User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*. ACM, 521–530.

UX Metrics: Deriving Country-Specific Usage Patterns of a Website Plug-In from Web Analytics

Florian Lachner^(), Florian Fincke, and Andreas Butz

LMU Munich, Human-Computer Interaction Group,
Amalienstr. 17, 80333 Munich, Germany
{florian.lachner,butz}@ifi.lmu.de, florian.fincke@campus.lmu.de

Abstract. Metrics for User Experience (UX) often involve traditional usability aspects, such as task success, but also mental aspects, such as interpretation and meaning. The actual experience of a user also highly depends on personal characteristics, such as the social and cultural background. In this paper, we investigate the relation between users' country of origin and their interaction patterns with an e-commerce website plug-in. We used a quantitative web analytics approach based on six UX-related metrics to evaluate the applicability of a quantitative UX evaluation approach in an international context. In a 34 day study we analyzed the usage patterns of 5.843 French, 2.760 German, and 5.548 Italian website visitors and found that they show significantly different patterns. This indicates that website metrics are a suitable means for cost-effective UX analysis on a large scale, which can provide valuable starting points for a further in-depth analysis.

Keywords: User experience · Cross-cultural design · User tracking · Data logging · Interfaces · Globalization · Localization

1 Introduction

The theory of User Experience (UX) goes back to the consideration of pleasure and emotions as part of a product's characteristics. Early approaches emerged from a user-centered design perspective, and the awareness of human factor professionals that user satisfaction is insufficiently considered in the concept of usability [26]. The consideration of pleasure and emotions was further increased by the focus on the interplay between affect and cognition. Due to this enhanced view on product design and development, aesthetics, pleasure, and usability became a balanced triad in the HCI community [40].

Nowadays, the primary goal of UX designers and engineers often is to create a pleasurable interaction between the user and the product that goes beyond traditional usability considerations [19]. It also has become common ground in the HCI community that experiences are subjective in nature and highly dependent

© IFIP International Federation for Information Processing 2017
Published by Springer International Publishing AG 2017. All Rights Reserved
R. Bernhaupt et al. (Eds.): INTERACT 2017, Part III, LNCS 10515, pp. 142–159, 2017.
DOI: 10.1007/978-3-319-67687-6_11

on the usage context [24,32]. Hence, a user’s experiences can be shaped and influenced based on his or her individual preferences (regarding aesthetics or ergonomics), mood, prior interactions, product brand, age, gender, and culture [7,12,16,29,30,34,42,51,52]. The cultural aspect becomes particularly interesting for global businesses, whose products or services can be accessed, evaluated, and purchased from all over the world [17,37,46,61].

In order to ensure the intended quality of UX, measurement tools and methods represent a crucial resource in UX design and research processes. However, there is still an ongoing debate about the applicability and effectiveness of qualitative and quantitative approaches for UX measurement [6,33,57]. Furthermore, researchers and designers have to balance information value, cost efficiency, and expenditure of time when gathering attitudinal (e.g., through lab studies or surveys) or behavioral data (e.g., through data logging or time measurement) [50,55,56].

In this paper, we analyze the relationship between the country of origin and the usage behavior of users of a website plug-in (see Fig. 1). We base our analysis on quantitative behavioral data, gathered through user tracking, to draw a conclusion on the applicability of web analytics metrics. Our dataset stems from a data logging study of a plug-in that was implemented in an e-commerce website plug-in.

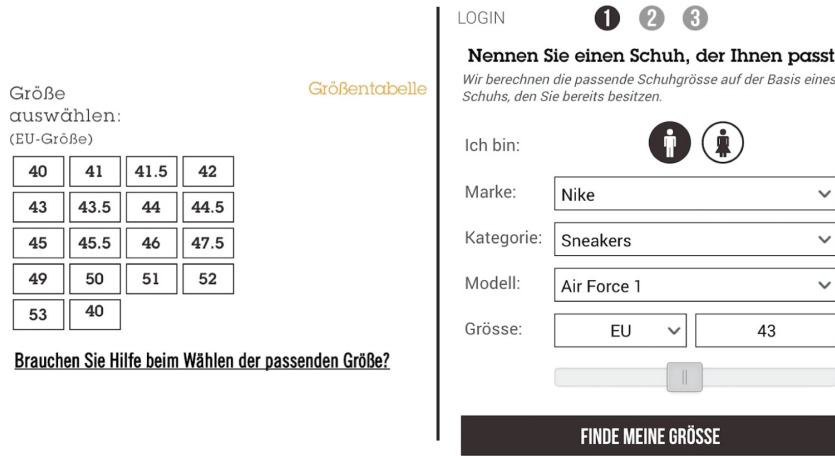


Fig. 1. Website plug-in (right) for shoe size recommendations and the link to it in the German online store (left).

Over the course of 34 consecutive days we tracked the behavior (i.e., plug-in interactions) of users located in France, Germany, and Italy based on six UX metrics, which we derived and adapted from the HEART framework of Rodden et al. [50]. Our study was motivated by the following research question:

Which differences in the user experience of a website (plug-in) can we identify between French, German, and Italian users simply through web analytics metrics?

Consequently, the contribution of this paper is twofold: First, based on the analysis of country-specific differences we identify associated relationships and hence suitable levers to efficiently target further qualitative in-depth analyses. Second, we adapt the quantitative UX framework of Rodden et al. [50] to our specific use case (i.e., e-commerce website plug-in) in order to examine the applicability of UX metrics that build upon large-scale website tracking data. Ultimately, we draw a conclusion how such a quantitative approach can support designers in saving time and money for cross-cultural UX evaluation and potentially localized interface adaptations. For our analysis, we, therefore, exclude a supplemental investigation of further factors, such as gender differences or device type. Our underlying goal is to foster an ongoing debate about cross-cultural UX design and about an appropriate balance of qualitative and quantitative UX measurement.

2 Background and Related Work

Despite the general agreement on its importance for human-centered design, researchers and practitioners still struggle to narrow down the broad field of UX to one unified definition [20]. The lack of a common definition of UX entails a large variety of research directions in the field of HCI, with foci ranging from usability to psychological needs and emotions [6, 32]. To locate our work in this ongoing discourse, the following sections illustrate the basic scope of (our understanding of) UX, some key aspects of UX measurement, as well as related work in the field of cross-cultural UX design.

2.1 The Scope of User Experience

The main difference between usability and UX is that UX researchers and designers can not merely focus on a product's characteristics (i.e., functionality, purpose, etc.) but also have to consider the user's needs and motivation as well as the context of use (i.e., the environment) [9, 13, 19, 24, 38]. Consequently, experiences do not only result from interacting with a product but also from a user's expectations, others' opinions, or from experiences with related technologies before the actual interaction. At the same time, experiences and associated feelings merely evolve over time through reflection on previous interactions, advertisements, and again through others' opinions [24, 27, 35, 52].

The scope of UX becomes even more complex for globally acting businesses: First, the concept of UX is differently understood between academia and industry as well as between different countries [31, 32]. Second, cultural differences in language, values, or needs raise various questions regarding the suitability of globally optimal or locally adapted designs of products and services [4, 37, 59].

In this paper, our goal is to analyze cultural differences in UX design. For this purpose, we simplify the origin of cultural differences to individual preferences caused by one's country of origin. Thus, we do not focus on further cultural allocations, such as age group or social background. Furthermore, UX in our context shall include both usability aspects as well as mental aspects, such as the interpretation of an e-commerce website plug-in. In order to answer our research question, we will, therefore, derive suitable web analytics metrics, which we call UX metrics.

2.2 Cross-Cultural Differences in UX Design

The need for cross-cultural considerations in interface design emerged more than two centuries ago, shortly after designers started to put an emphasis on the usability aspects of their designs (see [39]). Initial discussions mainly focused on the use of colors, language, as well as icons and symbols [5, 53]. However, since then usability theories and measures in the HCI community rather marginally focused on cultural design preferences [14]. Nevertheless, with the further increasing interest in experiences of product interactions, researchers in the HCI community once again started to raise questions about cross-cultural design preferences (see [8, 21, 47]). In fact, various studies have already been able to identify cultural differences in UX design in different use cases.

Athinen et al. [1], for example, investigated culturally sensitive design for a mobile wellness application. In their study, they interviewed 16 people (8 from Finland and 8 from India) to identify similarities and differences in the understanding of wellness and its consequences for the design of a mobile application. They found that Finns and Indians have a different understanding of goal setting, which is an important aspect for the associated mobile application. Similarly, Walsh and Vaino [60] argue for cross-cultural UX considerations for mHealth applications, while Al-Shamaileh and Sutcliffe [2] demonstrate varying preferences in the design of health-related websites in the UK and Jordan.

Furthermore, Frandsen-Thorlacius et al. [14] were able to detect differences in the understanding of the concept of usability for Danish and Chinese users. Using a questionnaire survey, the authors were able to derive that Chinese users preferentially value visual appearance, satisfaction, and fun, whereas Danish users rather focus on effectiveness, lack of frustration, and efficiency. Reinecke and Gajos [48] were, likewise, able to analyze visual preferences of websites based on a comprehensive study of 2.4 million ratings from almost 40 thousand participants.

However, cultural differences are not limited to the evaluation of products and services. Lallemand et al. [31] point out discrepancies in the understanding of the concept of UX based on a survey amongst 758 researchers and practitioners from 35 nationalities. Gereá and Herskovic [15] additionally expand this study to Latin America. Nowadays, researchers want to further link cultural studies and product design, particularly through the integration of Hofstede's (see [22]) cultural dimensions in HCI [36, 37, 43, 46, 58].

2.3 Qualitative Vs. Quantitative UX Measurement

Because experiences are such a complex phenomenon, UX researchers and practitioners utilize a whole set of measurement approaches to anticipate, test, and improve a product's UX. However, there is no common agreement whether qualitative or quantitative approaches should be favored [6, 33, 49]. On the one hand, qualitative approaches (gathered through, e.g., interviews) provide rich and detailed insights for in-depth analysis [54], on the other hand, quantitative approaches (gathered through, e.g., questionnaires) can reduce costs and time effort [23, 57].

Apart from this, UX measurement methods are primarily based on attitudinal data (i.e., data related to a user's feelings and emotions) [31, 50]. In contrast, the HEART framework [50] represents a first step towards the integration of behavioral data (i.e., actual activities of users - traditionally used in usability testing, see [3, 10, 25, 41]), in UX measurement. The framework includes five metrics, focusing on both usability and UX-related aspects [50]:

- *Happiness*: referring to, e.g., satisfaction and ease of use.
- *Engagement*: describing the user's level of involvement.
- *Adoption*: addressing customer acquisition.
- *Retention*: analyzing recurring users.
- *Task success*: covering traditional usability aspects.

The framework does not aim to describe UX as a whole but to strategically direct UX measurement processes based on large-scale data, particularly when working in teams. Therefore, one has to define a suitable measurement goal and approach per metric (e.g., the number of visits per week for *Engagement*, the error rate for *Task success*) depending on the respective product or service.

We understand their approach as an initial step towards including behavioral data from usability testing in UX measurement. Therefore, we aim to evaluate its applicability for our use case, i.e., the analysis of UX-related, country-specific usage patterns of French, German, and Italian users from web analytics. However, to ensure a suitable implementation of UX metrics in our collaboration partner's development process, we slightly customized our UX metrics based on the HEART framework.

3 Methodology and User Study

In order to examine the applicability of UX-oriented web analytics metrics for identifying country-specific user behaviors, we partnered with a company that provides a customizable website plug-in for online shoe stores. The plug-in allows customers to identify their correct shoe size based on the comparison with the size of another model.

3.1 Setting and Procedure

For our study, we tracked the plug-in interactions of a globally acting online shoe store. The analyzed plug-in (see Fig. 1 right) is integrated in the store’s website and accessible through a link below the actual selection of the shoe size (see Fig. 1 left). The overall goal of the plug-in is that customers can enter information about a shoe that they already own in order to identify the correct size of the shoe they want to buy. To ensure a problem-free implementation in different countries, the plug-in was translated by professional translators for all countries.

Once a customer clicks on the link, the plug-in opens and asks for the customer’s gender as well as the brand, category, model, and size of a comparative shoe (i.e., plug-in steps one to five). This information is used to identify the correct size for the customer depending on the shape and differences in size of the desired shoe. The comparative data is taken from our partner’s internal database. As a sixth plug-in step, users can request (i.e., click) a shoe size recommendation. After receiving all the information, the recommended size is stored for 90 days and additionally displayed within the link’s text label once a customer accesses the online store again. Thus, it is not necessary to open and use the plug-in repeatedly.

For post-hoc analysis, all tracked data points (plug-in openings, plug-in interactions, recommendation requests, and adding products to the website shopping cart) were anonymized and securely stored at our partner’s server infrastructure for long-term evaluations through client-based tracking. Client-based tracking (i.e., Javascript-based for plug-in interactions and cookie-based tracking for long-term analysis of recurring users) was pursued to minimize data traffic in order to ensure a smooth and pleasant plug-in implementation. Shoe recommendations were tracked through server-based tracking. The country of origin was identified by the client’s IP address.

3.2 Study Data and Analysis

Observations. We ran our study for 34 consecutive days. During this time, no special offer or promotion was announced at the client’s online store in order to ensure the comparability of our analysis. Over the course of our study people from 200 countries visited the client’s website, whereof people from 121 countries accessed the plug-in. For our investigation we focused on France, Germany, and Italy (277,551, 141,897, and 172,887 website loadings leading to 5843, 2760, and 5548 plug-in openings, respectively). Overall, about one third (31,4% in France, 30,4% in Germany, and 37,2% in Italy) of all website visitors per country accessed the website on a mobile device, two thirds (68,6% in France, 69,6% in Germany, and 62,8% in Italy) on a desktop device.

UX metrics. Our quantitative analysis of the plug-in interactions was based on six metrics (see Table 1) that we derived and adapted from the HEART framework [50]. Our metrics were consciously labeled with a distinguishing term in

order to highlight the objective of each metric. Furthermore, the particular term allowed our collaboration partner to align strategic initiatives and development efforts.

Table 1. UX metrics used for the analysis of plug-in interactions.

UX metric	Definition and objective	see HEART [50]
Adoption	No. of openings (link clicks) to measure user acquisition	<i>Adoption</i>
Complexity	Time per data input to analyze complexity per plug-in step	<i>Engagement</i>
Task success	No. of total recommendations to track plug-in effectiveness	<i>Task Success</i>
Continuity	No. of successful inputs per step to retrace plug-in continuity	<i>Task Success</i>
Trust	No. of recommended orders to derive trust in suggestions	<i>Happiness</i>
Mastery	No. of suggested orders without plug-in opening (recurring users) to derive long-term trust	<i>Retention</i>

First of all, we tracked the user **Adoption**, i.e., the number of users that click on the link to the plug-in as well as the **Complexity** of the plug-in (based on the process time per plug-in step). In order to analyze the effectiveness of the plug-in, we defined the two metrics, **Task Success** (overall number of final recommendations) and **Continuity** (successful completions per plug-in step). These four metrics describe usability aspects of the plug-in.

For the interaction with the online shoe store plug-in, we wanted the associated UX to be a pleasant interaction with the service that results in a trustworthy shoe size recommendation. The goal of the plug-in recommendation, therefore, is that customers identify the correct size of a shoe and trust the plug-in even when the recommendation differs from the size of the comparative shoe. An additional feature of the plug-in is that the recommended shoe size is stored and shown in the plug-in link when users complete all plug-in steps and access the website again within 90 days (see Fig. 2).

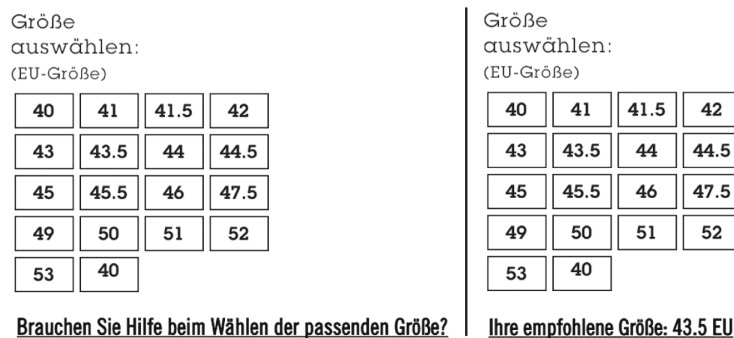


Fig. 2. Link (in the German online store) to the plug-in without recommendation (left) and with recommendation for recurring users (right).

Against this background, we defined the metric **Trust** to understand if users rely on the shoe size recommendation of the plug-in (i.e., put the recommended shoe size into the website's shopping cart). We, therefore, only considered users who ordered a recommended shoe size that differed from the initially entered size of the comparative shoe and excluded users whose recommended size corresponded to the size of the selected comparative shoe. Thus, we could evaluate if users clearly relied on the plug-in's recommendation. We adapted the metric happiness from the HEART framework to our use case as it was not desired to establish a direct communication with the user. All users who successfully clicked through all steps received a recommendation whereas we defined a pleased user as a user that relied on the recommended size for his/her final order. In order to draw conclusions on the long-term experience with the recommendation service, we defined the metric **Mastery**. This metric refers to the number of orders (of recommended shoe sizes) from recurring users that did not open the plug-in again but relied on the suggestion of a suitable size based on their previously entered information. The information was stored in a client-side cookie for 90 days as described before. For this purpose, the recommended shoe size was shown in the link's text label. Once again, we only considered orders that included differing shoe sizes.

All in all, we see these metrics as suitable measuring points for the UX evaluation of equivalent recommendation plug-ins (with the objective to minimize recurring interactions) in an e-commerce context. In further use cases, researchers and designers will have to question their generalization and adapt the metrics accordingly (e.g., when a repetitious interaction is aspired).

Data analysis. We conducted a statistical analysis (using SPSS version 20.0) to identify varying usage behaviors between French, German, and Italian users. We used the Chi-Square Test in order to analyze the association between the country and the UX metrics of *Adoption*, *Task Success*, *Continuity*, *Trust*, and *Mastery*. In order to evaluate the UX metric *Complexity* we used two-way ANOVA and post-hoc Sidak as well as an ANOVA test. We excluded outliers in the process times for the analysis of the metric *Complexity* according to Grubbs [18]. An identified outlier was also excluded from the analysis of previous plug-in steps to ensure consistency within our results. For all analyses we defined a significance level of 5%.

4 Results

The analysis of our data set using the previously defined UX metrics yielded a number of differences in the usage behaviours of the website plug-in between French, German, and Italian users. Thus, we were able to derive significant differences in the adoption rate, dropout rate per plug-in step, the temporal usage patterns, and the reliance on recommendations as described below.

4.1 Country-Specific Adoption, Dropout, and Recommendation Rate

First of all, it should be noted that we found a relationship between the country of origin and the *Adoption* rate (see Table 2), i.e., number of plug-in openings to measure user acquisition ($\chi^2(2) = 714.327$, $p = .000$): 2.11% for French users (277,551 website loadings, 5,843 openings), 1.95% for German users (141,897 website loadings, 2,760 openings), and 3.21% for Italian users (172,887 website loadings, 5,548 openings).

The analysis of *Continuity* (i.e., number of successful inputs per plug-in step to retrace usage continuity) provided insights in the relationship between country of origin and successful completions per plug-in step. We found a relationship for the plug-in steps where users had to select their gender ($\chi^2(4) = 28.267$, $p = .000$), the brand of a comparative shoe ($\chi^2(4) = 10.166$, $p = .038$), an associated model ($\chi^2(4) = 22.019$, $p = .000$), and click to receive a shoe size recommendation ($\chi^2(2) = 6.781$, $p = .034$), as summarized in Table 2.

Except for the last step, where users had to click to receive a recommendation, we included users who successfully completed the respective step (success), closed the plug-in or browser (failure), and users who went back to the respective plug-in step after having already moved on to further plug-in steps (detour) in the analysis of the usage *Continuity*. Thus, we were able to derive usage patterns per plug-in step: Generally, in the first plug-in step (i.e., selection of the gender) users from all three countries showed the highest dropout rate (including only successful and failed completions): 22.51% for France (1315 failed users), 24.82% for Germany (685 failed users), and 26.71% for Italy (1482 failed users). In addition, most users who went back to a previous plug-in step chose to start from the beginning of entering the comparative data, more precisely by selecting the brand (the second plug-in step) of a comparative shoe (see Table 2).

In addition, we were able to identify a relationship of *Task Success* (i.e., number of total recommendations to understand plug-in effectiveness) and country of origin ($\chi^2(2) = 13.332$, $p = .001$). Users from France showed the highest rate of successful recommendations (52% out of 5843 plug-in openings), followed by Germany (49% out of 2760 plug-in openings), and Italy (46% out of 5548 plug-in openings).

4.2 Divergent Temporal Usage Patterns

The goal of the metric *Complexity* was to identify temporal differences along the process steps in order to diagnose key hurdles of the plug-in. We used the z-score transformation to make the data normal before conducting the (two-way) ANOVA and post-hoc Sidak test, as our dataset (process time per plug-in step) did not represent a normal distribution according to the Kolmogorov-Smirnov test. We used post-hoc Sidak test as all users interacted with the plug-in independently. We then used the two-way ANOVA and a post-hoc Sidak test to analyze the effect of the country of origin on the time spent on each step along the plug-in process. Thus, we found out that there was an effect between country of origin and the process time per plug-in step ($F(10,2) = 10.427$, $p = .000$,

Table 2. Chi-Square results (χ^2) based on the UX metrics Adoption, Continuity, Task Success, Trust, and Mastery for French (FRA), German (GER), and Italian (ITA) users.

UX metric	Plug-In Step	Country of Origin				χ^2	Cramer's V	
		<i>FRA</i>	<i>GER</i>	<i>ITA</i>	<i>Total</i>			
Adoption	<i>Success</i>	5843	2760	5548	14.151	714.327	.035	
	<i>Failure</i>	271.708	139.137	167.339	578.184			
	<i>Total</i>	277.551	141.897	172.887	592.335			p = .000
Continuity	Gender	<i>Success</i>	4528	2075	4066	10,669	28.267	.032
		<i>Failure</i>	1315	685	1482	3482		
		<i>Detour</i>	14	8	19	41		
		<i>Total</i>	5857	2768	5567	14,192		
	Brand	<i>Success</i>	4105	1917	3733	9755	10.166	.022
		<i>Failure</i>	183	72	134	389		
		<i>Detour</i>	320	120	308	748		
		<i>Total</i>	4608	2109	4175	10.892		
	Category	<i>Success</i>	3908	1797	3540	9245	4.218	n.s.
		<i>Failure</i>	312	155	283	750		
		<i>Detour</i>	178	87	196	461		
		<i>Total</i>	4398	2039	4019	10.456		
Model	<i>Success</i>	3354	1516	2906	7776	22.019	.034	
	<i>Failure</i>	664	325	728	1717			
	<i>Detour</i>	27	18	44	89			
	<i>Total</i>	4045	1859	3678	9582			p = .000
Size	<i>Success</i>	3159	1417	2698	7274	8.256	n.s.	
	<i>Failure</i>	213	108	228	549			
	<i>Detour</i>	6	4	12	22			
	<i>Total</i>	3378	1529	2938	7845			p = .083
Rec.	<i>Success</i>	3038	1350	2560	6948	6.781	.031	
	<i>Failure</i>	125	70	145	340			
	<i>Total</i>	3163	1420	2705	7288			p = .034
Task success	<i>Openings</i>	5843	2760	5548	14.151	13.332	.025	
	<i>Rec.</i>	3038	1350	2560	6948			
	<i>Total</i>	8881	4110	8108	21.099			p = .001
Trust	<i>Yes</i>	10	3	9	22	21.232	.193	
	<i>No</i>	381	113	53	547			
	<i>Total</i>	391	116	62	569			p = .000
Mastery	<i>Yes</i>	158	17	103	278	42.130	.136	
	<i>No</i>	1094	421	500	2015			
	<i>Total</i>	1252	438	603	2293			p = .000

$\eta^2 = .011$). In our study, Italian users significantly differed in their temporal usage patterns along all plug-in steps from French users ($p = .000$) as well as from German users ($p = .022$). French and Germany did not differ significantly ($p > .050$) (see Fig. 3).

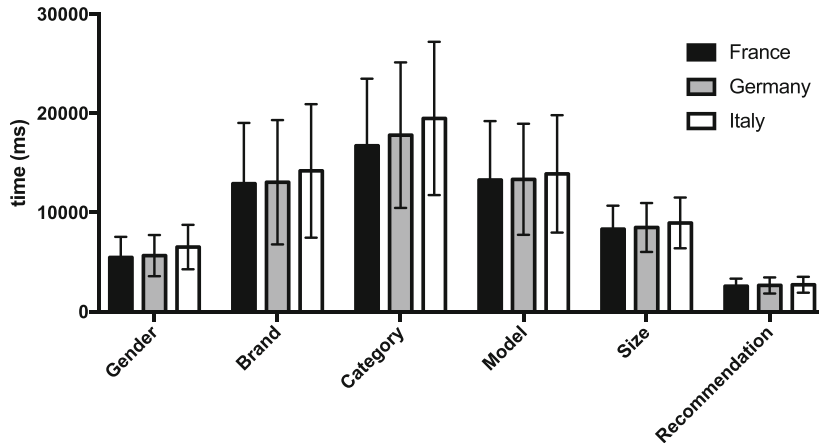


Fig. 3. Average process times for each plug-in step per country including the standard deviation.

Next, we conducted an ANOVA and post-hoc Sidak test to investigate whether the process times significantly vary per plug-in step. Thus, we found out that there is an effect of country of origin for the first plug-in step to select the gender ($F(12,2) = 10.774$, $p = .000$, $\eta^2 = .012$) as well as the third plug-in step to select a category of a comparative shoe ($F(12,2) = 6.342$, $p = .002$, $\eta^2 = .007$). For the other plug-in steps (i.e., brand, model, size, and recommendation) we could not identify significant differences ($p > .050$). More precisely, for the first step (i.e., gender) the process time (i.e., the mean) of Italian users differed from French users ($p = .000$) as well as from German users ($p = .008$). Furthermore, the mean of the process time of Italian users to select a category varied from the process time of French users ($p = 0.001$). On average, Italian users needed more time for each plug-in step.

4.3 Varying Reliance on Recommendations

Based on the two metrics *Trust* (number of recommended orders) and *Mastery* (number of recommended orders of recurring users without opening the plug-in) we analyzed the usage behaviors of plug-in users directly related to the recommendation service. The objective of these metrics is to understand whether the country of origin is related to the reliance of users on the shoe size recommendation as well as with the understanding of recurring users (who already successfully clicked through the whole plug-in process and should understand

that their suitable size is directly represented in the plug-in opening link) that they do not have to open the plug-in again.

We found out that there is a relationship between country of origin and the *Trust* in the recommendation of the plug-in ($\chi^2(2) = 13.983, p = .001$). Furthermore, the country of origin is related to the understanding of the link’s text label recommendation (*Mastery*) for recurring users ($\chi^2(2) = 42.130, p = .000$).

In our study, French and German users showed a comparable trust rate (i.e., number of users who ordered a differing shoe size based on the recommendation and excluding users whose initially entered shoe size equalled the recommended size hence no conclusion on the user’s trust can be drawn) of 2.56% (FR: 10 out of 391, 114 additional users excluded) and 2.59% (GER: 3 out of 116, 60 additional users excluded). However, from 62 Italian users that got a differing recommendation, 9 users (14.52% excluding 27 additional users) relied on the plug-in and added a differing shoe size into the website’s shopping cart (see Fig. 4).

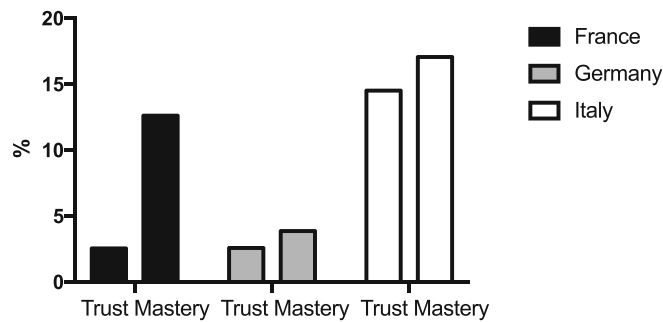


Fig. 4. Percentage of users per country who relied on the plug-in suggestion (Trust) or the suggestion in the link’s text label (Mastery) when the recommended size differed from the comparative shoe size.

With regards to recurring users relying on the recommendation of the plug-in link, the number of French users (adding a differing shoe size into the website’s shopping cart without opening the plug-in and once again excluding users for whom no conclusion can be drawn as the recommended size equalled the initially entered size) increased to 12.62% (158 out of 1252, excluding 486 additional users) and the number of reliant Italian users increased to 17.08% (103 out of 603, excluding 133 additional users). German users, however, remained at a rather low rate of 3.88% (17 out of 438, excluding 250 additional users) (see Fig. 4).

4.4 Summary and Interpretation

The analysis of UX metrics allowed us to understand country-specific usage patterns of French, German, and Italian users. Users from all three countries

showed distinct adoption and dropout rates as well as, in particular, significant associations with the plug-in steps gender, brand, model, and recommendation. In addition, we identified significant differences in the overall task success rates. Furthermore, the country of origin is related to the temporal usage patterns along the plug-in steps, with Italian users being the slowest.

Finally, the analysis of the UX-focused metrics *Trust* and *Mastery* showed lower rates for French and German users compared to users from Italy. However, recurring users from France strongly increased their long-term trust rate (i.e., *Mastery*) due to the suggestion in the plug-in link's text label. The described UX metrics helped our project partner to efficiently focus on selected plug-in steps as the analysis highlighted country-specific relationships with a low effect size that are worth paying attention (Cramer's V between .10 and .20 and $\eta^2=.01$) compared to country-specific relationships with a marginal effect size (Cramer's V between .00 and .10 and $\eta^2 <.01$) according to Rea & Parker [45] and Cohen [11]. The localization of all plug-in steps will increase development time and costs. Through the focus on selected and significant plug-in steps with at least low effect sizes, our project partner was able to allocate research and development resources more efficiently.

In order to identify localized interfaces for different countries, designers and researchers need to analyze suitable aspects in further in-depth studies. First, the interface of the website plug-in can be localized and evaluated recurrently for each country to minimize the dropout rate for the critical plug-in steps. One might, for example, prefer text-based icons for the selection of the gender. Second, with regards to the differing process times the plug-in design can be complemented with additional information in order to balance process times per step, dropout rate, and backward steps. Third, it is important to investigate the differences in the *Trust* and *Mastery* rate. German users, for example, might not want to receive suggestions within the link's text label but prefer to receive an individual recommendation each time. Thus, the overall plug-in and link design should be rearranged. Therefore, further qualitative in-depth investigations in the future will allow us to clarify our interpretations.

5 Conclusion, Limitation, and Future Work

In this paper, we demonstrated the applicability of web analytics metrics to analyze differences in the usage behavior and UX of an e-commerce website plug-in between French, German, and Italian users. We were able to identify significant relationships between the country of origin and the adoption rate as well as dropout rate of several plug-in steps. In addition, users from France, Germany, and Italy showed different temporal usage patterns as well as trust in the plug-in's recommendation. Although our work focused on the analysis of an e-commerce plug-in, further country-specific usage patterns have already been identified for Q&A websites such as Yahoo Answers (see [28]) as well as StackOverflow and Superuser (see [44]).

However, narrowing down the complex scope of UX to a selection of six customized website analytics metrics based on client-side user tracking can only

be a first step. Overall, it will be necessary to further investigate and analyze the applicability of user tracking for UX measurement due to its quantitatively descriptive nature (see [24,32,41]). Inspired by traditional usability approaches (i.e., logging data) we see our work as a starting point to efficiently guide in-depth UX analyses, complementary to qualitative evaluations with a focus on attitudinal data. Additionally, client-based tracking might not holistically track all website visitors due to, e.g., blocked website cookies. It is, by nature, not possible to track how many website visitors block cookies. We, therefore, limited our analysis to recurring users of plug-in interactions and not website visits. Furthermore, the collaboration with our industry partner did not allow any modification of the original website. Consequently, it was not possible to add a registration process to track the user behavior across different devices.

Based on our research, future studies should add further metrics and qualitative in-depth analysis of country-specific usage patterns, test our findings through locally adapted user interface studies, and investigate the impact of server-based tracking on both the users' UX and the validity of web analytics metrics. Furthermore, the investigation of user-level data (i.e., the consolidated usage data of individual users) might allow conclusions about more detailed user behaviors. Ultimately, to set up a holistic UX-focused user tracking process, it is necessary to compare the effect of cross-country differences with and in contrast to further aspects, such as gender and device type.

We conclude that user tracking can be an efficient way to identify UX-related levers for culturally sensitive design adaptations of website plug-ins. At the same time, we agree with Vermeeren et al. [57] and Law et al. [33] that an exclusive focus on quantitative UX measurement (through, e.g., web analytics metrics) might ignore relevant insights of qualitative measurement approaches. Consequently a balance of various measurement tools and approaches should be promoted. In culturally sensitive development processes, the research and design team can implement UX-focused user tracking to identify suitable levers for country-specific design adaptations. Once significant differences in the usage behaviors for certain steps of a website plug-in have been identified, researchers and developers can, e.g., efficiently set-up subsequent A/B-tests and investigate the impact on the click behavior for different designs. This includes but is not limited to more or less information for such plug-in steps, different designs (colours, fonts, etc.) or simply a different user flow through the plug-in. Changes in the design can then be analyzed through further user tracking and supplemental qualitative evaluations.

In summary, our work was guided by the motivation to pursue a quantitative approach based on web analytics metrics to identify UX-related, country-specific usage behaviors of a website plug-in. We aim to foster an ongoing discussion about cross-cultural UX design as well as a suitable balance between qualitative and quantitative UX measurement - following up on the investigation of large-scale behavioral data. In particular, however, we want to emphasize that the challenging need of globally acting companies to analyze country-specific preferences and usage patterns requires cost-efficient and quickly adaptable UX

measurement tools. In this light, we perceive our work as a constructive starting point for further cross-cultural investigations based on large-scale behavioral data.

References

1. Ahtinen, A., Ramiah, S., Blom, J., Isomursu, M.: Design of mobile wellness applications: identifying cross-cultural factors. In: Proceedings of the 20th Australasian Conference on Computer-Human Interaction (OZCHI), pp. 164–171. ACM Press (2008)
2. Al-Shamaileh, O., Sutcliffe, A.: Investigating a multi-faceted view of user experience. In: Proceedings of the 24th Australian Computer-Human Interaction Conference (OZCHI), pp. 9–18. ACM Press (2012)
3. Andreasen, M., Nielsen, H., Schröder, S., Stage, J.: What happened to remote usability testing? An empirical study of three methods. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 1405–1414. ACM Press (2007)
4. Aykin, N.: Overview: where to start and what to consider. In: Aykin, N. (ed.) Usability and Internationalization of Information Technology, pp. 3–20. Lawrence Erlbaum Associates Inc., New Jersey (2005)
5. Barber, W., Badre, A.: *Culturability: The Merging of Culture and Usability* (1998). <http://research.microsoft.com/en-us/um/people/marycz/hfweb98/barber/>
6. Bargas-Avila, J., Hornbæk, K.: 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 (CHI), pp. 2689–2698. ACM Press (2011)
7. Battarbee, K., Koskinen, I.: Co-experience: user experience as interaction. *CoDesign* **1**(1), 5–18 (2005)
8. Beaton, J., Kumar, R.: Indian cultural effects on user research methodologies. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA), pp. 4267–4271. ACM Press (2010)
9. Bjørneseth, F., Dunlop, M., Strand, J.: Dynamic positioning systems: usability and interaction styles. In: Proceedings of the 5th Nordic Conference on Human-Computer Interaction (NordiCHI), pp. 43–52. ACM Press (2008)
10. Chang, T.-H., Yeh, T., Miller, R.: GUI testing using computer vision. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 1535–1544. ACM Press (2010)
11. Cohen, J.: *Statistical Power Analysis for the Behavioral Sciences*. Academic Press, New York (1998)
12. Dunlop, M., Hamilton, I., Komminos, A., Nicol, E.: Shake ‘N’ Tap: a gesture enhanced keyboard for older adults. In: Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services (MobileHCI), pp. 525–530. ACM Press (2014)
13. Dunlop, M., McGregor, B., Elliot, M.: Using smartphones in cities to crowdsource dangerous road sections and give effective in-car warnings. In: Proceedings of the SEACHI 2016 on Smart Cities for Better Living with HCI and UX (SEACHI), pp. 14–18. ACM Press (2016)

14. Frandsen-Thorlacius, O., Hornbæk, K., Hertzum, M., Clemmensen, T.: Non-universal usability? A survey of how usability is understood by Chinese and Danish users. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 41–50. ACM Press (2009)
15. Gereaa, C., Herskovic, V.: Measuring user experience in Latin America: an exploratory survey. In: Proceedings of the Latin American Conference on Human Computer Interaction (CLIHIC). ACM Press (2015)
16. Gordon, M., Ouyang, T., Zhai, S.: WatchWriter. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 3817–3821. ACM Press (2016)
17. Gorman, T., Rose, E., Yaaqoubi, J., Bayor, A., Kolko, B.: Adapting usability testing for oral, rural users. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 1437–1440. ACM Press (2011)
18. Grubbs, F.: Procedures for detecting outlying observations in samples. *Technometrics* **11**(1), 1–21 (1974)
19. Hassenzahl, M., Tractinsky, N.: User experience - a research agenda. *Behav. Inf. Technol.* **25**(2), 91–97 (2006)
20. Hassenzahl, M.: User experience (UX): towards an experiential perspective on product quality. In: Proceedings of the 20th Conference on l'Interaction Homme-Machine (IHM), pp. 11–15. ACM Press (2008)
21. He, Y., Zhao, C., Hinds, P.: Understanding information sharing from a cross-cultural perspective. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA), pp. 3823–3828. ACM Press (2010)
22. Hofstede, G., Hofstede, G.J., Minkov, M.: *Cultures and Organizations. Software of the Mind*. McGraw-Hill, New York (2010)
23. Hofffeld, T., Keimel, C., Hirth, M., Gardlo, B., Habigt, J., Diepold, K., Tran-Gia, P.: Best practices for QoE crowdtesting: QoE assessment with crowdsourcing. *IEEE Trans. Multimedia* **16**(2), 541–588 (2014)
24. ISO DIS. 2009. 9241–210. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. International Standardization Organization (ISO). Switzerland (2009)
25. Jewell, C., Salvetti, F.: Towards a combined method of web usability testing: an assessment of the complementary advantages of lab testing, pre-session assignments, and online usability services. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA), pp. 1865–1870. ACM Press (2012)
26. Jordan, P.: Human factors for pleasure in product use. *Appl. Ergon.* **29**(1), 25–33 (1998)
27. Karapanos, E., Zimmerman, J., Forlizzi, J., Martens, J.: User experience over time: an initial framework. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 729–738. ACM Press (2009)
28. Kayes, I., Kourtellis, N., Quercia, D., Iamnitchi, A., Bonchi, F.: Cultures in community question answering. In: Proceedings of the 26th ACM Conference on Hypertext & Social Media, pp. 175–184. ACM Press (2015)
29. Komminos, A., Nicol, E., Dunlop, M.: Designed with older adults to Support Better error correction in SmartPhone text entry. In: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI), pp. 797–802. ACM Press (2015)

30. Lachner, F., Nägelein, P., Kowalski, R., Spann, M., Butz, A.: Quantified UX: towards a common organizational understanding of user experience. In: Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI). ACM Press (2016)
31. Lallemand, C., Gronier, G., Koenig, V.: User experience: a concept without consensus? Exploring practitioners' perspectives through an international survey. *Comput. Human Behav.* **43**, 35–48 (2015)
32. Law, E., Roto, V., Hassenzahl, M., Vermeeren, A., Kort, J.: Understanding, scoping and defining user experience. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 719–728. ACM Press (2009)
33. Law, E., Van Schaik, P., Roto, V.: Attitudes towards user experience (UX) measurement. *Int. J. Hum Comput Stud.* **72**(6), 526–541 (2014)
34. Lindley, S., Wallace, J.: Placing in age: transitioning to a new home in later life. *Trans. Comput.-Hum. Interact.* **22**(4), 1–40 (2015)
35. Lindley, S.: Making time. In: 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW), pp. 1442–1452. ACM Press (2015)
36. Malinen, S., Nurkka, P.: The role of community in exercise: cross-cultural study of online exercise diary users. In: Proceedings of the 6th International Conference on Communities and Technologies (C&T), pp. 55–63. ACM Press (2013)
37. Marcus, A., Gould, E.: Crosscurrents: cultural dimensions and global web user-interface design. *Interactions* **7**(4), 32–46 (2000)
38. Mekler, E., Hornbæk, K.: Momentary pleasure or lasting meaning ? Distinguishing eudaimonic and hedonic user experiences. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 4509–4520. ACM Press (2016)
39. Nielsen, J.: Designing for international use (panel). In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 291–294. ACM Press (1990)
40. Norman, D.: Emotional design. *Ubiquity* **45**(4), 1–1 (2004)
41. Obrist, M., Roto, V., Väänänen-Vainio-Mattila, K.: User experience evaluation - do you know which method to use?. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 2763–2766. ACM Press (2009)
42. Obrist, M., Wurhofer, D., Gärtner, M., Förster, F., Tscheligi, M.: Exploring children's 3DTV experience. In: Proceedings of the 10th European Conference on Interactive TV and Video (EuroiTV), pp. 125–134. ACM Press (2012)
43. Oliveira, N.: Culture-aware Q&A environments. In: Proceedings of DC CSCW Companion, pp. 101–104. ACM Press (2015)
44. Oliveira, N., Andrade, N., Reinecke, K.: Participation differences in Q&A sites across countries: opportunities for cultural adaptation. In: Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI). ACM Press (2016)
45. Rea, L.M., Parker, R.A.: Designing and Conducting Survey Research: A Comprehensive Guide. Wiley, Hoboken (2014)
46. Reinecke, K., Bernstein, A.: Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces. *Trans. Comput.-Hum. Interact.* **18**(2), 1–29 (2011)
47. Reinecke, K., Bernstein, A.: Predicting user interface preferences of culturally ambiguous users. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA), pp. 3261–3266. ACM Press (2008)

48. Reinecke, K., Gajos, K.: Quantifying visual preferences around the world. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 11–20. ACM Press (2014)
49. Reyal, S., Zhai, S., Kristensson, P.: Performance and user experience of touchscreen and gesture keyboards in a lab setting and in the wild. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 679–688. ACM Press (2015)
50. Rodden, K., Hutchinson, H., Fu, X.: Measuring the user experience on a large scale: user-centered metrics for web applications. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 2395–2398. ACM Press (2010)
51. Rödel, C., Stadler, S., Meschtscherjakov, A., Tscheligi, M.: Towards autonomous cars: the effect of autonomy levels on acceptance and user experience. In: Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI), pp. 1–8. ACM Press (2014)
52. Roto, V., Law, E., Vermeeren, A., Hoonhout, J.: UX White Paper. Bringing clarity to the concept of user experience, Result from Dagstuhl Seminar on Demarcating User Experience (2011)
53. Sun, H.: Building a culturally-competent corporate web site: an exploratory study of cultural markers in multilingual web design. In: Proceedings of the 19th Annual International Conference On Computer Documentation (SIGDOC), pp. 95–102. ACM Press (2001)
54. Swallow, D., Blythe, M., Wright, P.: Grounding experience: relating theory and method to evaluate the user experience of smartphones. In: Proceedings of the 2005 Annual Conference on European Association of Cognitive Ergonomics (EACE), pp. 91–98. ACM Press (2005)
55. Tuch, A., Trusell, R., Hornbæk, K.: Analyzing users’ narratives to understand experience with interactive products. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 2079–2088. ACM Press (2013)
56. Tullis, T., Albert, B.: Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics, 2nd edn. Morgan Kaufmann, Waltham (2013)
57. Vermeeren, A., Law, E., Roto, V., Obrist, M., Hoonhout, J., Väänänen-Vainio-Mattila, K.: User experience evaluation methods: current state and development needs. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction (NordiCHI), pp. 521–530. ACM Press (2010)
58. Walsh, T., Nurkka, P., Walsh, R.: Cultural differences in smartphone user experience evaluation. In: Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia (MUM), pp. 1–9. ACM Press (2010)
59. Walsh, T., Nurkka, P.: Approaches to cross-cultural design: two case studies with UX web-surveys. In: Proceedings of the 24th Australian Computer-Human Interaction Conference (OZCHI), pp. 633–642. ACM Press (2012)
60. Walsh, T., Vainio, T.: Cross-Cultural Design for mHealth Applications. In: Proceedings of the 23rd Australian Computer-Human Interaction Conference (OZCHI). ACM Press (2011)
61. Yatani, K., Novati, M., Trusty, A., Truong, K.: Review spotlight: a user interface for summarizing user-generated reviews using adjective-noun word pairs. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 1541–1550. ACM Press (2011)

Investigating Perceptions of Personalization and Privacy in India

Hanna Schneider^(), Ceenu George, Malin Eiband, and Florian Lachner

LMU Munich, Munich, Germany
hanna.schneider@ifi.lmu.de

Abstract. Technological products are increasingly equipped with data collection and personalization mechanisms that allow them to adapt to an individual user's needs [4]. However, the value and perception of these practices for users is still unclear. This field trip proposal investigates users' mental models of personalization as well as perceived benefits and drawbacks using semi-structured interviews. The interviews make use of the critical incident technique and drawing tasks. We expect that findings from the field trip will result in rich understanding of the perspective of a collectivist society on personalization and privacy. Results of the field trip can, hence, be contrasted to the results of an equivalent study conducted in Germany, an individualistic society. The overall goal of our studies is to highlight differences in user needs of collectivist and individualistic societies for researchers and practitioners who develop highly personalized systems.

Keywords: Privacy · Personalization · Value-driven design

1 Introduction

As data collection and processing techniques improve, technological products are increasingly equipped with personalization mechanisms that allow them to adapt to an individual user's needs [4]. One of the main purposes of personalization and customization is to evoke or foster a feeling of individuality or “me-ness”, which is especially important in individualistic and particularly Western societies [4,8]. However, the notions of individuality and privacy are perceived very differently, and often in a negative light, in societies with a collectivist world view [7], e.g. in the Arab Gulf [1], in Africa [3], or in India [2]. The goal of this fieldtrip is to investigate users' perceptions of benefits and drawbacks of personalization in India. Currently, we are conducting an equivalent study in Germany. The pre-study in Germany indicated differing mental models of personalization and varying privacy sensibility when using personalized products or services. We, therefore, want to further investigate this relation. However, to objectively address this research question, we want to understand privacy concerns of personalization in both individualistic and collectivist countries.

© IFIP International Federation for Information Processing 2017
Published by Springer International Publishing AG 2017. All Rights Reserved
R. Bernhaupt et al. (Eds.): INTERACT 2017, Part IV, LNCS 10516, pp. 488–491, 2017.
DOI: 10.1007/978-3-319-68059-0_57

Both studies will be conducted in similar settings hence we expect that findings from the field trip in India (a more collectivist country in contrast to Germany, an individualistic country [7]) will enrich our study with an alternative perspective to personalization and privacy and enable us to critically reflect on values and assumptions that underlie the design of personalized technologies. We hope that our findings will inform researchers and practitioners who develop personalized systems of similarities and differences in user needs between these two societies.

2 Project Plan

2.1 Goals

- Defining the necessity of personalization in a collectivist country
- Understanding needs for personalization and privacy in different contexts
- Analyzing culturally sensitive aspects of personalization and privacy

2.2 Locations

For this study, we propose two different contexts, namely (1) urban and (2) rural areas. We believe that this would provide an interesting setting to study the perception of personalization and privacy, as we believe that the perception will vary based on distance to the city. Participants should also be from a broad age group and gender should be equally distributed. We will split up into teams, each team targeting a different context. However, to select specific locations around Mumbai, we would appreciate advice from the organisational team of the conference or local universities. We are flexible on the exact locations as long as we can target the proposed geographical group of people.

2.3 Method

Our study plan is focused on qualitative data collection techniques. We combine semi-structured interviews, based on the critical incident technique [5], and drawing tasks to elicit users' mental models of personalization [9]. The interview script consists of questions about demographic data as well as users' technology use, their mental model of personalization and their perception of benefits and drawbacks. The ten main questions of the semi-structured interviews are:

1. Which specific websites did you visit last week?
2. What was your goal when you visited the websites?
3. Did you use different devices?
4. Did you notice that either the content or the interface of the webpage were tailored to you in any way?
5. Did you find this tailoring to you good/bad/helpful/useful? How did you like it?
6. Do you have any concerns about this tailoring?

7. Can you give us an example of a website you visited and you were concerned about the tailored website?
8. Do you think the page knows something about you, and if so, what?
9. How do you think this tailoring was done? Could you please draw and explain your thoughts?
10. Is there anything that you would have (not) liked to be tailored to you?

To better understand the context of use, interviewers may ask additional questions and note down observational data as needed. Moreover, as we assume that individual attitudes towards individualism and collectivism will be of importance for the analysis of the qualitative data, we will use standardized quantitative scales such as INDCOL [10] or the Culture Orientation Scale [6] to better understand our participant sample.

2.4 Participants, Recruiting and Ethical Considerations

All interview questions were designed according to ethical guidelines and all data will be stored and analysed anonymously. Participants will be informed about the goals of our study and asked for their consent before the interview. They will be paid Rs.200 as a token of compensation.

2.5 Schedule

Day 1

- 9:00–10:00: Team meeting, briefing, splitting up in smaller groups
- 10:00–17:00: Groups go out to their locations and conduct interviews and observations

Day 2

- 9:00–11:00: Groups share their findings + insights
- 9:30–12:00: Clustering of insights + definition of focus group topics
- 12:00–13:00: Lunch
- 13:00–14:00: Focus groups and scenario sketching
- 14:00–15:00: Summary of results + extracting of recommendations
- 15:00–15:30: Wrap-up + definition of next steps

3 Expected Outcome

Findings will help to contrast the views of personalization and privacy as prevalent in Western societies. The findings will further be used to create recommendations and guidelines for personalization and privacy for technological products targeting the Indian market.

4 Field Trip Participants

The field trip will be carried out by researchers from various disciplines and both academia and industry:

Hanna Schneider (Organizer, LMU Munich)
 Florian Lachner (Organizer, CDTM Munich)
 Elaine Brechin Montgomery (Facebook)
 Alfred Kobsa (University of California, Irvine)
 Panayiotis Zaphiris (Cyprus University of Technology)
 Corinne Longman (Google)
 Beth Bergen (Google)
 Anjali Kukreja (HealthKart Gurgaon, India)
 Pratiksha Dixit (Srishti Institute of Art, Design and Technology)

References

1. Abokhodair, N., Vieweg, S.: Privacy & social media in the context of the arab gulf. In: Proceedings of the 2016 ACM Conference on Designing Interactive Systems, DIS 2016, NY, USA, pp. 672–683 (2016). <http://doi.acm.org/10.1145/2901790.2901873>
2. Asokan, A.: The space between mine and ours: exploring the subtle spaces between the private and the shared in india. In: Ethnographic Praxis in Industry Conference Proceedings, vol. 2008, pp. 175–191. Wiley Online Library (2008)
3. Bidwell, N.: Ubuntu in the network: humanness in social capital in rural africa. *Interactions* **17**(2), 68–71 (2010)
4. Blom, J.: Personalization: a taxonomy. In: CHI 2000 Extended Abstracts on Human Factors in Computing Systems, pp. 313–314. ACM (2000)
5. Flanagan, J.C.: The critical incident technique. *Psychol. Bull.* **51**(4), 327 (1954)
6. Gudykunst, B., Singelis, T., Trafimow, D., Realo, A., Triandis, C., Street, E.D.: Converging measurement of horizontal and vertical individualism and collectivism. *J. Pers. Social Psychol.* **74**(1), 118–128 (1998)
7. Hofstede, G., Hofstede, G.J., Minkov, M.: *Cultures and Organizations. Software of the Mind*. McGraw-Hill, Maidenhead (2010)
8. Kalyanaraman, S., Sundar, S.S.: The psychological appeal of personalized content in web portals: does customization affect attitudes and behavior? *J. Commun.* **56**(1), 110–132 (2006)
9. Kang, R., Dabbish, L., Fruchter, N., Kiesler, S.: “my data just goes everywhere:” user mental models of the internet and implications for privacy and security. In: Eleventh Symposium on Usable Privacy and Security (SOUPS 2015), pp. 39–52. <https://www.usenix.org/conference/soups2015/proceedings/presentation/kang>
10. Singelis, T.M., Triandis, H.C., Bhawuk, D.P., Gelfand, M.J.: Horizontal and vertical dimensions of individualism and collectivism: a theoretical and measurement refinement. *Cross Cult. Res.* **29**(3), 240–275 (1995)

A Bermuda Triangle? - A Review of Method Application and Triangulation in User Experience Evaluation

Ingrid Pettersson*
Volvo Cars and Chalmers
University of Technology
Göteborg, Sweden
ingrid.pettersson@volvocars.com

Florian Lachner*
LMU Munich
Media Informatics Group
Munich, Germany
florian.lachner@ifi.lmu.de

Anna-Katharina Frison*
Technische Hochschule
Ingolstadt and Johannes
Kepler University Linz
anna-katharina.frison@thi.de

Andreas Riemer
Technische Hochschule
Ingolstadt and Johannes
Kepler University Linz
ansreas.riener@thi.de

Andreas Butz
LMU Munich
Media Informatics Group
Munich, Germany
butz@ifi.lmu.de

ABSTRACT

User experience (UX) evaluation is a growing field with diverse approaches. To understand the development since previous meta-review efforts, we conducted a state-of-the-art review of UX evaluation techniques with special attention to the triangulation between methods. We systematically selected and analyzed 100 papers from recent years and while we found an increase of relevant UX studies, we also saw a remaining overlap with pure usability evaluations. Positive trends include an increasing percentage of field rather than lab studies and a tendency to combine several methods in UX studies. Triangulation was applied in more than two thirds of the studies, and the most common method combination was questionnaires and interviews. Based on our analysis, we derive common patterns for triangulation in UX evaluation efforts. A critical discussion about existing approaches should help to obtain stronger results, especially when evaluating new technologies.

CCS Concepts

•Human-centered computing → User studies; Empirical studies in HCI; HCI theory, concepts and models;

Author Keywords

User experience; UX; evaluation; triangulation; mixed methods; review; meta-analysis

* First three authors contributed equally

INTRODUCTION

User experience (UX) has attracted increasing interest in recent years. One comparable indicator, at least for academia, is

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

© 2018 ACM. ISBN 978-1-4503-5620-6/18/04...\$15.00

DOI: <https://doi.org/10.1145/3173574.3174035>

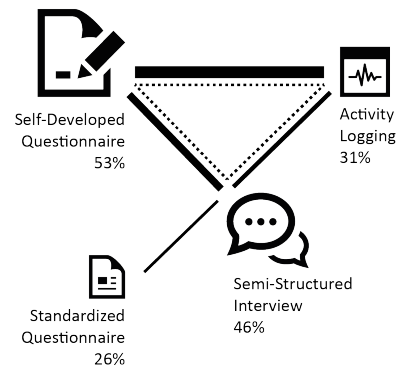


Figure 1. Commonly applied methods and triangulation strategies found in our literature study. The size of the nodes (methods) and links (combinations) is proportional to the number of occurrences.

the number of papers published. In Google Scholar, for example, the search term “user experience” roughly returns 20,800 results for the publication year 2010, increasing to more than 32,500 for the year 2016 (+56% in total). While UX considerations have become embedded in research and design processes, they still remain a challenging and strongly discussed area for both researchers in academia and practitioners in industry.

The focus of UX, for example defined as “A person’s perceptions and responses that result from the use or anticipated use of a product, system or service” [21], has shifted from simply acknowledging usability and performance aspects of interactive products, towards the emotional, hedonic aspects of interaction. Reaching beyond usability, UX has been extended to incorporate hedonic qualities such as aesthetics, stimulation and identification [37]. UX also is a dynamic concept influenced by contextual aspects, such as place, social and temporal aspects of use as well as the users’ specific emotional states [47, 23]. It is thus clear that UX offers a much more holistic and dynamic take on interaction with products than pure usability.

As an academic discipline, UX has evolved in the multi-disciplinary intersection of fields such as cognitive science, design, psychology and engineering. Not only the definition of the term UX is discussed, but also the question how to study and evaluate UX. A number of surveys of available methods and methodological gaps were conducted around the turn of the present decade (see, for example, [90, 107, 80, 9]). However, as the field has evolved over the years, we see the need to reassess the field of UX evaluation. We want to take the thoughts of the previous meta-analyses of Bargas-Avila and Hornbæk [9], Vermeeren et al. [107], Roto et al. [90] and Obrist et al. [80] - published around 2010 - further, to examine the progress until today. Our goal is to update the knowledge and provide an analysis of the current characteristics of empirical studies in UX from 2010 to 2016, as well as an outlook on possible future developments. In particular, we are interested in how the multi-dimensionality of UX is approached by using (or not using) method triangulation, i.e., applying two or more methods, to obtain well-founded evaluation results from different perspectives. Our analysis therefore includes trends regarding the number of UX publications, UX dimensions studied, study contexts, method application, and triangulation patterns derived from the analysis of method application.

Building on the current state of the art, we point out potential gaps and future directions. Is the topic of UX evaluation still a Bermuda Triangle of disparate research approaches, or is there land in sight for an evolved UX evaluation practice?

RELATED WORK

Evaluation has been identified as one of the core pillars of academic UX research [58, 63, 103]. Alves et al. [5] demonstrated in 2014 that experience evaluation plays an important role for UX practitioners in industry and concludes that “most practitioners believe that UX evaluations have a strong to decisive impact on the user interface”. The approaches used for empirical UX evaluation have, however, been debated (see for example [13, 63, 90]) as a result of the different epistemological directions of the research area.

It has been claimed that evaluating UX requires new assessment methods and approaches [9, 50] and a wide range of UX evaluation methods have been developed over the years (e.g., UX curve [56], AttrakDiff [39], or the UX-q [61]). Bargas-Avila and Hornbæk [9] pinpoint three areas of intense debate regarding evaluation: the types of products studied, the dimensions of UX, and the methodologies applied, reigniting “the debate between qualitative and quantitative approaches” [9]. Law et al. [63] point out the epistemological differences between engineering approaches to UX where “to measure is to know” and the humanities, where it may be considered naive and simplistic to assume that a fuzzy concept such as experience can be readily reduced and measured.

The ambition of sound empirical evaluations [63] applicable to UX thus led to the meta-review efforts in the late 2000’s and early 2010’s. These reviews concerned both the underlying theory and models of UX [58, 64] as well as the area of UX evaluation [5, 9, 80, 90, 107]. A number of common methodological gaps in UX evaluation approaches were identified at this point of time:

- In previous work, the underlying, assumed dimensions of UX have been found to often be unclear and/or interchangeable with usability [64]. The often very vague link between evaluations and UX theory is problematic; not fully understanding and formulating what is evaluated makes improvements difficult.
- Existing evaluation work was largely founded on self-defined questionnaires and few UX methods were satisfactorily validated for the cause [9, 107].
- There were only very few practices for evaluating UX at early stages of design processes [90, 107, 9]. These need formative methods, providing feedback that can provide information on what and why to revise to improve a design already early in the development [107].
- There was a lack of UX method triangulation that addresses the multidimensionality of experience. The lack of rational ways of collecting data from multiple methods was found by several reviews [9, 90, 107]. Although “benefits in terms of a rich picture of UX and higher scientific quality by collecting data with a combination of UX evaluation methods are well recognized” [107], there was a call for developed knowledge of which methods work well together and how to effectively collect and analyze the data from different sources.

Triangulation

As mentioned, the lack of multi-dimensionality in the evaluation formats was a commonly identified issue in previous meta-reviews. Law et al. [65] write “employing quantitative measures to the exclusion of qualitative accounts of user experiences, or vice versa, is too restrictive and may even lead to wrong implications”. Arhippainen et al. [8] demonstrate how applying several methods in practice can help researchers to learn about users and their ways to express experiences, and to catch “user experience information piece by piece by utilizing different methods”. However, they conclude that there is a general lack of knowledge generation in the area of UX and using multiple methods. Furthermore, it has been suggested that experience data is accessible in layers [108] and that therefore more of an experience can be understood by, for example, moving between “the expressible” by interviews [93], to tacit behaviors by observation techniques and latent experience data of “knowing, dreaming and feeling” by, for example, generative sessions [108].

Outside the UX field, the roots of method triangulation approaches traces back to the “paradigm wars” [104] and grew in popularity during the 1980s in social, behavioral and human sciences [44] to bridge different epistemological standpoints in research; applying triangulation approaches served as a way to overcome differences in approaches to knowledge production, enabling both the abilities of qualitative research in understanding the subjective, as well as the quantitative to determine statistical trends and connections. Denzin [19] outlined four types of triangulation in order to study a phenomenon: (a) data triangulation (i.e., the use of a mix of data sources in a study), (b) investigator triangulation (i.e., a number of researchers researching the same phenomenon), (c) theory triangulation

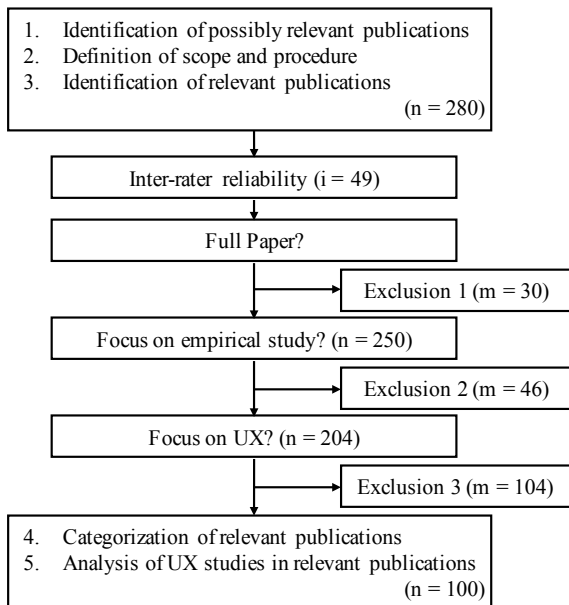


Figure 2. Procedure of our literature review, adapted from the QUOROM statement.

(i.e., the use of a number of theories used for interpreting results of a study), and (d) methodological triangulation (i.e., the use of more than one method to study a phenomenon). Creswell [18] describes the overarching two different types of employing two or more methods; either *sequential* (firstly either a quantitative or qualitative method is used, and the other type is used in a following study to explain, explore or validate the results) or *concurrent* (where two or more methods are employed within the same study to cross-validate findings). Employing triangulation to study a subject has been claimed to contribute to a more reliable, holistic and well-motivated understanding of phenomena [44], and to counteract inherent biases from data sources, investigators and especially methods.

Thus, triangulation can be claimed to lead to higher confidence in results and also to uncover unexpected results. This is true especially for multi-dimensional topics, such as UX, that may need multiple approaches, bringing depth as well as breath to the understanding. In their review of empirical studies, Bargas-Avila and Hornbæk found a “sad lack of reference” [9] between the groups of qualitative and quantitative methodological approaches. We are following up on the topic of triangulation approaches in UX evaluation - has there been progress? What can we learn from strategies currently applied?

STUDY AIM AND LITERATURE STUDY SETUP

UX evaluation continues to be a popular, but often debated, topic in conferences as well as in journals. Our goal is to provide an updated overview of UX evaluations in academia, from the years that have passed since the previous efforts of literature reviews, e.g., [9, 80, 90, 107]. We are interested in the methodologies employed, the dimensions of UX that are

Conference / Journal	Database (ACM DL)	Search Results (total)	Relevant Publications
CHI	7,482	137	50
UbiComp	2,265	34	10
DIS	1,092	40	14
CSCW	1,701	17	8
UIST	839	13	3
ICMI	843	1	0
IUI	837	20	7
TOCHI	278	18	8
Total	15,337	280	100

Table 1. Number of search results (excluding results from extended abstract, companion, and adjunct proceedings) and relevant publications for our analysis (per venue from 2010 to 2016).

studied, the products studied and the stages of development process in which the studies takes place. Furthermore, we investigated the application of triangulation within UX studies, whether it has increased since the turn of the decade, and identified current approaches for handling multidimensionality.

By evaluation, we mean a focus on assessing specific designs, from early concept ideas, over prototypes to finished products, in order to inform a design process. As our aim is to study the overall UX field, based on the academic literature, we did not select a specific practical model or view of UX as a basis for paper selection. Accordingly, we reviewed papers stating to evaluate UX as a part of their description of the study.

We decided to pursue a systematic analysis approach, based on a representative sample of publications in the field of Human-Computer Interaction and UX, to derive suitable insights regarding triangulation approaches in academic UX studies. Our approach (see Figure 2) is based on a procedure similar to Bargas-Avila & Hornbæk [9] and Lachner et al. [57] hence similarly adapted from the QUOROM statement [77], which specifies guiding principles on how to conduct meta-analyses including a quantitative data synthesis and a clarifying flow diagram.

Step 1: Identification of possibly relevant publications

Source selection. Academic work related to UX evaluation is spread across multiple scientific journals and conferences and continues to gain interest.

To limit the scope of our analysis we decided to only use the ACM Digital Library (DL) as a research database, including 476,316 records (207,571 from 2010 to 2016) in total at the time we conducted our review.

Furthermore, besides being a rich source of UX research, it contains influential conferences and highly ranked journals such as CHI, DIS and TOCHI. We further narrowed down the scope through our selection of suitable target conferences and journals. We based the selection of suitable venues on the h5-index indicated by Google Scholar as well as the relevance for our study aim hence omitted proceedings that target a specific domain (e.g., robotics, mobile, etc.). As a consequence, eight ACM conferences were selected (see Table 1).

Search procedure. We used the search query “*user experience*” AND “*evaluation OR method OR measure OR assessment OR study*” in any field and limited the search results to the period 2010 to 2016. We excluded the current year (2017),

since the publication year was not yet finished and thus not yet all possible relevant papers available. Next, we selected the conference proceedings (respectively the journal) of all selected venues and excluded all extended abstract, adjunct, and companion proceedings if they were listed individually (based on [9]). False positives were excluded at later stages. Our procedure led to 280 papers in total (see Table 1).

Step 2: Definition of scope and procedure

Exclusion criteria. As a next step, we defined criteria to exclude publications that were out of scope of our study aim. A publication was excluded if (1) it was not a full paper, or (2) if there was no trace of an empirical study, i.e., not including an empirical study of any kind, not including a clear description of the evaluation process or the study results, or (3) if the authors of the paper did not clearly state that evaluating UX (in any form) of a product or service was the aim of the empirical study. For the third exclusion criteria, it is important to note that we only included publications where the authors directly linked their study to UX while we excluded publications where the term “user experience” was only mentioned in the abstract, key words, related work, and/or introduction and not as a part of describing the specific study. We did not judge the authors’ views on UX; if authors claimed to study user experience, they were included in our detailed categorization.

Screening categories. Next, we selected the categories (based on [9]) that we used to analyze all relevant publications. Name and type of method were categorized (e.g., self-developed questionnaire, standardized questionnaire, free interview, semi-structured interview, activity tracking, live observation) and information if the method was referenced or not was noted. The (data) type of the methods in each publication (qualitative, quantitative, or both), task orientation (explorative use where the user was free to explore without guidance, or task-oriented), and if a motivation for the use of triangulation was stated (if two or more methods/types of data were applied). We also noted place of study (lab/field) and period of use (single-session/shortterm/longterm). For each study, we noted if there were references to UX literature and theory, as well as which dimension of UX that was studied. Thereby we differentiate between the consideration of following aspects: generic UX (experience is studied as an own construct without mentioning directly what is to be collected or measured), pragmatic quality (usability, functionality), hedonic quality (“psychological well-being through non-instrumental, self-oriented product qualities” [20, 36]), aesthetics/appeal, satisfaction, affect/emotions, enjoyment/fun engagement/flow, frustration, motivation and other constructs. We elaborate further on the categories in the results section.

Procedure. A common understanding and acceptable inter-rater reliability of the definition and interpretation of the exclusion criteria and the screening categories was ensured by four cross-checks before the final screening, containing of in total 60 papers mutually reviewed in full. Each of the four cross-check rounds consisted of an independent analysis as well as a joint telephone conference of the first three authors to discuss 15 papers from each of the venues CHI ’16, CHI ’15, DIS ’16, and DIS ’14. These cross-check helped to decide how

to interpret the inclusion or exclusion of papers and how to assign the defined categories. Most prominently, we sharpened selection criteria 3 and decided to exclude evaluations that mainly target the analysis of experiences with technology in general, e.g., with the aim to derive UX theory, as compared to the evaluation of a specific product or product type. Whereas Hayashi and Hong [41], for example, state that “the overall goal of the studies was to investigate the user experience in using an authenticator [...]” (included in our review), Tuch et al. [101] start their survey study with the question “Bring to mind a single outstanding positive experience you have had recently” (omitted). Similarly to the latter, Mekler & Hornbaek [72] aim “to identify hedonic and eudaimonic components of [...] experiences”. Both latter examples were excluded from our review for stringency to our meta-study by a clear focus on the directed evaluation/exploration of a product/service type or case.

Step 3: Identification of relevant publications

Inter-rater reliability. We considered the exclusion of publications as crucial for the subsequent analysis. To ensure the reliability of the selection process, the first three authors conducted a final screening test round. In the test round, each author individually screened the same set of 49 papers (17,5% of all possibly relevant publications). The set consisted of randomly selected papers of each conference/journal and year. The inter-rater reliability for the exclusion was found to be $\alpha = .8307$, 95% in a CI of (0.7161, 0.9345). According to Krippendorff [53], values for α higher than .8 can be seen as a satisfactory.

Procedure. For the final screening procedure, we split all publications of papers between the first three authors with weekly meetings to discuss borderline papers.

Exclusion. Firstly, 30 publications were excluded because they were not full papers. Second, 46 papers were excluded because they were not empirical studies of a specific product or product type. Third, 102 papers were excluded for not relating the concept of UX to the empirical study (see Figure 2).

Step 4: Categorization of relevant publications

After identifying all relevant publications, the three authors categorized the same set of publications according to the defined screening categories. Once again, weekly meetings were held to handle unsure cases of categorizations.

Step 5: Analysis of UX studies in relevant publications

Interim analysis. After about half of the time needed for the categorization in step 4, we organized a workshop at DIS 2017 [84] to discuss initial insights with researchers in the UX field. At the workshop, we presented first insights of our review, including, for example, types of products studied, UX dimensions addressed, referenced UX theory, employed methods, and triangulation approaches. Together with all workshop participants, we interpreted our initial findings at that time. These interpretations provided an initial basis for our final screening and analysis process.

Final analysis. Finally, we finished screening and categorizing all relevant publications. The final screening also included

one more cross-check for borderline papers. For the analysis, we first looked at general developments in the field of UX evaluation. Second, we specifically examined mixed method approaches/triangulation patterns.

RESULTS

Below we report our results in following structure: We begin by describing general insights about the development of the amount of UX publications over time (1), the studied dimensions (2), and context (3). Then, we present our analysis of the applied methodology (4), which ends up in detailed insights about common triangulation patterns (5) in UX studies.

The Development of User Experience

In general, we see a temporal development of the papers we rated as relevant (empirical studies with a focus on UX) with a percentage increase of 283% from 2010 to 2016. This means that in total 40% (N=100) (of all identified full papers (N=250)) were identified as relevant for a detailed categorization. However, the percentage of all full papers that have been excluded because they did not focus on UX (in total 41%) converges over the years with the percentage of papers we rated as relevant. Looking at the linear trend lines, there is only a slight decrease (-2.5%) of papers rated as relevant, and a slight increase of “no UX papers” (5.1%). Thus, neither a positive nor a negative development was observed over time (in terms of amount of relevant papers, i.e., the numbers of publications containing the search words increased continuously) but the percentage of papers found relevant in relation to the total numbers of papers containing the search words, remained fairly stable. This means that there is a continuous growth of UX studies. Below, we will discuss some insights regarding the UX dimensions we studied, the context and the methodology we used.

UX Dimensions

UX is mostly studied as a general construct

In the detailed analysis and categorization of relevant papers (N=100), we found that the category that we summarized as *Generic UX* was the most frequently evaluated UX dimension (56%, see Table 2 for a full list of all UX constructs we used). In the *Generic UX* category, UX authors understood UX as a general construct and did not specify which aspects they studied in detail. For example, Woo et al. [111] describes: “we conducted a qualitative user study to understand people’s

UX dimensions	%**	Examples
Generic UX	56	[111],[75],[78]
Pragmatic Quality	22	[1], [109], [99]
Aesthetics/Appeal	7	[96]
Hedonic Quality	6	[99]
Satisfaction	4	[106]
Affect/Emotion	4	[54]
Enjoyment/Fun	4	[69]
Engagement/Flow	3	[4]
Frustration	2	[86]
Motivation	1	[100]
Other Constructs	16	e.g. trust [82]

Table 2. Dimensions of UX research. Note: *multiple dimensions in one study possible, based on all relevant papers (N=100).

experiences with DIY smart home products”. In Bargas-Avila and Hornbæk’s study of papers from 2005-2009 [9], *Generic UX* was also the main experiential dimension, yet slightly less prominent (in 41% of all papers). Consequently, there has been an increase from 2010 to 2016 of papers that evaluate UX as a broader construct. In 2010, only 1 out of 6 papers studied UX as a general construct [49]. In 2016, 52% of the papers did not define any additional concrete dimensions.

In addition, 22% of all relevant papers measured the pragmatic quality, by constructs of usability, ease of use, and/or efficiency. The reasoning behind focusing on the pragmatic quality differed, e.g., “to better understand the impact on the User Experience, we conducted a lab-based user study to evaluate the effectiveness of different time series visualizations that use varied interaction techniques, visual encodings and coordinate systems for four tasks [...]” [1]. Other constructs which were understood as a dimension of UX included Aesthetics/Appeal (7%), hedonic quality (6%) and satisfaction (5%). Enjoyment/Fun and Affect/Emotion, both considered as core dimensions in [9], are only investigated in 4% of all papers. Engagement/Flow, Frustration and Motivation were also rarely studied.

Rare links to UX theory

The high percentage of papers using a vague description of UX is also reflected in the theoretical frameworks of the papers. Overall, only 17% of all papers use a definition of UX, 83% do not. Furthermore, established UX theory papers, e.g., [21, 23, 64, 65, 37, 38] are only referenced extensively in 8% of all papers, to some extent in 17% and in 75% not at all. Häkkinen et al. [32] wrote “[...] Although there is hardly a unified definition for UX [62, 64], it is widely agreed that UX goes beyond usability and instrumental aspects [64]. A definition presented in [38] describes UX as ‘a momentary, primarily evaluative feeling (good-bad) while interacting with a product or service’ [...]”, as an example of a paper containing extensive references to UX theory. Such extensive descriptions were rarely found in all relevant papers.

Context

Wide range of products studied

While in 2005 to 2009, art was the most frequently evaluated product [9], we found a very wide range of products studied from 2010 to 2016. 20% of all papers evaluated individual products - too diverse to create their own categories, e.g., icons [96] or interactive museum installations [42]. The most frequently studied product types were mobile phone/app (15%), followed by interactive games (13%), web tools (12%) and websites (10%). The UX of new technologies such as connected services/IoT (4%) and VR/AR (2%) are increasing, but the number of papers is still small, see Table 3.

Few studies focus on early product development stages

In total, 56% of all studied products are presented and discussed as prototypes (of which 96% are high-fidelity prototypes), 39% are finished products or beta versions. Earlier stages of concepts were rarely evaluated. Only two papers used a wizard-of-oz setting and only one paper analyzed stories in a narrated form.

Studied Products	%*	Examples
Mobilephone/App	15	[54], [67]
Interactive Game	13	[7], [3]
Webtool	12	[30], [83]
Website	10	[75], [100]
Professional Software	9	[99]
Audio/Video/TV	6	[32]
Connected Service/ IoT	4	[111]
Non-digital Product	3	[45]
Vehicle	3	[40]
VR/AR	3	[81]
Wearable	2	[24]
Other Products	20	e.g., icons [96]

Table 3. Products studied in UX research. Note: *multiple products in one study possible, based on all relevant papers (N=100).

Equal share of field and lab studies

In contrast to [9], where only 21% regarded the context of the product in their study, from 2010 to 2016, almost the half (45%) of the publications described a field study and thus involved the context in their investigations. Ghellal et al. [29], for example, studied the experience of an augmented reality game within a horror and vampire genre “merging a fictional universe and the physical environment into one pervasive experience, centering around a variety of augmented reality activities played out at sunset”. At the same time, 41% were lab studies. Remote studies were conducted by only 8% and a mixed setup (lab and field) by 4% (all percentages rounded). We found that professional software tools (77.8%; N=7), mobile phone/apps (53.3%; N=8) and connected services (75%; N=3) were mainly investigated in a field study, while interactive games (61%, N=8) are more frequently evaluated in a controlled lab setting.

Still limited “truly” longitudinal studies

Furthermore we can report that 63% of all selected UX studies are evaluating the UX within a single session. These sessions were mostly conducted in the lab (61.9%; N=39), or within a field study (25.4% N=16). However, at least 34% of all UX studies used a long term setup (several weeks). Of these, 76.4% are conducted at the field, 14.7% assess UX remotely and 5.9% are performed in a lab. Only 3% are analyzed in a short term setting (i.e., several days, thus longer than a single session but not for weeks or longer). In total we can speak of a positive development since 2010, where the studies stretching over several weeks were only available in “some papers” [9]. However, “truly” longitudinal studies which “cover typical product life cycles over several months and years” [9] are still missing at large, with important exceptions such as [55].

Method Application

Focus on traditional methods

In order to evaluate UX, a variety of methods from related fields, as well as newly developed methods have been employed over the years. With regards to method deployment in our data set (see Table 4), we observed that self-developed questionnaires were used in more than half of all papers (53%), 46% conducted semi-structured interviews, 31% employed activity logging, 26% used a standardized questionnaire, and 19% observed their users (see Figure 1). Probes (i.e., additional material given to the users to elicit experiences,

Method Type**	%*	Examples
Self-Developed Questionnaire	53	[81], [32], [75]
Semi-Structured Interviews	46	[24], [32],[111]
Activity Logging	31	[3], [99],[49]
Standardized Questionnaires	26	[40], [59]
Live User Observation	19	[96], [7]
Videorecording	16	[99], [45]
Free Interview	9	[102]
Think Aloud Feedback	6	[82]
Diaries	6	[7]
Focus Groups	5	[7]
Online Feedback	3	[66]
Probes	3	[79]
Physio-psychological	2	[25]
Others	5	e.g., sticky labels to capture context [68]

Table 4. Methods used in UX research. Note: *multiple methods in one study possible, based on all relevant papers (N=100) **data is collected without judging, whether the measurement is right for assessing UX.

such as the possibility to express experiences through video, photo or drawings) and objective measures such as physio-psychological methods were rarely used. Self-developed questionnaires are also the method which is most commonly used stand-alone (11%), followed by semi-structured interviews (9%) and standardized questionnaires (6%).

Broad range of specific methods

When looking at established methods that focus on a specific evaluation scenario, an analysis shows that there is a broad range; 40% use a unique method, which no other study employs in our dataset (see Table 5). The NASA-TLX questionnaire, developed to assess workload [35], was (perhaps surprisingly) the most frequently used method in all UX studies (7%). Next, the AttrakDiff questionnaire [39] (5%) and a second version of it (2%) were employed. The System Usability Scale was used in 3% of the papers. As a consequence, we cannot report a high consensus in methods in general.

Specific Method**	%*	Examples
NASA-TLX	7	[59]
AttrakDiff	5	[54]
System Usability Scale	3	[24]
AttrakDiff 2	2	[49]
User Engagement Scale	2	[4]
Aesthetics scale	2	[87]
PANAS	2	[52]
Others	40	e.g., SAM [76]

Table 5. Specific methods used in UX research. *multiple specific methods in one study possible, based on all relevant papers (N=100) **data is collected without judging, if the measurement is the right for assessing UX.

Data type is mostly mixed

Our systematical categorization shows that 32% of the data collected in UX studies is solely quantitative (e.g., activity logging, questionnaires, psycho-physiological data) and 22% solely qualitative (e.g., interviews, observations). However, combinations of different methods based on the same data type, meaning either two or more qualitative (respectively quantitative) methods, are part of these numbers. The bigger part of the papers (46%) applied both quantitative and qualitative measurements, meaning that the studies used two or more methods, i.e., a triangulation approach. This is further analyzed in the next section, both for method triangulation as well as also data triangulation.

Triangulation Patterns

Analyzing the methodology of the 100 selected papers, we can observe that the majority (72%) uses a triangulation approach. While 21% mix only the method (e.g., Campbell et al. [14] use activity logging during the interaction with weblog posts and a post-use self-developed questionnaire focusing on ease of use, enjoyment, and intention to return), in 46% also the data type (quantitative and qualitative methods) was triangulated. Only two papers mix theory [89] or user groups [83].

We were able to cluster and identify 8 insights about triangulation patterns based on our analysis of general method combinations (see Figure 2) and data type per temporal stage of assessment, i.e., before, during, or after the evaluated interaction (see Table 6) as described below.

Gaining deeper insights motivates triangulation

Our analysis shows that many authors justify the use of multiple methods. In 32 (44%) of all papers that use any kind of triangulation (N=72) the authors state a motivation for the use of multiple methods. Besides, e.g., Ardito et al. [7] who justify their approach based on related work in the field of triangulation, most authors briefly mention that their aim of applying more methods or data types was to get deeper insights (e.g., [17] or [105]). Additionally, a second main reasoning behind method triangulation was to better understand the results of other applied methods, e.g., using post-use interviews to make sense of observations (see [31]) or post-use interviews to make sense of video recordings (see [67]). A total of 23% of all papers reported only positive results from the study, while the rest reported mixed results or mainly negative outcomes of the evaluation. When including only papers that applied triangulation, 19% reported on only positive outcomes.

Tendency towards triangulating a few methods

Our analysis also shows that the majority of method combinations is only based on a small set of different methods (see Figure 1). More precisely, we saw that self-developed questionnaires are most frequently used together with activity logging (23%) or semi-structured interviews (20%). Furthermore, semi-structured interviews are often combined with activity logging (15%) or with standardized questionnaires (11%). A typical triangle of methods (not necessarily applied in isolation) is the combination of self-developed questionnaires, semi-structured interviews, and activity logging (10%). The most frequent triangulation pattern, which is used standalone (without any additional methods) is self-developed questionnaire and activity logging (9%), whereas 6% additionally use semi-structured interviews. Thus, compared to previous research [9], we see a substantial increase of the use of activity logging as a complement to more traditional self-reporting. However, when qualitatively reviewing the content of the method descriptions, results, analysis and discussions, we often found weak links between the conclusions drawn from both sources.

Preference for quantitative data

From all methodological approaches, post-assessment based on quantitative methods represents the most common approach (see Table 6). However, from 16 publications which use questionnaires after the experience, there was only one publication

No.	Comment	Temporal Stage			%*
		before	during	after	
1	only one stage / quantitative			■	16
2	during/after: quantitative		■	■	11
3	post-use triangulation			□■	9
4	only one stage / qualitative			□	7
5	during/after: mixed		□	□■	6
6	during/after: qualitative		□	□	6
7	all stages: qualitative	□		□	5
8	during/after: mixed		□■	□	5
9	only one stage / qualitative			□	4
10	during/after: mixed		□■	□■	4
11	all stages: quantitative	■	■	■	4
12	during/after: mixed		□■	■	3
13	during/after: mixed		□	■	3
14	during/after: mixed		■	□■	2
15	during/after: mixed		■	□	2
16	all stages: mixed	■	□	■	2
17	all stages: mixed	□	□■	□■	2
18	all stages: mixed	□■	■	□■	1
19	all stages: mixed	■	□■	□	1
20	all stages: mixed	■	■	□■	1
21	all stages: mixed	■	□■	□■	1
22	all stages: mixed	■	■	□	1
23	all stages: mixed	□	□	□■	1
24	before/after	■		■	1
25	before/after	■		□■	1
26	before/after	□		□■	1
Sum		22%	65%	96%	100

Table 6. Overview of the identified data type combinations in all selected publications based on the studied data type(s) per temporal stage of assessment (quantitative method(s) = ■, qualitative method(s) = □). Note: *combinations are numbered according to their frequency of occurrence, based on all relevant papers (N=100).

in which the authors combine standardized and self-developed questionnaires to evaluate UX. The other 15 papers use either self-developed or standardized questionnaires. Nevertheless, the triangulation of quantitative methods during and after the interaction represent also the second most common approach (see Table 6, no. 2). Furthermore, quantitative methods are often used in the test stage of a design process (in total 15 out of 27 publications that use approaches no. 1 or 2, during or after interaction, see Table 6, e.g., [40] or [11]) and for lab studies (in total 17 out of 27 publications that use approach no. 1 or 2, see Table 6, e.g., [2] or [3]).

Thus, as already observed in previous research [64] we can still see the strong links of many UX studies to momentary performance metrics and usability-style experiments.

Infrequent pre-/post-evaluation

Table 7 shows an emphasized interest in evaluations during and after the interaction. In particular, 96% of all studies assess UX after the product usage, whereas 65% of all studies assess the UX during the interaction. Although 22% of the relevant papers investigate UX before the actual use, only three publications from recent years pursue an expectation-focused approach focusing on the analysis of pre-use and post-use evaluation. Furthermore, 19% assess UX in all temporal stages. Uriu et al. [102], for example, conducted interviews before and after the assessed interaction plus video recording during and after the interaction. Their goal was to study the UX of a whole cooking support system. Current research emphasizes the focus on pre-use and post-use evaluation (e.g., [74]). Expectation is a key aspect of an experience, yet still rarely analyzed in academic studies.

Before	During	After	%	Examples
	♦	♦	42	[4], [7], [87]
♦	♦	♦	32	[92], [60], [88]
	♦		19	[68], [100]
♦			4	[27]
		♦	3	[66]
22%	65%	96%		

Table 7. Temporal aspects of UX evaluation: before, during and after interaction. Note: *combinations are sorted by frequency of occurrence, based on all relevant papers (N=100).

Interviews and questionnaires for post-use evaluation

Interviews and questionnaires are not only two of the most common method types that were used in all relevant publications, but also the preferred triangulated methods, as previously stated. In our analysis, we see a tendency towards post-use triangulation of interviews and questionnaires, either as a stand-alone data type triangulation (see no. 3 in Table 6, e.g., [34] or [12]) or in combination with the additional evaluation before and/or during the interaction. In total, one quarter of all analyzed publications triangulate only questionnaires and interviews to evaluate the experience afterwards.

Vermeeren et al. [107] also observed that scale-based questionnaires, often have a follow-up interview to better understand research findings. In contrast, Alves et al. [5] more recently outlined that in practice, companies prefer observation and think-aloud over questionnaires and interviews.

Tendency towards more methods for exploration

It would seem plausible that the more data we collect, the more insights we can derive. Vermeeren et al. [107] question why researchers always want more data and suggest to rather focus on suitable combinations of methods. Our analysis confirmed the tendency towards applying more methods in UX studies, as only 28% of all relevant papers base their user study on only one method. Furthermore, the studies that pursued an explorative approach (i.e., with the main goal to explore a product or prototype freely rather than evaluating a specific task) tend to be based on more methods than task-oriented studies. From all 16 publications that use 4 or more methods in their empirical study, 12 (75%) pursue an exploratory user study (e.g., [17] or [46]). When we had a closer look at the papers that only applied 1 method in their study, we saw that 11 (39%) out of 28 publications focus on Generic UX (e.g., [2] or [81]). We agree with Vermeeren et al. [107] that it is not a cause in it self to add more methods, but careful consideration of combining the right method is key. We, however, consider the tendency towards more methods to be a generally positive trend in the exploratory studies to better understand the results.

Extensive long term studies and formative evaluation

The joint evaluation of expectations, UX during the interaction, and post-use UX is usually time-consuming and costly, but pointed out as an important key understanding of user experiences [47, 70]. Based on our analysis, we can see that such holistic evaluations of UX before, during, and after the interaction are often conducted for long term studies. From all relevant publications in our analysis, 19 analyze all temporal stages whereof 13 (68%) focus on a long term evaluation.

We had assumed to find more holistic studies addressing all temporal stages of UX evaluation, since their importance and value have been highlighted before [48]. However, similar to our findings, the review of Bargas-Avila and Hornbæk [9] in 2010 highlighted 17% of papers which analyzed all temporal stages and 3% that focused their evaluation on pre-use and post-use experience.

Diverse approaches for evaluations of accumulative UX

Almost one fifth (19%) of all relevant papers analyze all temporal stages, including the experience before the interaction (i.e., expectations), during the actual use, and post-use UX. Although about half of the publications that analyze all temporal stages focus either only on quantitative methods or only on qualitative methods (see Table 6), we identified a variety of different data type triangulation approaches. While Shin et al. [95], for example, base their study on mixed data pre-use, quantitative during use and mixed data post-use, Park et al. [82] use quantitative methods before and after the interaction and qualitative methods during the interaction. Further combinations are summarized in Table 6.

DISCUSSION

During the process of writing this paper, we presented our results at a workshop [84] and discussed them with UX experts of academia and industry (N=8). This helped us to critically analyze and assess existing approaches of UX evaluation method application from a practical and non-biased perspective.

How is the UX Research Field Evolving?

While the overall number of UX studies is increasing, only a quarter of the papers make any kind of reference to UX-specific literature. This raises the following question: Is the theory of UX already taken for granted or is it too vague or unknown? We had to exclude 104 papers which used the term “User Experience” as a buzzword but did not address the topic at all (based on our understanding). Often the term was even used within the title or the author keywords but nowhere else in the paper. One of the workshop participants stated: “UX is gaining attention, everybody wants to say they do UX – even though they don’t do it”. The fact that theory was often only vaguely addressed likely had consequences on the quality of evaluation. UX evaluation still appears to struggle with the same issues as in previous meta-reviews (e.g., lack of theory, lack of validated methods, overlap with usability), and one could question whether the field is maturing or disintegrating.

Furthermore, the lack of reuse of UX-specific methods was apparent in our data. The reasons behind this could not be revealed in our empirical data, but perhaps the nature of specific experiences (e.g., of a mobile health service [10] or a navigation system [43]) may not appear to be translatable to more generic methods for approaching the evaluation topic, and the researchers turn to, for example, self-developed questionnaires rather than reusing existing, validated ones. UX is a diverse topic and hence it may be misleading to look for a “one solves it all” method, but rather choose more specific methods for the specific type of experience and/or triangulation to accomplish a useful UX evaluation.

How Can we Exploit Data-Driven Methods for UX?

Surprisingly, the data-driven and objective method “activity logging” belongs to the 4 most frequently used methods in UX studies. But is it really a valid measurement to assess users’ experiences? This was also discussed at the workshop, and one participant stated: “*Activity logging is only used to report data, but relationships are rarely investigated*”. Similarly, we found a large number of papers employing user observations, but often it was not clear how observations were analyzed and how they actually contributed to the results. Exceptions were, of course, found, which contributed to a better understanding of the experience (see, for example, [94]). We believe that there is much more work to be done in this area, to guide technological efforts that can be of value when studying UX. Dove et al. [22] write: “*It is no longer enough for UX designers to only improve user experience by paying attention to usability, utility, and interaction aesthetics.*” and suggest that much more can be done to improve UX by employing machine learning for offering new value, such as personalization of systems by learning from user interactions. We found very little efforts in this area in relation to evaluation, and look forward to further progress of data-driven methods to help us understand UX, finding patterns and relationships instead of isolated data. We also found very few studies addressing different other emerging technologies and contexts of evaluation, such as UX evaluation by, for example, AR or VR, although it appears to be on the rise (see for example [28, 33] for recent studies), as well as the UX of interacting with virtual actors (see [15, 110] for examples) and automated systems [85, 73, 26].

For data-driven methods, we believe that triangulation of qualitative insights and quantifiable measures is important, either by sequential triangulation (i.e., using qualitative data to understand identified patterns in quantitative data or the other way around [18]) or concurrent triangulation, to better grasp and validate the UX data as it is gathered.

Why are Early UX Evaluations still rare?

Michalco et al. [74] and Kujala et al. [55] suggest to study the relation between expectations and UX as expectation disconfirmations, which has a significant effect on the overall UX. However, our study revealed that as of today, only few publications investigate the relation between expectations and the post-use experience (assessment before and after the experience). The infrequent comparison between expectations and after-use evaluation is also related to the persisting lack of UX evaluation at early design process stages, although this is often claimed to be important in a design process [90, 107, 9]. We had assumed to find more of these studies, but we did not, even if the examples of methodological approaches of Wizard of Oz (i.e., a human controlling the interface to respond to a user in a test setting) used in the field [17, 31] were very informative.

We conclude that most studies are focused on tangible and complete or almost complete designs; early stage evaluation relies heavily on the imagination of the study participant and is a step researchers may be unwilling to take. In addition to this, we were interested to see if there were approaches uncovering

not only the expressible and readily available responses to an experience, but also the tacit and the latent aspects of experiences (c.f. [93]). Of all papers, 4% used extra stimuli/probes during the evaluation, such as the possibility to express the experience in video and audio material (see [68]). We think that these are interesting approaches deserving further exploration.

Parallel Analysis rather than Triangulation?

Our analysis of sequential and concurrent triangulation patterns (cf. Creswell, [18]) demonstrated the complexity and variety of UX evaluation. In this area we could see most progress in the research field, but also identified several methodological gaps. In the papers we reviewed, there were examples of well executed sequential triangulation. Sequential triangulation holds the possibility of ensuring a systematically user experience-driven process, where initial findings can be followed up by additional data for further explanation, validation or exploration [18]. As examples of studies where data was explored sequentially, Leong et al. [68] used an initial diary study as later basis for further explanation during interviews with the participants. Hart et al [34] challenged their results from questionnaires based on unexpected findings in qualitative interview data. Hayashi and Hong [41] validated their primary data source by deriving from the quantitative data that the participants had a reasonable amount of exposure of the system to evaluate it qualitatively. Results like these serve an important role in building confidence for the data validity. Triangulation of sequential exploration of themes found in initial user studies were common, such as Lederman et al. [66] who employed sequential exploration by first understanding a product space qualitatively, triangulated with the evaluation of a designed prototype. This was a more commonly applied type of triangulation, than for example validation of data points across data sources.

We found very few examples of concurrent triangulation that carefully matches quantitative with qualitative data to derive a truly joint analysis, where results can be questioned or strengthened based on correlations or the lack thereof in the data. In many papers reviewed that applied triangulation, different types of data are gathered, but rarely cross-analyzed. Many studies left us with questions whether there was not more to be learned from the data with regards to correlation or differences between different types of data. There appears to be a growing understanding in UX research that using more than one method is beneficial, but a well-grounded knowledge about how to systematically cross-analyze data appears less widespread. Actively analyzing overlaps and differences in, for example, qualitative and quantitative data, that can add to richer and better validated knowledge of the evaluated topic, is still rare. Valuable exceptions are however for example Woo et al. [111] who make connections between data points over time as well as from different formats, to strengthen the outcomes.

Many studies provide a very short and general motivation of triangulation, if any, but there are exceptions, such as Kim et al. [51] who describe the process of applying grounded theory for the understanding the nature of the experiences and numeric data of questionnaires and log data of system usage for providing descriptive statistics linked to the themes.

Ardito et al. [7] use a thorough motivation of their triangulation of data sources such as observations, questionnaires and focus groups, later cross-analyzing and making connections between the sources. Perhaps the “sad lack of reference” [10] between the quantitative and qualitative is beginning to lessen, but a widespread understanding how both approaches can contribute to each other apparently has not been accomplished yet. There could have been many more good examples of better integration of data during the analysis of results, as triangulation approaches have been claimed to lead to deeper understandings and sometimes unexpected results.

Outlook on UX Research

From our review, we conclude that there are a number of specifically open questions to be further addressed in the UX field: There is a need to provide **further guidelines and practical examples for effective combinations** of different methods, i.e., triangulation strategies in UX evaluation. These could further exemplify how results from studies can be explained either by the use of sequential triangulation, e.g., understanding a pattern found in quantitative data by employing further methods gathering qualitative data, or by concurrent triangulation, e.g., reinforcing findings with two or more types of data that are cross-analyzed. Along more theoretical descriptions, also practical examples of addressing UX from a triangulation perspective would be inspirational and serve as a palette of examples of methods to use.

More studies are required which analyze the relation between **expectations** and UX, building on the importance to satisfy user’s expectations to achieve high UX [55]. We believe that further incorporating expectations in UX evaluation studies is an interesting thread to follow up on in more evaluative papers. There is a lack of empirical papers addressing “multi touch-point” or “**multi-device**” experiences [91], although important for many products today [97], i.e., services that do not only link to one type of product and context of use. As this is an increasing part of our daily lives, it has gained interest in ACM publications (see for example [6, 16]). We also believe there is much work to be done to address the specific nature of multi-device UX. The topic also connects to multi-user environments, which were addressed in 16% of the studies. To encompass the needs of studying emerging experiences of, for example, the IoT, the palette of evaluation strategies must be expanded to encompass multi-device and multi-user experiences. The methods used for studying a single person or product in one context need to be challenged and expanded or accompanied by other approaches. Up to 2016, we found hardly any work concerning **up and coming technological approaches** such as machine learning, that, however, appears to be applicable in UX evaluation, given the possibilities of technological development [22, 98].

LIMITATIONS AND FUTURE WORK

In our review, sources outside ACM were excluded, which may have provided a bias towards approaches founded in engineering and human factors perspectives rather than, for example, a design, psychology or a marketing perspective on UX. Directing the search to, for example, more design-oriented conferences and journals will most likely provide

further insights into the state of UX evaluation. Further work could also encompass additional sources inside ACM that were now excluded, such as the NordiCHI conference series. Especially, there is a need to constantly keep the analysis up-to-date with each finished publication year, starting with 2017. UX is still a developing research field which needs to be observed continuously.

Given the substantial amount of research in the field, it would also make sense to narrow the focus of further meta-reviews. For example, it could be specific types of experience dimensions and types of data collected for these, enabling deep-dives into more isolated questions. An example of such a deep-dive is provided by Mekler et al. [71] on quantitative evaluations of enjoyment from interactive games. Even if more specificity is needed in meta-reviews, we also look forward to other reviews of where the field of UX is heading, not only in terms of academic evaluation but also, for example, in theoretical foundations and industrial practice (as only 16% of the reviewed papers had 1 or more authors with an industrial affiliation).

CONCLUSION

To analyze the current state of UX evaluation in academia, we systematically identified 280 relevant papers, out of which 100 papers were finally selected for full review. In the continuously growing number of papers over the years 2010-2016, an increasing diversity in this inherently multi-dimensional field can be found. This is, of course, an asset: a product may (at different stages of development) benefit from both “macro”- and “micro”-perspective evaluations [57]. However, we see that many of the challenges reported in earlier meta-reviews still remain, such as the weak links between theory and evaluation, little attention to expectations in UX, and a tendency towards self-defined questionnaires and post-use evaluation. Progress could be identified in the use of triangulation, by inclusion of more methods as well as a larger number of studies performed in field contexts. Thus, we can perhaps see land on the horizon, but conclude that UX evaluation currently still remains sort of a Bermuda Triangle, often depending on personal perceptions of UX rather than aggregated theory.

We see that interest and efforts in the UX field still persist, and we look forward to further work. Areas which need to be addressed are evaluation approaches to multi-device experiences, machine learning, upcoming technology for virtual experiences, and addressing expectations in UX. Method triangulation needs to be used more coherently; for stronger results in UX studies, we recommend to integrate and structure data better. For example, a well-defined structuring of observational data, improved cross-analysis of qualitative and quantitative data, and a solid definition, which aspect of UX to evaluate, will bring results forward.

REFERENCES

1. *Muhammad Adnan, Mike Just, and Lynne Baillie. 2016. Investigating time series visualisations to improve the user experience. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 5444–5455.

2. *Sarah Fdili Alaoui, Baptiste Caramiaux, Marcos Serrano, and Frédéric Bevilacqua. 2012. Movement qualities as interaction modality. In *Proceedings of the Designing Interactive Systems Conference*. ACM, 761–769.
3. *Florian Alt, Stefan Schneegass, Jonas Auda, Rufat Rzayev, and Nora Broy. 2014. Using eye-tracking to support interaction with layered 3D interfaces on stereoscopic displays. In *Proceedings of the 19th international conference on Intelligent User Interfaces*. ACM, 267–272.
4. *David Altimira, Florian Floyd Mueller, Jenny Clarke, Gun Lee, Mark Billinghurst, and Christoph Bartneck. 2016. Digitally augmenting sports: An opportunity for exploring and understanding novel balancing techniques. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 1681–1691.
5. Rui Alves, Pedro Valente, and Nuno Jardim Nunes. 2014. The state of user experience evaluation practice. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. ACM, 93–102.
6. Craig Anslow, Pedro Campos, Andrés Lucero, Laurent Grisoni, Mirjam Augstein, and James Wallace. 2016. Collaboration Meets Interactive Surfaces and Spaces (CMIS): Walls, Tables, Mobiles, and Wearables. In *Proceedings of the 2016 ACM on Interactive Surfaces and Spaces (ISS '16)*. ACM, New York, NY, USA, 505–508.
7. Carmelo Ardito, Maria F. Costabile, Antonella De Angeli, and Rosa Lanzilotti. 2012. Enriching Archaeological Parks with Contextual Sounds and Mobile Technology. *ACM Transactions on Computer-Human Interaction* 19, 4 (2012), 1–30.
8. Leena Arhippainen, Minna Pakanen, and Seamus Hickey. 2013. Mixed UX Methods Can Help to Achieve Triumphs. In *Proceedings of CHI 2013 Workshop 'Made for Sharing: HCI Stories for Transfer, Triumph and Tragedy'*. ACM, 83.
9. Javier A Bargas-Avila and Kasper Hornbæk. 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*. ACM, 2689–2698.
10. *Eric PS Baumer, Sherri Jean Katz, Jill E Freeman, Phil Adams, Amy L Gonzales, John Pollak, Daniela Retelny, Jeff Niederdeppe, Christine M Olson, and Geri K Gay. 2012. Prescriptive persuasion and open-ended social awareness: expanding the design space of mobile health. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 475–484.
11. *Luca Benedetti, Holger Winnemöller, Massimiliano Corsini, and Roberto Scopigno. 2014. Painting with bob: Assisted creativity for novices. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM, 419–428.
12. *Gilbert Beyer, Florian Alt, Jörg Müller, Albrecht Schmidt, Karsten Isakovic, Stefan Klose, Manuel Schiewe, and Ivo Haulsen. 2011. Audience behavior around large interactive cylindrical screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1021–1030.
13. Mark Blythe, Marc Hassenzahl, Effie Law, and Arnold Vermeeren. 2007. An analysis framework for user experience (UX) studies: A green paper. *Towards a UX Manifesto* (2007), 1.
14. *Amy Campbell, Christopher Wienberg, and Andrew Gordon. 2012. Collecting relevance feedback on titles and photographs in weblog posts. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces - IUI '12*. ACM Press, New York, New York, USA, 139.
15. Heloisa Candello, Claudio Pinhanez, and Flavio Figueiredo. 2017. Typefaces and the Perception of Humanness in Natural Language Chatbots. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3476–3487.
16. Marta E. Cecchinato, Anna L. Cox, and Jon Bird. 2017. Always On(Line)? User Experience of Smartwatches and Their Role Within Multi-Device Ecologies. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3557–3568.
17. *Sandy Claes, Karin Slegers, and Andrew Vande Moere. 2016. The Bicycle Barometer: Design and Evaluation of Cyclist-Specific Interaction for a Public Display. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 5824–5835.
18. John W Creswell. 2013. *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
19. Norman K Denzin. 1978. Triangulation: A case for methodological evaluation and combination. *Sociological methods* (1978), 339–357.
20. Sarah Diefenbach, Nina Kolb, and Marc Hassenzahl. 2014. The 'hedonic' in human-computer interaction: history, contributions, and future research directions. In *Proceedings of the 2014 conference on Designing interactive systems*. ACM, 305–314.
21. ISO DIS. 2009. 9241-210: 2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. *International Standardization Organization (ISO)*. Switzerland (2009).
22. Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX Design Innovation: Challenges for Working with Machine Learning As a Design Material. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 278–288.

23. Jodi Forlizzi and Katja Battarbee. 2004. Understanding experience in interactive systems. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*. ACM, 261–268.
24. *Jutta Fortmann, Erika Root, Susanne Boll, and Wilko Heuten. 2016. Tangible Apps Bracelet: Designing Modular Wrist-Worn Digital Jewellery for Multiple Purposes. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM, 841–852.
25. *Jérémy Frey, Maxime Daniel, Julien Castet, Martin Hachet, and Fabien Lotte. 2016. Framework for electroencephalography-based evaluation of user experience. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2283–2294.
26. Anna-Katharina Frison, Philipp Wintersberger, Andreas Riener, and Clemens Schartmüller. 2017. Driving Hotzenplotz: A Hybrid Interface for Vehicle Control Aiming to Maximize Pleasure in Highway Driving. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, 236–244.
27. *Chi-Wing Fu, Wooi-Boon Goh, and Junxiang Allen Ng. 2010. Multi-touch techniques for exploring large-scale 3D astrophysical simulations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2213–2222.
28. Wei Gai, Chenglei Yang, Yulong Bian, Chia Shen, Xiangxu Meng, Lu Wang, Juan Liu, Mingda Dong, Chengjie Niu, and Cheng Lin. 2017. Supporting Easy Physical-to-Virtual Creation of Mobile VR Maze Games: A New Genre. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5016–5028.
29. *Sabiha Ghellal, Ann Morrison, Marc Hassenzahl, and Benjamin Schaufler. 2014. The remediation of nosferatu: exploring transmedia experiences. In *Proceedings of the 2014 conference on Designing interactive systems*. ACM, 617–626.
30. *Nitesh Goyal and Susan R Fussell. 2016. Effects of Sensemaking Translucence on Distributed Collaborative Analysis. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 288–302.
31. *Erik Grönvall, Sofie Kinch, Marianne Graves Petersen, and Majken K Rasmussen. 2014. Causing commotion with a shape-changing bench: experiencing shape-changing interfaces in use. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 2559–2568.
32. *Jonna R Häkkinä, Maaret Posti, Stefan Schneegass, Florian Alt, Kunter Gultekin, and Albrecht Schmidt. 2014. Let me catch this!: experiencing interactive 3D cinema through collecting content with a mobile phone. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 1011–1020.
33. Daniel Harley, Aneesh P. Tarun, Daniel Germinario, and Ali Mazalek. Tangible VR: Diegetic Tangible Objects for Virtual Reality Narratives. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. ACM, New York, NY, USA, 1253–1263.
34. *Jennefer Hart, Alistair G Sutcliffe, and Antonella De Angeli. 2013. Love it or hate it!: interactivity and user types. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2059–2068.
35. Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 50.
36. Marc Hassenzahl. 2004. The interplay of beauty, goodness, and usability in interactive products. *Human-computer interaction* 19, 4 (2004), 319–349.
37. Marc Hassenzahl. 2005. The thing and I: understanding the relationship between user and product. *Funology* (2005), 31–42.
38. Marc Hassenzahl. 2008. User experience (UX): towards an experiential perspective on product quality. In *Proceedings of the 20th Conference on l'Interaction Homme-Machine*. ACM, 11–15.
39. Marc Hassenzahl, Michael Burmester, and Franz Koller. 2003. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Mensch & Computer 2003*. Springer, 187–196.
40. *Renate Häuslschmid, Sven Osterwald, Marcus Lang, and Andreas Butz. 2015. Augmenting the Driver's View with Peripheral Information on a Windshield Display. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*. ACM, 311–321.
41. *Eiji Hayashi and Jason I Hong. 2015. Knock x knock: the design and evaluation of a unified authentication management system. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 379–389.
42. *Eva Hornecker. 2016. The To-and-Fro of Sense Making: Supporting Users' Active Indexing in Museums. *ACM Transactions on Computer-Human Interaction (TOCHI)* 23, 2 (2016), 10.
43. *Shaohan Hu, Lu Su, Shen Li, Shiguang Wang, Chenji Pan, Siyu Gu, Md Tanvir Al Amin, Hengchang Liu, Suman Nath, Romit Roy Choudhury, and others. 2015. Experiences with eNav: A low-power vehicular navigation system. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 433–444.

44. R Burke Johnson, Anthony J Onwuegbuzie, and Lisa A Turner. 2007. Toward a definition of mixed methods research. *Journal of mixed methods research* 1, 2 (2007), 112–133.
45. *Jinyung Jung, Seok-Hyung Bae, and Myung-Suk Kim. 2013. Three case studies of UX with moving products. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 509–518.
46. Mohsen Kamalzadeh, Christoph Kralj, Torsten Möller, and Michael Sedlmair. 2016. TagFlip: Active Mobile Music Discovery with Social Tags. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*. ACM, 19–30.
47. Evangelos Karapanos, John Zimmerman, Jodi Forlizzi, and Jean-Bernard Martens. 2009. User experience over time: an initial framework. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 729–738.
48. Evangelos Karapanos, John Zimmerman, Jodi Forlizzi, and Jean Bernard Martens. 2010. Measuring the dynamics of remembered experience over time. *Interacting with Computers* 22, 5 (2010), 328–335.
49. *Tuula Kärkkäinen, Tuomas Vaittinen, and Kaisa Väänänen-Vainio-Mattila. 2010. I don't mind being logged, but want to remain in control: a field study of mobile activity and context logging. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 163–172.
50. Joseph 'Jofish' Kaye. 2007. Evaluating experience-focused HCI. In *Proceedings of CHI 2007 Extended Abstracts on Human factors in computing systems*. ACM, 1661–1664.
51. *Tanyoung Kim, Hwajung Hong, and Brian Magerko. 2010. Design requirements for ambient display that supports sustainable lifestyle. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems*. ACM, 103–112.
52. *Martin Knobel, Marc Hassenzahl, Melanie Lamara, Tobias Sattler, Josef Schumann, Kai Eckoldt, and Andreas Butz. 2012. Clique Trip: feeling related in different cars. In *Proceedings of the designing interactive systems conference*. ACM, 29–37.
53. Klaus Krippendorff. 2004. *Content Analysis: An Introduction to Its Methodology* (second ed.). Sage.
54. *Sari Kujala and Talya Miron-Shatz. 2013. Emotions, experiences and usability in real-life mobile phone use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1061–1070.
55. Sari Kujala and Talya Miron-Shatz. 2015. The Evolving Role of Expectations in Long-term User Experience. In *Proceedings of the 19th International Academic Mindtrek Conference (AcademicMindTrek '15)*. ACM, New York, NY, USA, 167–174.
56. Sari Kujala, Virpi Roto, Kaisa Väänänen-Vainio-Mattila, Evangelos Karapanos, and Arto Sinnelä. 2011. UX Curve: A method for evaluating long-term user experience. *Interacting with Computers* 23, 5 (2011), 473–483.
57. Florian Lachner, Philipp Nägelein, Robert Kowalski, Martin Spann, and Andreas Butz. 2016. Quantified UX : Towards a Common Organizational Understanding of User Experience. In *Proc. NordiCHI 2016*.
58. Carine Lallemand. 2015. *Towards consolidated methods for the design and evaluation of user experience*. Ph.D. Dissertation. University of Luxembourg, Luxembourg.
59. *Carine Lallemand and Guillaume Gronier. 2012. Enhancing User eXperience during waiting time in HCI: contributions of cognitive psychology. In *Proceedings of the Designing Interactive Systems Conference*. ACM, 751–760.
60. *Airi Lampinen, Vilma Lehtinen, Coye Cheshire, and Emmi Suhonen. 2013. Indebtedness and reciprocity in local online exchange. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 661–672.
61. Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*. Springer, 63–76.
62. Effie Law, Virpi Roto, Arnold POS Vermeeren, Joke Kort, and Marc Hassenzahl. 2008. Towards a shared definition of user experience. In *CHI'08 extended abstracts on Human factors in computing systems*. ACM, 2395–2398.
63. Effie Lai-Chong Law. 2011. The measurability and predictability of user experience. In *Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems*. ACM, 1–10.
64. Effie Lai-Chong Law, Virpi Roto, Marc Hassenzahl, Arnold POS Vermeeren, and Joke Kort. 2009. Understanding, scoping and defining user experience: a survey approach. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 719–728.
65. Effie Lai-Chong Law, Paul van Schaik, and Virpi Roto. 2014. Attitudes towards user experience (UX) measurement. *International Journal of Human-Computer Studies* 72, 6 (2014), 526–541.
66. *Reeva Lederman, Greg Wadley, John Gleeson, Sarah Bendall, and Mario Álvarez-Jiménez. 2014. Moderated online social therapy: Designing and evaluating technology for mental health. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 1 (2014), 5.
67. *Yong-Ki Lee, Youn-kyung Lim, and Kunpyo Lee. 2016. Timelessness. *Proceedings of the 2016 ACM Conference on Designing Interactive Systems - DIS '16* (2016), 73–83.

68. Tuck W Leong, Frank Vetere, and Steve Howard. 2012. Experiencing coincidence during digital music listening. *ACM Transactions on Computer-Human Interaction (TOCHI)* 19, 1 (2012), 6.
69. *Craig M MacDonald and Michael E Atwood. 2014. What does it mean for a system to be useful?: An exploratory study of usefulness. In *Proceedings of the 2014 conference on Designing interactive systems*. ACM, 885–894.
70. John McCarthy and Peter Wright. 2004. Technology as experience. *interactions* 11, 5 (2004), 42–43.
71. Elisa D Mekler, Julia Ayumi Bopp, Alexandre N Tuch, and Klaus Opwis. 2014. A systematic review of quantitative studies on the enjoyment of digital entertainment games. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 927–936.
72. Elisa D Mekler and Kasper Hornbæk. 2016. Momentary Pleasure or Lasting Meaning?: Distinguishing Eudaimonic and Hedonic User Experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 4509–4520.
73. Thomas Meneweger. 2017. Experiencing Automation: UX and Practices with Automated Systems. In *Proceedings of the 2017 ACM Conference Companion Publication on Designing Interactive Systems (DIS '17 Companion)*. ACM, New York, NY, USA, 407–408.
74. Jaroslav Michalco, Jakob Grue Simonsen, and Kasper Hornbæk. 2015. An Exploration of the Relation between Expectations and User Experience. *International Journal of Human-Computer Interaction* 7318, July (2015), 150701125618002.
75. *Hannah J Miller, Shuo Chang, and Loren G Terveen. 2015. I LOVE THIS SITE! vs. It's a little girly: Perceptions of and Initial User Experience with Pinterest. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 1728–1740.
76. *Pejman Mirza-Babaei, Lennart E Nacke, John Gregory, Nick Collins, and Geraldine Fitzpatrick. 2013. How does it play better?: exploring user testing and biometric storyboards in games user research. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1499–1508.
77. David Moher, Deborah J Cook, Susan Eastwood, Ingram Olkin, Drummond Rennie, and Donna F Stroup. 1999. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet* 354, 9193 (1999), 1896–1900.
78. *Timothy Neate, Michael Evans, and Matt Jones. 2016. Designing visual complexity for dual-screen media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 475–486.
79. *Marianna Obrist, Rob Comber, Sriram Subramanian, Betina Piqueras-Fiszman, Carlos Velasco, and Charles Spence. 2014. Temporal, affective, and embodied characteristics of taste experiences: A framework for design. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 2853–2862.
80. Marianna Obrist, Virpi Roto, and Kaisa Väänänen-Vainio-Mattila. 2009. User experience evaluation: do you know which method to use?. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2763–2766.
81. *Thomas Olsson and Markus Salo. 2012. Narratives of satisfying and unsatisfying experiences of current mobile augmented reality applications. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2779–2788.
82. *S. Joon Park, Craig M MacDonald, and Michael Khoo. 2012. Do you care if a computer says sorry?: user experience design through affective messages. In *Proceedings of the Designing Interactive Systems Conference*. ACM, 731–740.
83. *Denis Parra, Peter Brusilovsky, and Christoph Trattner. 2014. See what you want to see: visual user-driven approach for hybrid recommendation. In *Proceedings of the 19th international conference on Intelligent User Interfaces*. ACM, 235–240.
84. Ingrid Pettersson, Anna-Katharina Frison, Florian Lachner, Andreas Rieger, and Jesper Nollhage. 2017. Triangulation in UX Studies: Learning from Experience. In *Proceedings of the 2016 ACM Conference Companion Publication on Designing Interactive Systems*. ACM, 341–344.
85. Ingrid Pettersson and Wendy Ju. 2017. Design Techniques for Exploring Automotive Interaction in the Drive towards Automation. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. ACM, 147–160.
86. *Erika Shehan Poole. 2012. Interacting with infrastructure: a case for breaching experiments in home computing research. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 759–768.
87. Katharina Reinecke and Abraham Bernstein. 2011. Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)* 18, 2 (2011), 8.
88. *Stefan Rennick-Egglestone, Sarah Knowles, Gill Toms, Penny Bee, Karina Lovell, and Peter Bower. 2016. Health Technologies' In the Wild': Experiences of Engagement with Computerised CBT. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2124–2135.

89. *Shyam Reyal, Shumin Zhai, and Per Ola Kristensson. 2015. Performance and user experience of touchscreen and gesture keyboards in a lab setting and in the wild. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 679–688.
90. Virpi Roto, Kaisa Väänänen-Vainio-Mattila, Effie Law, and Arnold Vermeeren. 2009. User experience evaluation methods in product development (UXEM'09). In *IFIP Conference on Human-Computer Interaction*. Springer, 981–982.
91. Virpi Roto, Heli Väättäjä, Effie Law, and Rachel Powers. 2016. Experience Design for Multiple Customer Touchpoints. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordicCHI '16)*. ACM, New York, NY, USA, Article 146, 3 pages.
92. *Yea-Kyung Row and Tek-Jin Nam. 2014. CAMY: applying a pet dog analogy to everyday ubicomp products. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 63–74.
93. Elizabeth B-N Sanders. 2002. From user-centered to participatory design approaches. *Design and the social sciences: Making connections* 1, 8 (2002).
94. *Magy Seif El-Nasr, Bardia Aghabeigi, David Milam, Mona Erfani, Beth Lameman, Hamid Maygoli, and Sang Mah. 2010. Understanding and Evaluating Cooperative Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 253–262.
95. *Jaemyung Shin, Bumsoo Kang, Taiwoo Park, Jina Huh, Jinhan Kim, and Juneha Song. 2016. BeUpright: Posture Correction Using Relational Norm Intervention. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 6040–6052.
96. *Johanna M Silvennoinen and Jussi PP Jokinen. 2016. Aesthetic Appeal and Visual Usability in Four Icon Design Eras. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 4390–4400.
97. Adalberto L Simeone, Ming Ki Chong, Corina Sas, and Hans Gellersen. 2015. Select & Apply: understanding how users act upon objects across devices. *Personal and Ubiquitous Computing* 19, 5-6 (2015), 881–896.
98. Samaneh Soleimani and Effie Lai-Chong Law. 2017. What Can Self-Reports and Acoustic Data Analyses on Emotions Tell Us?. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. ACM, New York, NY, USA, 489–501.
99. *Jacopo Staiano, María Menéndez, Alberto Battocchi, Antonella De Angeli, and Nicu Sebe. 2012. UX_Mate: from facial expressions to UX evaluation. In *Proceedings of the Designing Interactive Systems Conference*. ACM, 741–750.
100. *S Shyam Sundar, Jeeyun Oh, Saraswathi Bellur, Haiyan Jia, and Hyang-Sook Kim. 2012. Interactivity as self-expression: a field experiment with customization and blogging. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 395–404.
101. Alexandre N Tuch, Rune Trusell, and Kasper Hornbæk. 2013. Analyzing users' narratives to understand experience with interactive products. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2079–2088.
102. *Daisuke Uriu, Mizuki Namai, Satoru Tokuhisa, Ryo Kashiwagi, Masahiko Inami, and Naohito Okude. 2012. Panavi: recipe medium with a sensors-embedded pan for domestic users to master professional culinary arts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 129–138.
103. Kaisa Väänänen-Vainio-Mattila, Virpi Roto, and Marc Hassenzahl. 2008. Towards practical user experience evaluation methods. *Meaningful measures: Valid useful user experience measurement (VUUM)* (2008), 19–22.
104. Koen van Turnhout, Arthur Bennis, Sabine Craenmehr, Robert Holwerda, Marjolein Jacobs, Ralph Niels, Lambert Zaad, Stijn Hoppenbrouwers, Dick Lenior, and René Bakker. 2014. Design patterns for mixed-method research in HCI. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. ACM, 361–370.
105. *Aditya Vashistha, Edward Cutrell, Gaetano Borriello, and William Thies. 2015. Sangeet swara: A community-moderated voice forum in rural india. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 417–426.
106. *Yolanda Vazquez-Alvarez, Matthew P Aylett, Stephen A Brewster, Rocio Von Jungenfeld, and Antti Virolainen. 2016. Designing Interactions with Multilevel Auditory Displays in Mobile Audio-Augmented Reality. *ACM Transactions on Computer-Human Interaction (TOCHI)* 23, 1 (2016), 3.
107. Arnold POS Vermeeren, Effie Lai-Chong Law, Virpi Roto, Marianna Obrist, Jettie Hoonhout, and Kaisa Väänänen-Vainio-Mattila. 2010. User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*. ACM, 521–530.
108. Froukje Sleeswijk Visser, Pieter Jan Stappers, Remko Van der Lugt, and Elizabeth BN Sanders. 2005. Contextmapping: experiences from practice. *CoDesign* 1, 2 (2005), 119–149.
109. *Qing Wang and Huiyou Chang. 2010. Multitasking bar: prototype and evaluation of introducing the task concept into a browser. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 103–112.

110. Rina R. Wehbe, Edward Lank, and Lennart E. Nacke. 2017. Left Them 4 Dead: Perception of Humans Versus Non-Player Character Teammates in Cooperative Gameplay. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. ACM, New York, NY, USA, 403–415.
111. *Jong-bum Woo and Youn-kyung Lim. 2015. User experience in do-it-yourself-style smart homes. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. ACM, 779–790.

*Note: Papers which were analyzed in this paper and used as examples in the text are highlighted in the references with *.*
A full list of analyzed papers can be reviewed under: <https://uxtriangulation.wordpress.com/>



PRIVACY AND PERSONALIZATION: THE STORY OF A CROSS-CULTURAL FIELD STUDY



Hanna Schneider, LMU Munich
Florian Lachner, LMU Munich
Malin Eiband, LMU Munich
Ceenu George, LMU Munich
Purvish Shah, IIT Guwahati
Chinmay Parab, Microsoft Research
Anjali Kukreja, HealthKart
Heinrich Hussmann, LMU Munich
Andreas Butz, LMU Munich

Insights

- In collaborative field research, very different perspectives might meet and challenge each other.
- It takes close listening, open-mindedness, good preparation, and improvisation to set up a field trip in an unfamiliar context. Close collaboration with locals is invaluable.
- Despite all preparation, be ready for surprises—and welcome them.

In September 2017, we had the chance to conduct a two-day field trip in Mumbai, India. This field trip provided us with a fresh perspective, inspiration, and great human encounters. With this article, we aim to provide insights and motivation, as well as details on planning and conducting an intercultural field trip for researchers considering a similar endeavor. We recount our experiences from the perspectives of both German and Indian participants, from the initial idea to planning and conducting the field trip in collaboration with locals. We also share our lessons and recommendations.

HOW IT ALL STARTED

*December 2016,
10 months before the field trip*
Our identities are complemented by a virtual counterpart, consisting of diverse social network profiles and a data footprint most of us leave when we are online. This accumulation of personal information is used to tailor the content or functionality of Web pages to our preferences, interests, knowledge, and other personal traits. The concept behind this adaptation, or *personalization*, is simple: Messages, objects, and experiences that we attribute to our self affirm our identity and differentiate us from others. As a

consequence, they seem generally more appealing, more valuable [1], and more persuasive [2]. Personalization makes use of these effects by reinforcing a sense of uniqueness through the interface. This sense of “me-ness” is appreciated by users [3]; moreover, they may benefit from information that is more relevant to them or from functionality that better suits their needs. However, personalization is often done *implicitly* by algorithms. Therefore, users might not understand *why* and *how*—or even might not be consciously aware *that*—something was adapted to them that may potentially infringe on their privacy.

To explore individual perceptions of privacy and personalization, we conducted interviews in Germany. Typical reactions to personalized content were Netflix recommendations rated positively and advertisements often seen negatively. The personalization of the Facebook news feed and Google search was met with both appreciation and skepticism. Even though most users enjoyed the benefits of personalization in certain contexts, they were left with some sense of unease. They also wondered where their identities were stored and traced on the Web when they could neither tell nor control what data was collected and how it was processed.

—Hanna, Florian, Malin, Ceenu, Heinrich, and Andreas
Researchers, Germany

Privacy concerns and the wish to control one’s online identity came up in most of our interviews. It was the consistency of these findings that made me wonder about our German cultural perspective, as they did not match my experiences at home. Growing up in Europe with Indian parents means having two personas: The first one is studying, working, and socializing in an individualistic culture; the other one comes to life at home, where the Indian, interdependent culture is maintained. Simply said, at home everything has to be shared and all information is public, while at work I was part of a team that investigates the various facets of privacy in HCI. When we discussed this among our research team, we came to the understanding that in order to validate our research into privacy and personalization, we needed to review it in another culture. Around this time,



Team of field-trip participants and organizers discussing first interview results.

we came across the opportunity to conduct a field trip in the context of the INTERACT 2017 conference taking place in Mumbai.

—Ceenu
Researcher, Germany

DESK RESEARCH

June 2017, four months before the field trip
We began with a closer look at cultural theory. In HCI literature, cross-cultural researchers commonly base their analyses on cultural dimensions. However, these cultural dimensions were defined before current digital technologies began to affect our lives as they do now. As a result, the perceptions of online privacy and personalization in different cultures are not yet represented in established cultural theory. Nevertheless, many well-researched dimensions are related to online privacy and personalization: For example, tolerance for uncertainty, equality of power distribution, and the individual or collectivist orientation of interpersonal relationships [4]; but also the physical distance that is perceived as comfortable [5], as well as the separation of private and public life [6]. Based on the study of these dimensions, one might, for example, hypothesize that individualistic cultures value personalized products or services more than collectivist cultures, or that online privacy concerns are higher in cultures that value offline privacy more as well. However, to truly understand how we can best encourage a culturally sensitive design process for personalized products or services, we would clearly

need a more in-depth understanding, as emphasized by researchers [4]. Research in cross-cultural settings, however, is a challenging endeavor. Physical distance hampers cross-cultural studies in general, and the researchers’ cultural biases can easily influence the interpretation of insights and results. The concept of a field trip allowed us not only to conduct research abroad under the guidance of local experts but also to jointly gather and interpret insights with researchers and practitioners from different cultures and backgrounds.

—Hanna, Florian, Malin, Ceenu, Heinrich, and Andreas
Researchers, Germany

SIGNING UP AND PREPARING THE TRIP

June 2017, four months before the field trip
I first read about privacy and personalization on the Internet when there were problems related to it in European countries. I was curious to explore this further and thus did more reading about the topic. I talked to my friends and relatives about personalization and, to my surprise, only a few people actually understood the term here in India. When I read about the related field trip on the INTERACT website, I thought it would be an interesting opportunity to delve deeper, and I was eager to be part of it.

—Purvish
Field-trip participant,
Department of Design,
IIT Guwahati

SPECIAL TOPIC

India has always been a culturally rich country. We have a plethora of diversity, yet that diversity is not celebrated. We have an inferiority complex, an undying thirst to copy Western culture. This, of course, is reflected in the products we create (both online and physical). Instead of inventing creative solutions to the local problems that our communities face, we are just trying to copy and paste solutions from across the world. At my organization, for example, a healthcare e-commerce firm, we wanted to create a personalized experience for our users. But personalization often comes at the cost of privacy. My purpose of attending the field trip was to get an understanding of how online privacy is being perceived in India; I felt that this field trip would give me a perfect start toward understanding this issue.

—Anjali Kukreja
Delhi, India

When we started planning a field trip on perceptions of privacy and personalization in Mumbai, our first approach was to conduct our studies just as we would have done in Germany, using semi-structured interviews, observation, and a drawing task. In the drawing task, study participants were asked to sketch how products or services (e.g., a social media app) could personalize their content. Our experience was that this set of methods would allow us to holistically tackle our research question.

However, we had read that ethnographic studies by nature are dependent on both contextual knowledge and a well-developed understanding of local peculiarities. We therefore decided to obtain feedback from the field-trip organizers, as we wanted to make sure that our study setting was not biased by our own culturally influenced assumptions. After several feedback loops with them, we realized that some of our questions had

assumed a certain understanding of the concepts of privacy and personalization that should not be expected.

We hence adjusted our interview guidelines to the local context and revisited culturally sensitive ethical considerations, for example, the proper wording for our consent forms. The organizers supported us tremendously by recruiting suitable participants in Mumbai, something we would not have had a chance to do remotely. They further helped by planning and scheduling logistics to suitable study locations and contacting local NGOs that supported the field trip.

—Hanna, Florian, Malin,
Ceemu, Heinrich, and Andreas
Researchers, Germany

HITTING THE ROAD

September 2017,
field trip

Before we began our research, I already had an idea about the type of responses we would get. It was a great challenge for me to not let this bias the interview process. I was driven by a vision to address the issue and create awareness about personalization in India; this field trip was my first step. The day when we were introduced to our fellow participants, I realized I was the youngest of them all, with absolutely no experience. I became nervous, but at the same time I was excited to “interact” with and learn from the professionals from Facebook and Google. They not only played a mentoring role for me but also were open to ideas from my side. Being a local, I was an important link for smooth communication. In total, four among the field-trip participants were locals, so we were able to split up into several groups to conduct more interviews.

—Purvish
Field-trip participant,
Department of Design,
IIT Guwahati

Having arrived the day before, we had already realized that traffic works a bit differently from what we were used to, so we gladly agreed to take a shuttle from the conference hotel to the conference venue on the first day of the field trip. We quickly met the eight other researchers and practitioners (from India, the U.K., and the U.S.) in a special meeting area for all field-trip participants and organizers.

That first day, we interviewed people in Powai, one of the wealthier areas in Mumbai. At the beginning it was quite challenging to approach people on the street and ask if they were willing to take part in our study. Luckily, four local students joined our field trip and hence spoke the language. Additionally, as the field trip was hosted in the name of the INTERACT 2017 conference at IIT Bombay, we had implicit support from a trustworthy, well-known local institution. When it came to providing personal data, most of the eight study participants whom we interviewed that day showed fewer concerns than our earlier subjects in Germany. One example: their readiness to provide personal data for the Indian national ID, the Adahar Card, which stores fingerprints and includes an iris scan. A much bigger concern of our participants was attacks from scammers and hackers—which none of our German participants had mentioned.

On the second day, we traveled to the lower-income area of Dharavi. With the help of a local NGO, we had seven prescheduled interviews at a public building to avoid invading personal spaces in the dense urban area. We observed that social media apps were used extensively and, again, that the main source of concern was attacks from hackers and scammers. Especially for women, uploading pictures was considered unsafe, a statement that we received from both male and female participants. As devices and passcodes were often shared, some participants recounted deleting their chat histories regularly to prevent family members, friends, or hackers from reading along. Implicit data collection, for example through tracking browsing behavior, was not associated with personalized content and thus not among the participants’ concerns, in contrast to our observations in Germany.

—Hanna, Florian, Malin,
Ceemu, Heinrich, and Andreas
German Researchers, now in Mumbai

When it came to providing personal data, most of the eight study participants whom we interviewed that day showed fewer concerns than our earlier subjects in Germany.

An interesting belief we observed among the participants was that social media apps, on which they spend a major part of their time, were not hosted by companies that generate revenue from users' online activities. Participants did not have a very clear understanding of how data was being generated by their activities on social media. Sharing images, liking posts, or browsing particular websites was not data that was relevant enough to be misused. Unless the information was being asked for explicitly (like passwords or emails), no other data was considered as being recorded. One participant said, "All my data information is with Mark Zuckerberg, and the government is paying Facebook for each account a citizen created."

Another point that came across in the discussions with the participants was the fact that when the information sharing was explicit, participants felt reluctant to share even the basic details; however, when it was hidden and not perceptible on the surface, participants did not seem to mind sharing their details at all. This was a great takeaway for my project at my workplace: As personalization in the healthcare industry involves a lot of information from the participants, the mode in which the data is gathered should be educational as well as secure for users.

—Anjali Kukreja
Delhi, India

REFLECTION

November 2017,

two months after the field trip

Even though—or maybe just because—so many aspects of our field trip in Mumbai turned out differently from what we initially expected, it was a great experience and source of inspiration. We will try to repeat this form of research and to expand it to new contexts. Some lessons and best practices will guide us and hopefully others along the way:

- Seek close collaboration with locals or NGOs who have deeper insights into the culture and situational conditions; ideally get in touch with local participants or SVs as early as possible.
- Learn about technological products and services that potential participants might be able to relate to via desk research or discussions with locals.
- Discuss methods and questions with your local contacts; ideally pilot your



Participants expressing their perception of personalization through drawing and thinking aloud.

questions; prepare alternative or back-up questions.

- Give up on the idea of getting everything right: Despite all preparations, you will likely discover that your questions are not understood as expected. This will be fine, however, as you are ultimately doing field trips because you want to be surprised.

We felt that the format and structure offered by INTERACT 2017 combined all the points mentioned here for conducting research in a sensitive, respectful way while providing the support needed. We highly recommend this format for inclusion in other conferences, and encourage other researchers to make use of any opportunity to conduct intercultural field research. From our experience, the step out of the comfort zone that is required to conduct research in different cultures more than pays off: We gained a different perspective on our research topic, feedback on the generalizability of design ideas and potential solutions, inspiration for new research directions, and, last but not least, we met great and inspiring people.

Of course, conducting research in a different culture also means overcoming hesitations and uncertainties and taking a more vulnerable role: During our field trip we depended on the expertise and willingness to help of local participants and fellow local researchers. Students knowing the local reality in our field trip contributed just as much or even more to its success as our own expertise and that of other experienced HCI researchers in our team. When you sign up for a field trip, the package includes

surprises, misunderstandings, and unpredictable situations. But if you are willing to improvise and take the plunge, you will most likely have a great experience.

—Hanna, Florian, Malin,
Ceenu, Heinrich, and Andreas
Researchers, Germany

ENDNOTES

1. Kahneman, D., Knetsch, J.L., and Thaler, R.H. Anomalies: The endowment effect, loss aversion, and status quo bias. *The Journal of Economic Perspectives* 5, 1 (1991), 193–206.
2. Petty, R.E., Wheeler, S.C., and Bizer, G.Y. Attitude functions and persuasion: An elaboration likelihood approach to matched versus mismatched messages. In *Why We Evaluate: Functions of Attitudes*. G.R. Maio and J.M. Olson, eds. Erlbaum, Mahwah, NJ, 2000, 133–162.
3. Kalyanaraman, S. and Sundar, S.S. The psychological appeal of personalized content in web portals: Does customization affect attitudes and behavior? *Journal of Communication* 56, 1 (2006), 110–132.
4. Hofstede, G., Hofstede G.J., and Minkov, M. *Cultures and Organizations: Software of the Mind*. McGraw-Hill, New York, 2010.
5. Hall, E.T. *Beyond Culture*. Doubleday, New York, 1989.
6. Trompenaars, F. and Hampden-Turner, C. *Riding the Waves of Culture. Understanding Cultural Diversity in Business*. 2nd Edition. Nicholas Brealey, London, 1998.

🔗 Hanna Schneider, Malin Eiband, Ceenu George, and Florian Lachner are Ph.D. students at LMU Munich focusing on privacy, personalization, and culturally sensitive user experience design, supervised by Heinrich Hussmann and Andreas Butz. Purvish Shah (IIT Guwahati), Chinmay Parab (Microsoft Research), and Anjali Kukreja (HealthKart) conduct studies and research on privacy and personalization in India.
→ hanna.schneider@ifi.lmu.de

Nurturing Empathy between UX Design Teams and Users in Digitally-Mediated User Research

Florian Lachner
LMU Munich
Munich, Germany
florian.lachner@ifi.lmu.de

Hanna Schneider
LMU Munich
Munich, Germany
hanna.schneider@ifi.lmu.de

Lisa Simon
LMU Munich
Munich, Germany
simon.lisa@campus.lmu.de

Andreas Butz
LMU Munich
Munich, Germany
butz@ifi.lmu.de

ABSTRACT

Creating an empathic, holistic understanding of the user experience and communicating it within the design team is a constant challenge in UX design projects. This paper explores the potential of digital tools to support designers and researchers in this task. We explored the needs of different stakeholders in semi-structured interviews and hosted an ideation workshop to generate design ideas for suitable software tools. Based on the resulting insights and ideas, we implemented a first prototype that balances individual feedback visualizations with detailed user profiles, a user journey and a communication feature. The prototype was assessed in seven focus groups with a total of 26 participants and with the AttrakDiff questionnaire. We found out that the persona view, the user journey view and the flexible filters of our prototype allowed designers to gain a insightful picture of users' experiences. Future work is needed to better understand how digitally-mediated empathic relationships evolve over the long term.

Author Keywords

User experience; evaluation; empathy; participatory design; interdisciplinary teams

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

EMPATHY IN UX RESEARCH

"Understanding the user", or user research, has been widely accepted as critical for good design by interaction designers and researchers alike. Time- and labour-intensive research methods, such as ethnography-inspired field work, diary studies, and cultural probes are arguably best suited to facilitate

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

NordiCHI'18, September 29–October 3, 2018, Oslo, Norway

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-6437-9/18/09...\$15.00

DOI: <https://doi.org/10.1145/3240167.3240182>

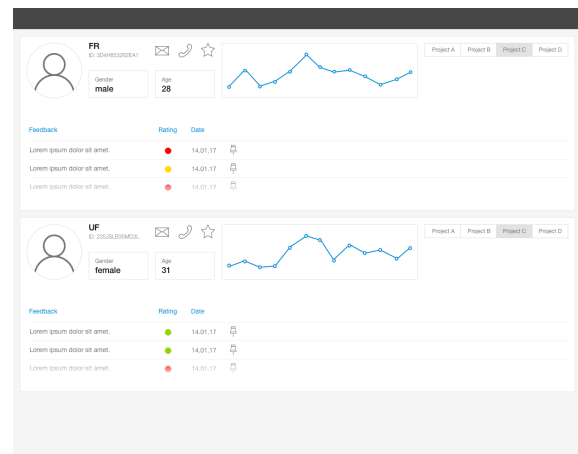


Figure 1. Mockup of our UX evaluation prototype showing two user profiles including contact details to allow user-researcher interactions.

an empathic relationship between designers and participants, often seen as the goal of user-centered design [56].

However, in the dilemma of staying within time and budget and living up to their own standards of user research, designers are often pushed towards cheaper and more practicable methods. To use available resources effectively, they are more and more relying on digital User Experience (UX) evaluation tools that collect, store and/or analyze UX questionnaire, interview, and logging data such as the commercial tools NVivo¹ or Dedoose². While these tools are a promising addition in the UX tool kit, this paper aims to explore how such tools can facilitate empathic relationships with users (e.g., see Figure 1).

Empathy has been described as vital to design practice [45, 52]. It "means that people are seen and understood from where they stand, not as test subjects but as persons with feelings" and it requires "personal contact and connection with the

¹<https://qsrinternational.com>

²<http://dedoose.com>

users" [37]. Moreover, the importance of empathy in design is increasing with growing ambitions of designers and growing expectations of users [37]: Instead of ensuring usability, new goals of interaction design include agency, fulfillment, joy, connectedness, empowerment, and social justice [2, 21, 22]. According to Mattelmäki and Battarbee [37] empathy is especially important when designers aim to go beyond designing for practical functions and focus on personal experiences and private contexts.

In this paper, we aim to investigate if and how empathic relationships can be created via digital UX evaluation platforms:

RQ "How can digital communication platforms help to create an empathic relationship between users and designers?"

In this paper, we derive the understanding of empathic relationships for UX design from Koskinen et al. [40], who define empathic design as an approach that helps designers to get closer to users, their environment, and their experiences. We report results from (1) nine interviews with designers and further stakeholders to understand how their methods create empathy, the constraints they face and the needs that are currently unmet by digital tools, (2) an ideation session with design stakeholders to explore the design space of digital tools to foster empathy between users and designers (3) and an implementation and evaluation of a first prototype. Throughout the project, we stayed in close contact with the interaction design community in our city and partnered with a design agency to truly understand our user group: interaction designers.

Our paper offers two main contributions: First, we explore possibilities to design digital user research tools in ways that facilitate empathy. Second, we implement and evaluate a prototype with several features to support empathy in the design process. From these experiences we draw conclusions for future research and design recommendations.

THE STATUS QUO OF THE UX DESIGN PROCESS

The term UX is often used as an umbrella term to describe the quality of use of interactive systems [24, 35]. Definitions range from a focus on psychological aspects to task-oriented user goals [1]. To holistically address our research question we base our understanding of UX in this work on the integrated view on pragmatic and hedonic qualities that Hassenzahl [24] postulated since the early days of UX considerations.

The fields of gathering, analyzing, and communicating UX feedback are diverse yet intertwined research areas with different perspectives and opinions. Below, we discuss current developments with a focus on empathic designer-user relationships. We present how UX more and more becomes an interdisciplinary topic of interest, summarize the ongoing discourse about the role of qualitative and quantitative data as well as the role of the user in UX evaluation and present established digital tools and methods for designers that are currently on the market.

The multidisciplinary endeavor of UX

In traditional organizational settings of technology-oriented companies, engineering departments were responsible for the technical development and implementation of their products

and ensured product success based on usability requirements. The communication and consideration of experiential factors was in the hands of the marketing department [53]. Nowadays, product life-cycles are getting shorter, development cycles move increasingly fast, and usability is taken for granted [44]. As a consequence, to guarantee a successful transformation from a design idea to a shipping product, the design must be understood and exemplified by many people in an organization [10]. More precisely, diverse backgrounds and roles are now involved in UX design processes, adding valuable but at the same time sometimes contradicting viewpoints.

Today there is a variety of stakeholders with diverse perspectives involved in the creation of a product's UX [1, 33, 27], ranging from user research, interaction design, product management, engineering, graphics design, and more. Despite differences in mindset, needs, and interests, people from diverse disciplines need to jointly work together to ensure a consistent UX [46]. Roto et al. [47] go even one step further and emphasize the importance of company-wide experience goals. However, due to time and budget constraints it is generally not feasible that everybody within an organization builds a relationship with users to profoundly understand their needs and concerns to ultimately ensure a suitable UX.

The role of data in UX evaluation

Measuring and evaluating the holistic UX in general is a major challenge for HCI researchers and hence intensively discussed in both academia and industry [1, 35]. A whole spectrum of design research methods exists, as evidenced by the views of both Roto et al. [48] and Pettersson et al. [43]. The associated debate about qualitative and quantitative data naturally evolved since the early stages of UX, when epistemological assumptions and theoretical concepts from different disciplines started to participate in the discourse about experience-related product aspects beyond usability [36]. Cooper et al. [10] note that the belief that only quantitative data yields objective and true results neglects the fact that numbers describing human experiences are subject to interpretation and tend to over-simplify. In addition, while every method has its advantages, hardly any single method can simultaneously meet practitioners' needs for practicability, understanding context of use, and scientific quality [54].

Vermeeren et al. [54] and Law [34] already recommended a while ago to pursue mixed method approaches or triangulation for UX evaluation. However, the debate about UX evaluation still centers around the legitimacy of quantitative and qualitative approaches. Whereas quantitative measures are generally favoured for communicating product success and benchmarking [33], qualitative approaches provide in-depth insights that allow designers to draw conclusions for product design [5, 7]. Dubberly et al. [14], for example, highlight that qualitative insights help designers to interpret the currently researched situation (i.e., analysis) to move towards the imagined future situation of a prototype or product (i.e., synthesis). Overall, Pettersson et al. [43] highlight a general trend towards mixed-method approaches combining different evaluation approaches and data types.

Existing methods and tools in UX evaluation

To collect both quantitative and qualitative data and to enable an easier implementation and communication of UX, numerous researchers have presented distinct methods and tools to enrich, amplify, and facilitate UX evaluation until now. Basic approaches focus on the pure evaluation of sensation based on Emocards or Emofaces [12] or facial recognition [11]. Additionally, more comprehensive methods consider the development of an experience over time, e.g., through a diary-style experience sampling (see [51]), self-reported experience sampling (see [30]), or a quantified visualization of UX over time (see [29] and [32]). Prominent and established approaches are standardized questionnaires that generally focus on different aspects of UX, e.g., AttrakDiff (focusing on hedonic and pragmatic product attributes through semantic differentials [25]), NASA TLX (particularly addressing workload [23]), SAM scale (an emotion focused pictogram-based questionnaire, see [26]) and many more.

Furthermore, with an increasing digitization of business, research, and development processes, designers can rely on diverse commercial software tools and services to support their daily work. These digital means range from simple survey tools (e.g., Google Forms³) over specialized design-oriented services, such as Persona (e.g., UserForge⁵) or Wireframing tools (e.g., Balsamiq⁶) to holistic qualitative and quantitative data analysis and visualization software (e.g., NVivo or Deoose). Recently, several researchers have also attempted to collect large quantities of contextual, qualitative UX data, e.g., through mobile diary studies [8, 9, 16, 50] or mobile ethnographies [42]. These tools mitigate a large part of the burden to plan and conduct qualitative diary studies or to visualize and analyze increasingly large data sets. However, the benefits of a digital image illustrating graphs, numbers and quotes that many digital software tools focus on come along with the risk to leave the individual user aside. Existing tools and methods of the same kind as the ones mentioned above primarily focus on the analysis of the actual feedback (e.g., graphs from survey results or digitized and automated content analysis of written feedback) than building a relationship with user. In contrast, Golsteijn & Wright [18] illustrate how narrative-centered approaches focusing on user relationships, for example, can strongly support researchers relationship with study participants and inform design ideas.

The role of the user in UX evaluation

Historically, products were generally evaluated based on task completion and performance, mainly driven by a usability mindset. At that time, the user was rather a mean to generate a quantified feedback to describe the quality of an interaction with a system as indicated in the previous paragraphs. The shift towards UX since the beginning of this century came along with a shift towards more qualitative tools and methods, putting the user more and more into the center of the design process [3]. As a consequence, according to Vermeeren et al. [54] one of the most common methods for UX evaluation is

³<https://www.google.com/forms/about/>

⁵<https://userforge.com>

⁶<https://balsamiq.com>

interviewing, whereas Alves et al. [1] demonstrate that observation and think-aloud are strongly endorsed in UX practice. Wright & McCarthy [56] took the role of the user a step further, as they point out how narratives and biography-based approaches, such as personas, scenarios, or probes can elicit rich experiences that help to derive insightful design artifacts and to truly understand the users' needs and objectives. In general, Batterbee & Koskinen [4] conceptualize three different approaches to UX based on the analysis how UX is interpreted and applied:

- *Measuring approaches* that focus on the measurable aspects of experience, e.g., through emotion detection or self-reporting.
- *Empathic approaches* that aim to foster a meaningful relationship between the user and the designer.
- *Pragmatist approaches* that base their analyzes on theoretical frameworks and models.

In our work, we take the perspective of an empathic approach, in which the role of the user is being part of a meaningful relationship with the designer.

Reflection on designer-user interaction in UX evaluation

To consider the diverse interests of different stakeholders in the design process, a useful UX evaluation tool that supports the creation of an empathic relationship with the user would ideally allow to gain qualitative insights and to quantify qualitative data if needed – depending on the individual needs of the different stakeholders. Consequently, UX evaluation tools should be designed in a way that they foster the relationship with the user and at the same time promote a shared mindset characterized by "instrumental judgement [...] fluid improvisation [...] rather than the mere performance of a codified method" as described by Gray [19]. In the words of Batterbee & Koskinen [4], we see great potential for digital software tools that balance a measuring approach based on 'numbers' and an empathic UX approach focusing on on users' emotions, particularly as researchers nowadays tend to stick with common evaluation approaches such as surveys, user logging, and interviewing [43].

A PARTICIPATORY DESIGN APPROACH

Our overall process and methodology focused on the creation of new design knowledge based on stakeholder interviews and a design artifact to address our RQ from different perspectives. According to the early work of Frayling [15] and Muller et al. [41], we understand our work as a process related to action research (i.e., a step-by-step documentation that communicates and contextualizes practical research results) and participatory design. Along our design process, we collaborated with a design agency and conducted semi-structured interviews, an ideation workshop, and finally developed a prototype that served as a design artifact/exemplar. Furthermore, we evaluated our prototype based on semi-structured interview, focus groups, and the AttrakDiff survey. The different steps of our approach are outlined in the following sections.

Interview Participant	Role(s) in the Design Process			
	User Researcher	Usability Engineer	Designer	Product Manager
I.1 (industry)	■	■		
I.2 (industry)	■	■		
I.3 (industry)	■		■	
I.4 (industry)	■		■	
I.5 (industry)	■	■		
I.6 (industry)	■			■
I.7 (industry)				■
I.8 (industry)	■	■		
I.9 (industry)	■			■

Table 1. Overview of the nine stakeholder interview participants (I.1 - I.9) according to their role and tasks in the design process.

INTERVIEW STUDY: DESIGNERS' NEEDS

To understand under which conditions empathic relationships between designers and participants flourish, we conducted semi-structured interviews with designers and further stakeholder in the design process.

Interview participants

We recruited nine participants for the interviews from our industry partner and additional snowball sampling (see Bieracki & Waldorf [6]). We mainly targeted people who are actively involved in user research activities. Table 1 shows an overview of all interview participants including their roles in the design process according to their own statements.

Interview procedure

In our semi-structured interviews that lasted between 45 and 60 minutes, we first asked about the participants' roles in the UX design process, what specific tasks they are responsible for, and what personal contact with (potential) users during the design process means for them in their work. We discussed which methods they used to involve users, if and how they maintain a relationship with users, which type of data they collect, and how they collect it. The interview guideline was driven by the motivation to understand current challenges in involving users throughout the different stages of the design process to finally develop a first prototype for a digital research tool that overcomes these hurdles. We also asked all participants how a digitally-mediated relationship between designers and users could be facilitated with the help of a digital research tool to derive first design ideas for the subsequent workshop. All interviews were transcribed, structured, and analyzed by one of the authors following a grounded theory approach (see [17] and [28]). We did not conduct an inter-rater reliability test at this stage as we did not aim to solely base our prototype design on the stakeholder interviews but rather use them as a starting point for the in-depth ideation workshop.

Interview results

Established methods: In general, our participants used diverse methods, ranging from interviews, focus groups, surveys (mainly UEQ and HUX) over card sorting and cultural probes to prototyping, co-creation, and lab studies (I.1 - I.9). Our

Workshop Participant	Role(s) in the Design Process			
	User Researcher	Usability Engineer	Designer	Product Manager
W.1 (academia)	■		■	
W.2 (industry)	■	■		
W.3 (industry)			■	■
W.4 (industry)		■		■
W.5 (industry)	■			
W.6 (industry)			■	

Table 2. Overview of the six ideation workshop participants (W.1 - W.6) according to their role and tasks in the design process.

interview participants used narratives, contextualized user stories, and personas to create empathy for users among the whole design team. They emphasized the need to read the story behind the feedback data (I.2, I.4, I.5) and that "users should be constantly involved at any phase in the design process" (I.2) and were keen to approach UX holistically by taking the user as well as the time and the context of use into account (I.2, I.6, I.7, I.8). However, a major hurdle to holistically evaluate UX is limited time and budget (I.1 - I.9).

Current challenges and opportunities to foster empathy with users: From the interview results, we were able to extract four main challenges or opportunities to nurture empathy between UX design teams and users in digitally-mediated user research, i.e., (1) maintaining empathic relationships with users, (2) driving the commitment of study participants to share their experience (i.e., give feedback), (3) creating a shared vision, i.e. understanding, of the relationship(s) with users in interdisciplinary design teams and (4) enhancing the versatility of a digital tool that leverages empathic relationships.

IDEATION WORKSHOP: DESIGN IDEAS

To collect ideas on how a digital communication platform can tackle these challenges, we organized a workshop in collaboration with our industry partner. The workshop helped us to involve all stakeholders and generate design ideas with practical value as suggested by Muller et al. [41]. For each question, the workshop participants brainstormed on how to address that challenge (see Figure 2).

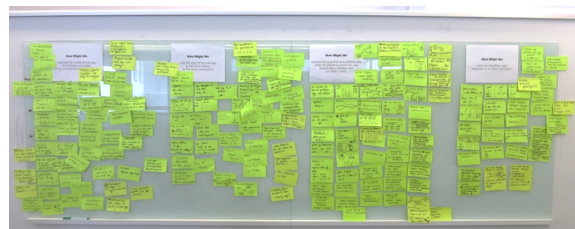


Figure 2. Whiteboard with all identified design ideas during our design workshop for each challenge.

Workshop participants

The workshop was conducted with six stakeholder (see Table 2) at our industry partner's office and was limited to two hours. Again, we recruited participants for the workshop from our industry partner based on their roles in the design process. Table 2 shows an overview of the workshop participants and their role(s) in the design process.

Workshop procedure

The starting point of the workshop was the challenge to create a digital tool that allows design teams to foster relationships with users. Many aspects about this tool were left deliberately open, for example, (1) how participants/users are recruited, (2) if users already used the product or a prototype that should be evaluated or not, (3) how messages would be exchanged between designers and users, or how designers would communicate with a whole user base of, e.g., registered test users. We started the workshop by presenting insights and challenges extracted from the interviews. For each of the four presented challenges, participants silently wrote down ideas for digital tools that tackle the presented challenge to foster empathy with users on post-it notes. Next, participants presented their ideas to the others while sticking their notes on the whiteboard. Finally, participants were asked to cluster ideas according to their similarity in an affinity diagram and to mark their favourite ideas with sticky dots (see Figure 2).

Workshop results

During the workshop, 183 partly redundant ideas were generated by the participants. After adding 32 ideas that we drew from the expert interviews, 215 ideas were subject to the dot-voting at the end of the workshop. Table 3 summarizes the highest-rated ideas with respect to each of the four challenges, i.e., (1) maintain empathy with users, (2) drive commitment of users to share their experiences, (3) create a shared vision in interdisciplinary design teams and (4) enhance the versatility of a digital tool that leverages empathic relationships.

PROTOTYPE STUDY: DIGITALLY-MEDIATED EMPATHY

To explore how digital communication platforms can help to create an empathic relationship between users and design teams we developed a first prototype for a digital UX tool based on our insights from our interview study and ideation workshop and evaluated the prototype in two studies (i.e., a pilot and a main study).

Objective and Feature Selection

With the help of the derived challenges from our interview study and the design ideas that we identified in our ideation workshop we created a prototype of a digital communication platform that focuses on creating and maintaining relationships with users. Guided by our RQ, our overall aim of the prototype was to develop a first proof-of-concept to explore the potential of digital tools to foster an empathic relationship between design teams and users. We did not intend to develop a market-ready research tool but to implement selected design ideas from our workshop to address our RQ. Consequently, together with our industry partner we defined several criteria to specify which features should be implemented in the prototype with the given scope and limited time of the project.

First, we determined to focus on the designer/researcher view of such a digital tool as we imagined that the user view will build on existing solutions and concepts to collect feedback with, e.g., a mobile application (see, e.g., [50]). Second, we decided to implement features that directly support interdisciplinary design teams' "orientation[s] to the other person [i.e., user]" [56]. Third, we agreed that a digital communication platform should not only represent an add-on in the method toolkit of UX teams but provide the established and valued features of existing solutions to analyze and communicate UX feedback. Based on these criteria, we derived four main features for our prototype:

- Persona and user journey view that allow the design team to truly understand the users' goals [10].
- Communication features - as empathy situates in the a context of communication [55].
- Flexible filters to additionally support UX teams in established tasks and processes for qualitative data analysis.
- Feedback visualizations as an effective form to communicate UX feedback within interdisciplinary teams.

In Table 3 we highlight the ideas related to these main features that we prioritized for the development of our prototype.

Prototype

As described above, the concept of the UX platform we envisioned to foster empathy in design teams built on existing means to collect feedback from users with a mobile application as presented in [50]. We therefore assume that an app installed on participants phones allows them to capture image and audio material and enter text about their experience with a product, service, or prototype. Moreover the app functions as a portal for the communication with the design team. The prototype described in the following paragraphs focuses on the application used by the design team. Its main goal is to visualize content generated by participants and fostering empathy between users and designers. Below, we describe the visual design, technical implementation as well as the four mentioned main features/elements of the first proof-of-concept.

Persona and user journey view: The profile section consolidates a profile overview for every study participant (see Figure 3 middle). Each profile can be expanded to display a timeline as well as a grid view of the participant's feedbacks. This profile section was conceptualized and implemented as the main discussion in the workshop centered around the creation of adequate user profiles and personas. According to Cooper and colleagues [10], such personas or profiles based on real data and real-world observations are "easier to relate to than feature lists and flowcharts".

Communication features: On the feedback page, the data is displayed in a tabular form with columns referring to the feedback (ID, rating, category, title, description, date, photo, mail and tags) as well as columns referring to the user (name, job, age, gender). By selecting the envelope icon in the mail column, every stakeholder of the design team can start a conversation with the user (in our prototype, the user's mail program is

(1) Maintain Empathy with Users	(2) Drive Commitment of Users	(3) Create a Shared Vision	(4) Enhance Versatility of a Dig. Tool
<ul style="list-style-type: none"> * interacting via a digital user research tool allow users to choose a persona that they feel represents them well (or to create a new persona) encourage users to use diverse input modalities especially voice messages, photos and videos, drawings or shaking of the mobile device show users a representation of designers in the design team and their specific interests, the history of their conversations with the design team and vice versa allow users to annotate and prioritize their feedback/message * designers need to be able to contact participants to inquire background information and follow up on feedback that turn out to be ambiguous user should be able to contact designers via chat or call and vice versa include (a wide range of) emoticons as an input modality 	<ul style="list-style-type: none"> pay incentives for high engagement in different stages of the project before the study, the users should be explicitly informed about the purpose of the study * giving feedback and sending messages must be very quick and easy (what information can be viewed by whom?) multiple people in the design team could engage with a participant to drive commitment, but their identity should be revealed for the participant mix open and closed questions to keep barriers to engage low appreciate any input, feedback, or information provided by users (e.g., by thanking them, compensating them financially) * show designers if engagement of participants declines, so that they can be 'ping' them ask 'why' questions and show that someone is truly interested create a sense of belonging (e.g., participants can view each others messages and can react or comment) and include reminders (possibly automatic or context-based) 	<ul style="list-style-type: none"> * use persona icons next to feedback/messages from a user leverage meta data (e.g., input duration) to understand the context in which the feedback was created * user profile visualization derived from all insights from a user should be linked to an individual profile (can be compared to users) * emotional user journey visualization including feedback history, rating and mood over the course of the study (also consider visualizing day time or day of the week) * persona visualization with a mix of information collected from users * adopt visualization of feedback to content and its modality (e.g., long text is presented as a blog, photos as gallery) and make filter methods transparent * highlight outliers: extreme users and use situations are often perceived as very insightful * link any aggregated visualization with single feedback sources include highlighting/messaging features for collaboration within the design team 	<ul style="list-style-type: none"> create a tool for both designers and users and foster a community * feedback should focus on the context of use and feedback over time a tool for setting up the study and for processing and presenting the feedback data * guiding the stakeholder through modular settings of a study (e.g., current phase in the design process, focus of the study, and selecting the participants) support the creation of a reusable study design * modalities should be adaptable to use cases (e.g., choose from text, video, sound, screenshots, location, pace, heart rate, temperature) * balance customizable visualizations: quantitative (e.g., bar charts) including filters/meta-data and qualitative (e.g., word cloud)

Table 3. Overview of the design ideas that were rated highest in our workshop and interview study for all identified hurdles to foster empathy with the help of a digital tool (ideas that have been taken into consideration for the development of the main features of our prototype are marked with *).

opened along with a new mail addressing the respective participant and citing his or her corresponding feedback). The communication features represent the core function of the tool to foster empathy with users and to drive commitment as expressed in our interviews and workshop. We see this as a vital part as professional empathy training, in general, is substantially based on communication training [56].

Flexible filters: The main page of the tool is the dashboard that allows the design team to access the three main pages: overview of all feedback data, profiles, and visualizations. A global filter enables the user to screen the feedback data across all pages of the application and a tag cloud depicts the 20 most frequent terms in the feedback. By selecting data points in the grid or in visualizations a preview of the feedback is displayed as an overlay (Figure 3 right). These filter options represent the basic analysis option of the tool as our industry partner emphasized that different stakeholders with different interests will and should use the tool.

Feedback visualizations: The visualizations page displays different types of visualization of the feedback data. A bar chart and a pie chart depict the overall frequency distribution of the feedback rating, a stacked bar chart illustrates the rating frequency distribution per feedback category, a timeline displays the feedbacks over a period of time, and a tag cloud,

as mentioned, the most frequently used words in the users' feedbacks. These visualizations were derived from the wish of our workshop participants to quickly gain an overview about the user feedback while at the same time highlight if/when engagement of participants declines over time.

Visual design: All features were first designed as a visual mockup to run a final iteration with our industry partner (see Figure 1). Most prominently, our industry partner highlighted the suitability of the user profiles and user journeys in the mockups, which, as a consequence, we prioritized in the final prototype. The proof of concept was implemented as a high fidelity/low resolution prototype (see Figure 3).

Technical implementation: The prototype was developed as a web-based application in Angular (version 4.0.1⁷) and deploys the D3 visualization library⁸. The grid view of the feedback data is built upon the JavaScript-based data grid "ag-grid"⁹. The tag cloud is based on an Angular module by Dario Costa¹⁰. Unlike design practice, we did not aim to develop a marketable

⁷<https://angular.io>, retrieved July 2017

⁸Michael Bostock, Vadim Ogievetsky, Jeffrey Heer. D3: Data-Driven Documents. In: IEEE Transactions on Visualization and Computer Graphics, 17:12, 2011, 2301-2309

⁹<https://ag-grid.com>

¹⁰<https://github.com/DarioCosta/tag-cloud>

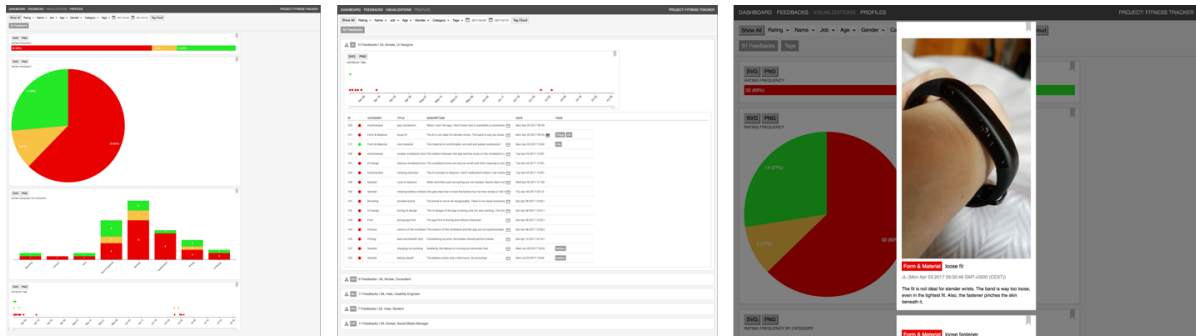


Figure 3. Screenshots of our UX evaluation prototype that visualizes qualitative data in aggregated (left) and detailed (right) views as well as by individual participants (middle).

product but a prototype that helps to stimulate a discourse around the research questions according to Zimmermann and colleagues [58].

Study Setting and Participants

Pilot Study: We conducted a pilot study with five employees of our industry partner to get first insights about a suitable implementation of our prototype and to identify target questions for our main study. Each participant had two tasks: First to analyze a given set of feedback data using a tool of their own choice. Second, the participants analyzed the same data set with our prototype (we changed the order of tasks for three participants to avoid order bias). The data set consisted of written feedback (distributed as a CSV file) and photos (distributed in a zip folder) evaluating the office environment of our industry partner. We conducted individual semi-structured interviews with all participants (taking about one hour each) and one group interview (taking about 15 minutes). Furthermore, all participants filled out the AttrakDiff questionnaire. The participants in this pilot study did not receive a compensation.

Main Study: To evaluate our prototype more holistically, we conducted another evaluation in form of a case study. We recruited seven focus groups with a total of 26 participants from our educational institution. All groups had just finished a product development course and hence were familiar with product development tools and processes. Furthermore, each group consisted of students with different study backgrounds, mainly from business administration, electrical engineering, physics, and computer science (or similar, see Table 4). Thus, we ensured an evaluation setting with a practical orientation in terms of interdisciplinary backgrounds and roles. For this study, we first asked the participants about their experience in the product development course as a warm-up. Then, we gave each group the same test feedback data set about a fitness tracker. The groups had the task to analyze the data and jointly generate key insights. Each focus group took about 45 minutes and was conducted as a combination of think-aloud and semi-structured group interviews. Additionally, all participants filled out the AttrakDiff questionnaire. The participants in this study were compensated with 10€, or roughly US\$12.

Inter-rater reliability: The interviews of the main study were recorded, transcribed, and analyzed through grounded theory.

The two authors who conducted the focus groups first discussed initial insights and jointly derived six coding categories according to our RQ and the results of our pilot study: "observing data", "analyzing data", "implementing the tool" (i.e., how to use it), "improvements for tool", "ideal use case" (i.e., when to use it), and "task comment / not meaningful" (i.e., comments related to the case study task, not a feedback to our tool). The categories were chosen to better understand how a digital tool that aims to nurture empathy with users could be used in different design phases and how it could be implemented in the design process. The categories helped us to cluster common feedback through inductive category formation according to Mayring & Fenzl [39] to finally derive overall insights and guidelines for digital research tools with the objective to leverage relationships with users. Two authors coded all focus groups to ensure inter-rater reliability (IRR). IRR was found to be $\alpha = 0.9009$, 95% in a CI of (0.8431, 0.9504). According to Krippendorff [31], values for α higher than .8 can be seen as satisfactory.

Results

The evaluation of our prototype was based on the pilot study and the main study described above. The overall goal was to receive feedback on our prototype and at the same time to further understand needs and requirements for design teams related to the empathic connections between designers and users. In the pilot study with our industry partner we focused on the identification of key topics and questions that we can then intensively discuss in our main study. The feedback in the smaller pilot study with five participants (numbered P.A, P.B, P.C, P.D, and P.E) mainly focused on how people use the prototype compared to established approaches. In our case, we primarily compared the use of our prototype to the use of the spreadsheet software Excel due to the established design processes of our industry partner. Thus, we were able to identify a first reasoning for using a digital communication platform that combines traditional feedback visualizations with distinct user profiles and user feedback journeys as "it helps that everything is connected" (P.B) and even allows researchers and designers to present and discuss unknown feedback data directly with the management board or customers (P.A). Figure 4 highlights the evaluation (i.e., consistently higher pragmatic quality) of

Focus Group	Gender and Age [years]	Background		
		Business Administration	Eng. and Nat. Sciences	Social Sciences
1	male / 27		■	
	male / 25		■	
	female / 19	■		
	male / 22		■	
2	female / 26		■	
	male / 22		■	
	female / 23	■		
	female / 22	■		
3	male / 26	■		
	male / 25		■	
	female / 26		■	
4	female / 26		■	
	male / 30		■	
	male / 24		■	
	female / 21	■		
5	male / 25	■		
	male / 21		■	
	female / 23		■	
	male / 22	■		
6	male / 23	■		
	male / 26		■	
	male / 26	■		
	male / 26			■
7	male / 22	■		
	male / 26		■	
	male / 19		■	

Table 4. Overview of the focus group participants according to their group association and educational background.

our prototype using the AttrakDiff questionnaire compared to traditional spreadsheet software.

Next, in the main study, we asked seven focus groups about the usefulness of specific features, hence were able to better understand the derived benefits related to digitally mediated connections with users. The main study revealed that the current version of our prototype was still considered neutral yet still with a higher pragmatic and hedonic value than the spreadsheet software (see Figure 5). We also found out that such a digital communication platform can help to analyze qualitative data on a large scale, as the visualizations were seen as "more useful when you have a lot of data" (Group 7), and even highlight unexpected insights, as "the tag cloud [...] is a great way to explore the data in a way that is not common which might lead to uncommon insights". (Group 1). With a closer look at the connection and relationship with users, our in-depth qualitative analysis of the focus groups pointed out three main benefits of our digital UX tool :

- Establish a relationship with users
- Supporting long-term studies at later stages
- Support collaboration to foster user interaction

Establish a relationship with users: The individual user profiles in our prototype helped to efficiently derive in-depth insights from the data set and relate to individual user feedback. A participant in Group 2 mentioned that he "used the link to the profiles when [he] saw that one person had a lot of feedback so [he] wanted to check what else he's writing". The profiles were also helpful, according to Group 7, "when you were looking for like conflicting feedback [...] so you could actually see how the opinion of one person changed". In general, the visualizations, tag cloud, and user profiles allowed our study participants to quickly get familiar with the data set. In particular, however, the profile section helped the teams to truly understand different needs of different users.

If you want to check the feedback for customer groups or if you want to check whether it is more intuitive for some people than for others, then I think that the profiles section is very useful. (Group 2)

Moreover, we were able to verify throughout the course of our project that the increasing multidisciplinary of UX processes is perceived as a challenge. Different users of UX tools have different goals and objectives and hence are interested in different feedbacks. The profile and notification feature in our prototype was built to consider diverse interests. Our study showed that "profiles are actually super important" (Group 1) and that the user journey view in the profile section was considered as very valuable, because "when you look back at a project, it's useful to have all the overviews and be able to track it over time" (Group 7). Furthermore, Group 2 exemplarily highlights another benefit of the notification feature strongly related to the need for tools that foster empathic connections with users.

If something was unclear and if it was a topic that a lot of other people talk about, then I would ask this person to tell me more [...] this [notification] feature is cool. (Group 2)

Supporting long-term studies at later stages: In our pilot study, the interview participants' opinions about when a UX tool that aims to foster an empathic connection with users is most useful varied from at the beginning to at the end of a development process. Several groups in our main study, however, mentioned that "in an early stage where you have paper prototypes [that are not close to the design of the] final product, [...] observing is more powerful" (Group 5). In Group 7, a participant added that such a UX tool is useful at later stages because then people have used the product for a longer time and one can get more insights. Overall, whereas traditional data gathering methods, such as interviewing, were seen as helpful to interact with users at early stages our prototype was considered as beneficial for later stages to enable the maintenance of empathic relationships with several users over a longer time.

I think [it would be useful] more for large-scale studies and not directly at the beginning, because there it is really important that you actually talk to the person because [...] you can directly ask a follow-up question [...]. If you talk to users it is always better than if you do it over a

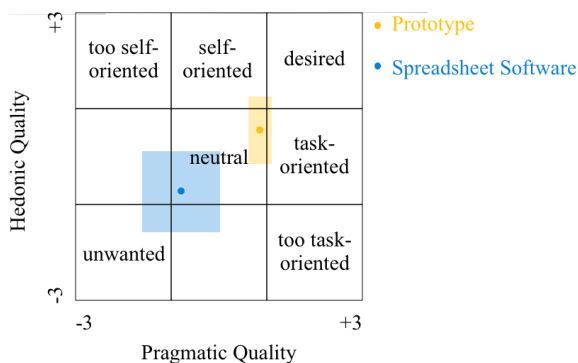


Figure 4. Results of the AttrakDiff questionnaire comparing the use of a spreadsheet software to the use of our UX evaluation prototype in our pilot study.

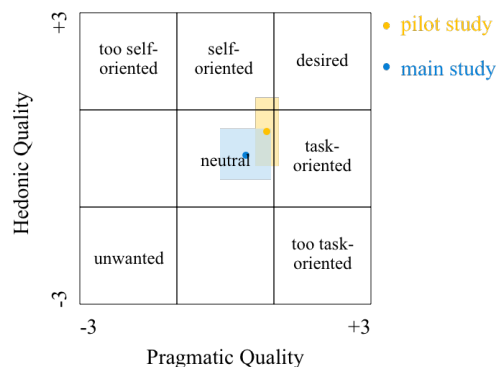


Figure 5. Results of the AttrakDiff questionnaire comparing our pilot study and main study.

screen, so if it is still not a big sample of people I would still do it face to face but if it's actually a big data [set] with many people I would definitely go for this [tool]. (Group 5)

Support collaboration to foster user interaction: The participants in our focus group were asked to jointly analyze the presented case study. As a consequence, communication and collaboration between the focus group members naturally evolved over the course of the study as soon as the participants got familiar with the feedback data. This revealed further insights for collaboration in the UX process, mainly formulated as suggested improvements for our prototype. Group 5 mentioned that it "would be nice to make the [tool] more social, [...] you might have pictures of people who have already looked at [certain issues] so you know "ah, he has seen this" [...] that would be nice". Putnam and Kolko [45] report similar insights and highlight the challenge that, in general, not every stakeholder in the design process usually interacts with users hence it is difficult to pass on empathic connections. Other groups in our study highlight the option to increase the involvement of users similar to the ideas from our ideation workshop, e.g., "in the dashboard view or executive summary [you might] get a warning that a test user is not giving feedback anymore" (Group 6). In addition, our focus group study revealed that such a tool can help to even track company wide UX goals, as emphasized by Roto et al. [47]:

Now imagine I am Apple and I have phones, fitness trackers like the Apple Watch and I have Macbooks; would there be an option to see general problems [...] not on a product but more on a company level? (Group 6)

REFLECTION ON DIGITALLY-MEDIATED EMPATHY

For our research questions, the guiding principles of participatory design proved to be valuable as opinions from and needs for different stakeholders had to be considered. We pursued a participatory design approach involving relevant perspectives in our interview study and to develop a design artifact to generate design knowledge for digitally-mediated empathy in UX design processes.

Our interview study and prototype study allowed us to endorse anticipated needs and benefits of a UX tool to foster empathic connections with users. We saw that the topic of maintaining empathy with users is very important for design teams yet at the same time challenging to address with digital software from the UX method toolkit. Existing UX tools and evaluation methods, e.g., Quantified UX (see [33]) or AttrakDiff (see [25]) demonstrate several benefits of digital UX tools, such as the support to prioritize UX initiatives and the comparison of different products. However, these UX evaluation approaches mainly focus on quantitative data and lack insights into rich qualitative feedback that can emerge from empathic relationships. Schneider et al. [50] emphasize that digital tools can further support the consolidation of context-rich user stories through the collection of context-sensitive qualitative data. In contrast, further tools that support the analysis and visualization of both quantitative and qualitative UX, such as NVivo¹ or Dedoose², do not yet provide features to develop and maintain a relationship with users.

Our prototype study outlined that distinct tool features, particularly the persona and user journey view as well as the communication feature, can be beneficial complements to traditional visualizations and analysis approaches. Although personas and user journeys are traditionally used in early design stages and challenging to apply in the design process (see Matthews et al. [38]) they were considered as helpful for long term studies and for later stages. At the same time, Salminen et al. [49] already showcased that further research is needed to understand how to best apply Persona profiles in general. In our work, we felt that in combination with the communication feature, the persona and user journey view can help to maintain a better overview of individual user feedback and maintain empathy beyond an initial requirement definition in early design stages. Interestingly, our study participants also suggested several additional add-on features and further use cases for such a tool that have not been implemented in the prototype due to time and feasibility constraints but have

¹<https://qsrinternational.com>

²<http://dedoose.com>

already been suggested in our ideation workshop. As a consequence, we feel strengthened in our effort to create a UX tool that supports digitally-mediated connections with users.

However, we realize that our work has some inherent limitations. The design of our prototype was strongly influenced by the opinions and voting of our interview and workshop participants. Although we invited participants with different backgrounds, they were mainly recruited through our industry partner. Consequently, the development of our tool was guided by the mindset, previous experiences, and established digital tools that were frequently used by our interview and workshop participants, particularly the spreadsheet software MS Excel. Furthermore, to ensure a practical tool and minimize complexity in our initial prototype we had to limit the number of implemented design ideas. We acknowledge that different interview and study participants, a different industry partner, or different selection criteria to derive the prototype from all design ideas might have led to a different design of the prototype. A different design of the prototype, in turn, might have revealed different insights and suggestions. At the same time, we are thankful for the close collaboration with the design agency as it allowed us to apply various analysis methods (i.e., stakeholder interviews, workshop, pilot study) that led to a holistic understanding of our research question. In addition, the collaboration with an industry partner ensured that we base the design of our prototype on real design and research experiences from different stakeholders rather than merely on theoretical hypotheses. Nevertheless, these experiences can be unique to our industry partner hence further analyses and different partners need to be considered. Ultimately, we want to motivate further research and the implementation as well as evaluation of further features in digital design tools as exemplarily summarized in Table 3.

We have chosen to evaluate our prototype with interdisciplinary student groups. Senior UX researchers or designers with several years of work experience, in comparison, might have different needs and wishes related to digital UX evaluation tools. Nevertheless, no focus group in our main study questioned the functionality of our prototype but rather suggested improvements and additional features. Consequently, we feel reassured in our research effort to foster a research and design culture that focuses on empathic connections with users. Furthermore, we would like to encourage further research in that field to investigate additional case studies, probes, reflections, or design rationale - according to Greenburg & Buxton [20] more appropriate in certain situations.

CONCLUSION AND FUTURE WORK

In this paper, we have presented a participatory design approach including an interview study and a prototype study for the development of a UX tool to foster empathic connections between users and design teams. In collaboration with a design agency, we conducted stakeholder interviews, hosted an ideation workshop, and developed a prototype that supports empathic interactions with users. Finally, we evaluated our tool in a pilot study with employees from our industry partner and a main study with interdisciplinary student teams from our institution in seven focus groups. Our goal was to alleviate

the task of establishing and maintaining empathic connections with users in interdisciplinary teams during a UX project.

In our stakeholder interview, we saw once more that there is a need for design knowledge and resulting tools in the field of empathic connections with users. At the same time, our evaluation demonstrated that UX software tools can go further than traditional spreadsheet software and support the analysis of qualitative data on a large scale. Interestingly, our study revealed that a UX evaluation tool based on visualizations and user profiles can highlight unexpected insights - an area that we consider very promising for future work as it can spark creativity and innovation. Furthermore, it was encouraging to hear that based on our research questions our study participants did not only positively evaluate the user profiles that allowed the development of closer user connections but already suggested additional features to support team collaboration and the communication with users.

In the future, we see a variety of opportunities to bring the empathic connections with users to the next level. Research questions that have not yet been addressed in this work are the implementation of natural language processing, machine learning, chatbots and further technical developments that designers start to integrate in their product or services recently (see, e.g., Dove et al. [13] and Yang et al. [57]). Such approaches can strongly assist the observation, analysis, and communication of UX insights but need to balance the danger of an increasing distance between designers and users and consequently losing empathy during the design process. Further challenges for empathic UX software tools that we did not focus on but need to be addressed in future studies are how the right balance of powers between designers as experts and users as embraced study participants look like, how test users would like to be reimbursed/motivated, the level of transparency that users demand, and how a trusted and rewarding relationship without biasing results can be maintained in the long run.

ACKNOWLEDGMENTS

We thank our partner agency designaffairs GmbH for the collaboration and their contribution to our research project with insightful stories and experiences related to design processes. Furthermore, we thank all study participants for their time, insights, and engagement throughout the study.

REFERENCES

1. Rui Alves, Pedro Valente, and Nuno Jardim Nunes. 2014. The State of User Experience Evaluation Practice. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14)*. ACM, New York, NY, USA, 93–102. DOI: <http://dx.doi.org/10.1145/2639189.2641208>
2. Shaowen Bardzell. 2010. Feminist HCI: Taking Stock and Outlining an Agenda for Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 1301–1310. DOI: <http://dx.doi.org/10.1145/1753326.1753521>
3. Javier A Bargas-Avila and Kasper Hornbæk. 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*. ACM, 2689–2698.
4. Katja Battarbee and Ilpo Koskinen. 2005. Co-experience: user experience as interaction. *CoDesign* 1, 1 (2005), 5–18.
 5. Nigel Bevan. 2008. Classifying and selecting UX and usability measures. In *International Workshop on Meaningful Measures: Valid Useful User Experience Measurement*, Vol. 11. 13–18.
 6. Patrick Biernacki and Dan Waldorf. 1981. Snowball sampling: Problems and techniques of chain referral sampling. *Sociological methods & research* 10, 2 (1981), 141–163.
 7. Michael Burmester, Kilian Jäger, Laura Festl, and Marcus Mast. 2011. Studien zur formativen Evaluation der User Experience mit der Valenzmethode. *Reflexionen und Visionen der Mensch-Maschine-Interaktion-Aus der Vergangenheit lernen, Zukunft gestalten* 9 (2011), 567–572.
 8. Scott Carter and Jennifer Mankoff. 2005. When participants do the capturing: the role of media in diary studies. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 899–908.
 9. Sunny Consolvo and Miriam Walker. 2003. Using the experience sampling method to evaluate ubicomp applications. *IEEE Pervasive Computing* 2 (2003), 24–31.
 10. Alan Cooper, Robert Reimann, David Cronin, and Christopher Noessel. 2014. *About Face. The Essentials of Interaction Design* (fourth ed ed.). John Wiley & Sons, Inc., Indianapolis, Indiana.
 11. Pieter Desmet. 2003. Measuring emotion: Development and application of an instrument to measure emotional responses to products. In *Funology*. Springer, 111–123.
 12. Pieter Desmet, Kees Overbeeke, and Stefan Tax. 2001. Designing products with added emotional value: Development and application of an approach for research through design. *The design journal* 4, 1 (2001), 32–47.
 13. Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX Design Innovation: Challenges for Working with Machine Learning As a Design Material. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 278–288. DOI: <http://dx.doi.org/10.1145/3025453.3025739>
 14. Hugh Dubberly and Shelley Evenson. 2008. On modeling The analysis-synthesis bridge model. *interactions* 15, 2 (2008), 57–61.
 15. Christopher Frayling. 1993. Research in art and design. (1993).
 16. Jon Froehlich, Mike Y. Chen, Sunny Consolvo, Beverly Harrison, and James A. Landay. 2007. MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones. In *Proceedings of the 5th international conference on Mobile systems, applications and services*. ACM, 57–70.
 17. Barney G Glaser, Anselm L Strauss, and Elizabeth Strutzel. 1968. The discovery of grounded theory; strategies for qualitative research. *Nursing research* 17, 4 (1968), 364.
 18. Connie Golsteijn and Serena Wright. 2013. Using narrative research and portraiture to inform design research. In *IFIP Conference on Human-Computer Interaction*. Springer, 298–315.
 19. Colin M Gray. 2016. "It's More of a Mindset Than a Method": UX Practitioners' Conception of Design Methods. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, New York, New York, USA, 4044–4055. DOI: <http://dx.doi.org/10.1145/2858036.2858410>
 20. Saul Greenberg and Bill Buxton. 2008. Usability Evaluation Considered Harmful (Some of the Time). In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. ACM Press, New York, New York, USA, 111. DOI: <http://dx.doi.org/10.1145/1357054.1357074>
 21. Steve Harrison, Phoebe Sengers, and Deborah Tatar. 2011. Making epistemological trouble: Third-paradigm HCI as successor science. *Interacting with Computers* 23, 5 (2011), 385–392.
 22. Steve Harrison, Deborah Tatar, and Phoebe Sengers. 2007. The three paradigms of HCI. In *Alt. Chi. Session at the SIGCHI Conference on Human Factors in Computing Systems San Jose, California, USA*. 1–18.
 23. Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 50. DOI: <http://dx.doi.org/10.1037/e577632012-009>
 24. Marc Hassenzahl. 2008. User experience (UX): towards an experiential perspective on product quality. In *Proceedings of the 20th Conference on l'Interaction Homme-Machine*. ACM, 11–15.
 25. Marc Hassenzahl, Michael Burmester, and Franz Koller. 2003. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Mensch & Computer 2003*. Springer, 187–196.
 26. Marc Hassenzahl and Daniel Ullrich. 2007. To do or not to do: Differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with computers* 19, 4 (2007), 429–437.
 27. Mats Hellman and Kari Rönkkö. 2008. Is User Experience supported effectively in existing software development processes?. In *Proc. of COST294-MAUSE Workshop on Valid Useful User Experience Measurement (VUUM)*. Reykjavik, Island. 32–37.

28. Lynn A. Isabella. 1990. Evolving Interpretations as a Change Unfolds: How Managers Contrive Key Organizational Events. *Academy of Management Journal* 33, 1 (mar 1990), 7–41. DOI: <http://dx.doi.org/10.2307/256350>
29. Evangelos Karapanos, Jean-Bernard Martens, and Marc Hassenzahl. 2010. On the retrospective assessment of users' experiences over time: memory or actuality?. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*. ACM, 4075–4080.
30. Evangelos Karapanos, John Zimmerman, Jodi Forlizzi, and Jean-Bernard Martens. 2009. User experience over time: an initial framework. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 729–738.
31. Klaus Krippendorff. 2004. *Content Analysis: An Introduction to Its Methodology* (second ed.). Sage.
32. Sari Kujala, Virpi Roto, Kaisa Väänänen-Vainio-Mattila, Evangelos Karapanos, and Arto Sinnelä. 2011. UX Curve: A method for evaluating long-term user experience. *Interacting with Computers* 23, 5 (2011), 473–483.
33. Florian Lachner, Philipp Naegelein, Robert Kowalski, Martin Spann, and Andreas Butz. 2016. Quantified UX: Towards a Common Organizational Understanding of User Experience. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI '16)*. ACM, New York, NY, USA, Article 56, 10 pages. DOI: <http://dx.doi.org/10.1145/2971485.2971501>
34. Effie Lai-Chong Law. 2011. The measurability and predictability of user experience. In *Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems*. ACM, 1–10.
35. Effie L-C. Law, Virpi Roto, Marc Hassenzahl, Arnold P.O.S. Vermeeren, and Joke Kort. 2009. Understanding, scoping and defining user experience: a survey approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 719–728.
36. Effie Lai Chong Law, Paul Van Schaik, and Virpi Roto. 2014. Attitudes towards user experience (UX) measurement. *International Journal of Human Computer Studies* 72, 6 (2014), 526–541. DOI: <http://dx.doi.org/10.1016/j.ijhcs.2013.09.006>
37. Tuuli Mattelmäki and Katja Battarbee. 2002. Empathy probes. In *PDC*. 266–271.
38. Tara Matthews, Tejinder Judge, and Steve Whittaker. 2012. How do designers and user experience professionals actually perceive and use personas?. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1219–1228.
39. Philipp Mayring and Thomas Fenzl. 2014. *Qualitative Content Analysis*. Springer.
40. Deana McDonagh. 2004. Empathic Design: User Experience in Product Design: I. Koskinen, K. Battarbee And T. Mattelmäki (Eds). *The Design Journal* 7, 3 (2004), 53–54.
41. Michael J Muller and Sarah Kuhn. 1993. Participatory design. *Commun. ACM* 36, 6 (1993), 24–28.
42. Matthias Muskat, Birgit Muskat, Anita Zehrer, and Raechel Johns. 2013. Generation Y: evaluating services experiences through mobile ethnography. *Tourism Review* 68, 3 (2013), 55–71.
43. Ingrid Pettersson, Florian Lachner, Anna-Katharina Frison, Andreas Riener, and Andreas Butz. 2018. A Bermuda Triangle? A Review of Method Application and Triangulation in User Experience Evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*.
44. B Joseph Pine, James H Gilmore, and others. 1998. Welcome to the experience economy. *Harvard business review* 76 (1998), 97–105.
45. Cynthia Putnam and Beth Kolko. 2012. HCI professions: differences & definitions. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2021–2026.
46. Christian Rohrer, Irene Au, Elissa Darnell, Nancy Dickenson, Shelley Evenson, and Klaus Kaasgaard. 2008. Design, marketing, strategy: where does user research belong?. In *CHI'08 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2241–2244.
47. Virpi Roto, Yichen Lu, Harri Nieminen, and Erdem Tural. 2015. Designing for user and brand experience via company-wide experience goals. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2277–2282.
48. Virpi Roto, Heli Väättäjä, Satu Jumisko-Pyykkö, and Kaisa Väänänen-Vainio-Mattila. 2011. Best practices for capturing context in user experience studies in the wild. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. ACM, 91–98.
49. Joni Salminen, Lene Nielsen, Soon-gyo Jung, Jisun An, Haewoon Kwak, and Bernard J Jansen. 2018. “Is More Better?”: Impact of Multiple Photos on Perception of Persona Profiles. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 317.
50. Hanna Schneider, Katharina Frison, Julie Wagner, and Andras Butz. 2016. CrowdUX: A Case for Using Widespread and Lightweight Tools in the Quest for UX. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, New York, NY, USA, 415–426. DOI: <http://dx.doi.org/10.1145/2901790.2901814>
51. Christie Napa Scollon, Chu-Kim Prieto, and Ed Diener. 2009. Experience sampling: promises and pitfalls, strength and weaknesses. In *Assessing well-being*. Springer, 157–180.

52. Leon D Segal and Jane Fulton Suri. 1997. The empathic practitioner: Measurement and interpretation of user experience. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 41. SAGE Publications Sage CA: Los Angeles, CA, 451–454.
53. Kaisa Väänänen-vainio mattila, Virpi Roto, and Marc Hassenzahl. 2008. Towards Practical User Experience Evaluation Methods. *5th COST294-MAUSE Open Workshop on Valid Useful User Experience Measurement* (2008), 1–4.
54. Arnold P.O.S. Vermeeren, Effie L-C. Law, Virpi Roto, Marianna Obrist, Jettie Hoonhout, and Kaisa Väänänen-Vainio-Mattila. 2010. User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*. ACM, 521–530.
55. Gert-Jan Vreeke and Ingrid L Van der Mark. 2003. Empathy, an integrative model. *New Ideas in Psychology* 21, 3 (2003), 177–207.
56. Peter Wright and John McCarthy. 2008. Empathy and Experience in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 637–646. DOI: <http://dx.doi.org/10.1145/1357054.1357156>
57. Qian Yang, Alex Scuito, John Zimmerman, Jodi Forlizzi, and Aaron Steinfeld. 2018. Investigating How Experienced UX Designers Effectively Work with Machine Learning. In *Proceedings of the 2018 Designing Interactive Systems Conference*. ACM, 585–596.
58. John Zimmerman, Erik Stolterman, and Jodi Forlizzi. 2010. An Analysis and Critique of Research through Design: towards a formalization of a research approach. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems - DIS '10*. ACM Press, New York, New York, USA, 310. DOI: <http://dx.doi.org/10.1145/1858171.1858228>

Culturally Sensitive User Interface Design: A Case Study with German and Vietnamese Users

Florian Lachner
LMU Munich
Munich, Germany
florian.lachner@ifi.lmu.de

Mai-Anh Nguyen
LMU Munich
Munich, Germany
mai.anh.nguyen@campus.lmu.de

Andreas Butz
LMU Munich
Munich, Germany
butz@ifi.lmu.de

ABSTRACT

Considering country-specific preferences in user interface (UI) design is a time-consuming task. We present a case study with German and Vietnamese users to explore how cultural theory can be applied in early design phases to support culturally sensitive design. We present an analysis of cultural dimensions and a comparison of German and Vietnamese question-and-answer (Q&A) websites. Based on the derived insights, we developed two UI concepts of a Q&A-website that differ in information architecture, navigation structure, and visual presentation. The prototypes were assessed with 14 German and 14 Vietnamese users in a think aloud setting. We were able to draw a conclusion about our initial analysis and the differing evaluation of the participants from the two countries due to their preferences regarding information retrieval, trust, and error handling. Our analysis provides first insights into the applicability of cultural theory in UI design but also opens up questions for further research.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; • **HCI design and evaluation methods** → *User studies*;

KEYWORDS

User Interface Design, Cultural Dimensions, Culturally Sensitive Design, Hofstede, Cross-Cultural Design, User Experience

ACM Reference Format:

Florian Lachner, Mai-Anh Nguyen, and Andreas Butz. 2018. Culturally Sensitive User Interface Design: A Case Study with German and Vietnamese Users. In *2nd African Conference for Human Computer Interaction (AfriCHI '18)*, December 3–7, 2018, Windhoek, Namibia. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3283458.3283459>

1 INTRODUCTION

Placing users' needs at the center of the design process has been accepted as crucial to ensure commercial success and customer loyalty for a long time [1, 64]. In addition, since the turn of the century, design teams and researchers are increasingly aware that users' needs

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

AfriCHI '18, December 3–7, 2018, Windhoek, Namibia

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6558-1/18/12...\$15.00

<https://doi.org/10.1145/3283458.3283459>

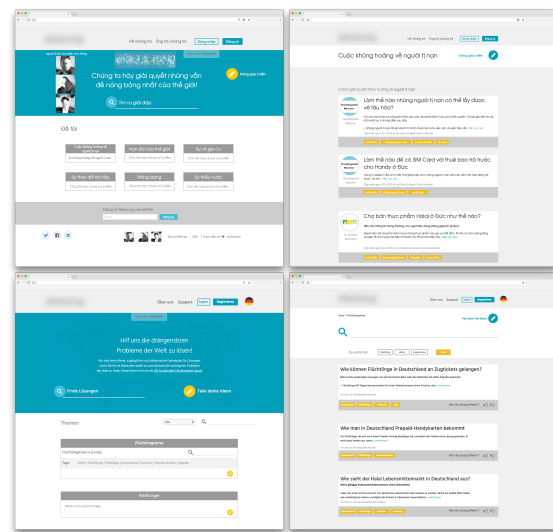


Figure 1: Screenshots of the main page (left) and an exemplary sub-page (right) of our two prototype versions targeting high Power Distance cultures (top, Vietnamese version) and low Power Distance cultures (bottom, German version). Photos and partner logos were anonymized for this paper.

do not only circle around a usable design but favour a pleasant and emotional experience [72]. With an increasing focus on pleasurable products and hedonic design attributes beyond traditional usability aspects, user-centered design became more and more complex. Particularly in our digital and widely connected online world, where competing services are just a few clicks away, user interface (UI) designers can no longer primarily concentrate on a clickable and usable design but need to take further UI elements such as visuals, content structure, information architecture, and more into account (see Figure 1). It is not sufficient anymore to understand ergonomic human factors in design. Instead, users' evaluation of a UI today depends on many further factors, such as the emotional state, prior experiences, expectations, age, gender, and culture [69, 79]. Researchers and designers from both academia and industry use diverse tools and methods, ranging from interviews and observations to questionnaires and data logging [7, 66, 69, 80, 93] to study, understand, and consider these influencing factors.

In general, design teams can observe and measure the impact of different factors by a suitable study setup and choice of study

participants. The consideration of culture, however, usually requires time-consuming and costly ethnographic analyses [37, 76, 102]. In particular, such first-hand research is often not feasible and scalable for smaller teams and companies that do not have the chance to rely on team members from respective target countries. Consequently, numerous researchers discuss the consideration of cultural theory, particularly Hofstede's cultural dimensions [34], during the design process to cope with the increasing effort and costs [20, 44, 68, 76, 98, 100]. At the same time, further perspectives state concerns about the blind application of cultural theory in design (see, e.g., Nwokoye et al. [65]). In an age of "postcolonial computing", where culture goes beyond national boundaries, Irani et al. [38] emphasize that the core of cross-cultural considerations is engagement. In this light, we want to address the research question:

"How can cultural theory support a culturally sensitive UI design process?"

To address our research question, we partnered with a small social start-up in order to identify country-specific UI design aspects of their website for the two pre-defined target countries Germany and Vietnam. We present results from (1) an analysis of cultural theory, particularly cultural dimensions according to Hofstede and colleagues [34], an analysis of German and Vietnamese question-and-answer (Q&A) websites and (2) a case study with 14 German and 14 Vietnamese participants to evaluate two UI concepts with differing information architecture, navigation structure, and visual presentation based on the previous analysis (see Figure 1).

This paper contributes an analysis of cultural theory and its application in UI design. For this purpose, we analyze cultural dimensions in the context of Human-Computer Interaction (HCI) and derive design insights for culturally sensitive UI design in a German and Vietnamese context. In addition, we discuss implications for HCI in further contexts (e.g., Africa) as well as future work for culturally sensitive design. We have chosen to establish the term *culturally sensitive design* for this work, similar to concept of value sensitive design according to Friedman et al. [21], to emphasize that we aim to foster a mindset that accounts for cultural differences during the design process.

2 BACKGROUND AND RELATED WORK

The role of culture in HCI has been widely discussed in both academia and industry. Companies in many industries have developed localization strategies for their products and marketing campaigns since a long time. In addition, academic studies underpin the impact of cultural differences on product design [74, 77]. However, a common theoretical basis and understanding of cultural aspects in HCI has not yet been established. In fact, researchers still controversially discuss how existing cultural theory can be applied in HCI (see Winschiers [98] and Ford & Kotzé [20]).

To better understand the scope of cultural issues in HCI and to embed our research question into the academic discourse we, first, summarize how the shift from a usability to a User Experience (UX) perspective changed the way we need to look at culture in HCI (see Figure 2). Second, we present basic conceptualizations of culture and some associated design approaches. Next, we reflect on how the established concept of cultural dimensions is generally applied in user-centered design processes.

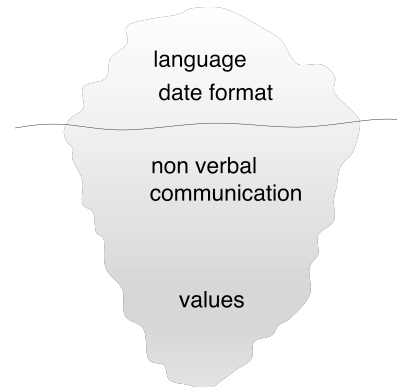


Figure 2: Different layers of culture according to Hofst [35].

2.1 From Usability to User Experience

The way we design and evaluate user interfaces has significantly changed within the last decades. Researchers and practitioners alike traditionally looked at UI from a usability perspective based on the credo that "to measure is to know" (see Law et al. [50]). As a result, the field of HCI was initially driven by engineers, designers, and researchers that focused on the analysis of a product's characteristics (e.g., functionality, ergonomics) [9, 48]. In fact, before the turn of the century, UI design was primarily based on a set of selected design heuristics, e.g., consistency, feedback mechanisms, and error prevention (see Nielsen & Molich [61]). Nowadays, however, usability is generally taken for granted while pleasurable and hedonic product attributes became crucial for customer loyalty and product success [5, 8, 72]. Consequently, the field of HCI and UI design is recently guided by a much more general understanding about users' experiences. Yet, definitions and evaluation approaches of the resulting concept of UX still range from a psychological perspective on human needs (see Hassenzahl [30] and Sheldon et al. [84]) to a task-oriented perspective on users' goals and motivation [31]. In addition, the shift from a usability to a UX-focused UI design process forces UI designers to take further aspects into account. Besides the experience during the actual interaction, UX can also be influenced by prior experiences or expectations. Furthermore, UX is unique to an individual user and rooted in a cultural context [79]. In this work, we base our analysis of culturally sensitive UI design on the differentiation of pragmatic and hedonic aspects - including both experience- and task-oriented product characteristics - according to Hassenzahl [30].

2.2 Culture in the Context of HCI

At the same time when HCI researchers and practitioners started to shift from a usability to a UX mindset, a discussion about the role of culture when designing for international user groups started to arise [62]. However, the landscape of cultural theory in the context of HCI is controversially discussed and lacks a common understanding or guidelines [36, 78]. Definitions of culture range from "a system of meaning that underlies routine and behaviour in everyday working life" [10] over "race and ethnicity as well as other variables

[...] manifested in customary behaviours, assumptions and values, patterns of thinking and communication style" [11] to "the collective programming of the mind that distinguishes the members of one group or category of people from another" [32]. A common denominator of these perspectives is the existence of visible, conscious variables (e.g., number formatting, currency, time, date formats, and language) and abstract, unconscious variables of culture (e.g., non-verbal communication, a sense of time, and physical distances) [36]. The iceberg meta model from Hofst [35] illustrates the different layers of culture (see Figure 2).

In the field of HCI and UI design, the multilayered nature of the concept of culture can be associated with traditional usability and in-depth UX considerations. While visible layers, i.e., language, meaning of colors, etc., can easily be adapted to create country- or culture-specific designs [6], the consideration of the abstract invisible layers (i.e., values) and its impact on UI design generally requires more time and resources [78]. Furthermore, the concept of culture can be addressed and studied on different levels, such as the national, regional, gender, generation, social class, or corporate level [33]. In this paper, we focus on the concept of national cultures, because research has evidenced its impact on the perception of what constitutes good design [78].

2.3 Cross-Cultural Design Approaches

With a change of design criteria from task-oriented usability heuristics to hedonic experience attributes, the method toolkit of designers and researchers grew proportionally in order to cope with the increasing complexity and dynamics of the design processes [7, 70]. Different methods or approaches are certainly more suitable in distinct contexts and situations. Nevertheless, Battarbee & Koskinen [8] cluster three general approaches for the consideration of UX in design process:

- *Measuring approaches* that base their analysis on quantifiable aspects of UX, e.g., emotion detection and self-reporting.
- *Empathic approaches* that aim to truly understand users' needs through meaningful user-designer relationships.
- *Pragmatist approaches* that focus on the application of theoretical frameworks and models.

In cross-cultural UI design, measuring approaches are usually based on quantitative analyses of different cultures through, e.g., log analysis (e.g., Lachner et al. [47]), evaluation of website aesthetics (e.g., Nordhoff et al. [63]), or international survey studies (e.g., Al-Shamaileh & Sutcliffe [3], Reinecke & Gajos [77], or Walsh & Nurkka [95]). An inherent limitation of such studies is the difficulty to understand why certain differences occur. In contrast, empathic approaches, such as ethnographic interviews (see, e.g., Schneider et al. [82]) or qualitative lab studies (e.g., Athinen et al. [2]) can provide rich stories and insights. However, such approaches are unavoidably time-consuming and costly or require further validation [37, 76, 97, 102]. Nevertheless, measuring-focused remote studies are often used in settings, in which on-site studies would be challenging to conduct due to the distance between researchers and study participants. Examples include the exploration of UX of a learning service in South Africa by researchers based in Finland and the UK [92] or the analysis of websites in Muslim countries by Malaysian researchers [59]. Empathy-focused ethnographic studies

are often favored in settings that require high ethical considerations, such as for research related to minorities or developing countries, e.g., the analysis of usability in Namibia [98], the investigation of apartheid-era narratives in South Africa [49], the evaluation of digital mobile maps in sub-Saharan Africa [88], or the study of mobile banking of low-literate, low-income users [58].

Pragmatic approaches based on theoretical frameworks have gained increasing interest in the HCI community in recent years. More and more researchers study the applicability of cultural dimensions in UX and UI design [52, 54, 75, 96]. Cultural dimensions are individual traits that characterize a culture's preferred reaction to general societal problems that are common to all cultures [32]. Several sociologists and anthropologists derived distinct cultural dimensions that have been established in academic research since then, namely Edward Hall [27–29], Geert Hofstede and colleagues [33], Kluckhohn & Strodtbeck [45], Shalom Schwartz [83], David Victor [94], and Trompenaars & Hampden-Turner [90]. Previous cultural considerations in HCI mainly builds upon the study of cultural dimensions according to Hofstede et al. [33] as their work represents the most comprehensive cross-cultural study that is mainly used in HCI (see, e.g., [52, 54, 67, 75, 76, 100]). Hofstede et al. [34] describe six cultural dimensions:

- *Power Distance (PD)*: The degree to which hierarchy and an unequal distribution of power is accepted in a society.
- *Individualism vs. Collectivism (ID)*: Preference towards a loosely-knit vs. tightly-knit social framework.
- *Masculinity vs. Femininity (MAS)*: The balance to which societies strive for status (masculine) or modesty (feminine).
- *Uncertainty Avoidance (UA)*: The degree to which societies feel uncomfortable with uncertain situations.
- *Long Term vs. Short Term Orientation (LTO)*: Preference towards traditions (short term) or societal change (long term).
- *Indulgence vs. Restraint (IN)*: The extent of vitality and fun (indulgence) vs. strict social norms (restraint).

Kahn et al. [43], for example, use Hofstede's cultural dimensions as an inspiration to develop automotive human-machine interfaces (HMI) for users from the UK and India, George et al. [23] identify differing preferences in web design for Australia as a whole and an indigenous Australian group in 2010 and validate ethnographic results with theoretical insights in 2012 [22], Jaramillo-Bernal et al. [41] develop a design framework based on Hofstede's dimensions, Yeo [103] uses Hofstede's dimensions to explain cultural differences in software development processes, Gould et al. [25] derive design guidelines for Malaysian and US websites from Hofstede's dimension website design, and Suadamara et al. [89] describe a process to integrate Hofstede's dimensions into the Technology Acceptance Model (TAM) according to Evers & Day [18].

2.4 Reflection on Cultural Dimensions for HCI

Our overall goal is to apply cultural theory in such a way that it supports the conceptual model of designers (or the represented model) to better understand how users from other cultures might interpret the way a product should be used, i.e., their mental models (see Cooper et al. [15]). While ethnographic on-site studies or internationally staffed design teams might best allow to validate the associated design elements, not every team can afford this due to

team size, time pressure, or costs. However, as the understanding of the concept of UX differs between different cultures [74], we argue that it is inevitable for good design to be aware of cultural differences and culturally sensitive design elements. An understanding of cultural theory represents a cost-efficient and sustainable way to eventually develop a culturally sensitive design process.

In general, as Hofstede's cultural dimensions represent dichotomous scales, studies such as the work from Mimouni & MacDonald [60], Reinecke & Bernstein [76], and Walsh & Nurkka [95] show that UI designs should differ in relation to the relative scale values. In contrast, Ford & Kotzé [20] and Ford & Gelderblom [19] argue that the design of a UI can generally be improved simply by focusing on high values for distinct cultural dimensions. The latter two, however, base their study on the analysis of a website's usability and human performance, whereas we additionally focus on further UX-oriented aspects. Although Marcus & Hamoodi [55] observed inconsistent cases for Hofstede's cultural dimensions, the work of Hofstede and colleagues has experienced most attention in HCI in recent years [76]. Consequently, we decided to base our evaluation study and associated research question on Hofstede's work. However, the majority of cross-cultural HCI studies focuses on post-hoc analyses of existing websites to draw a conclusion on the relation between culture and design (see, e.g., Gevorgyan & Porter [24] or Oliveira et al. [67]). In contrast, we want to investigate how we can translate theoretical insights derived from the analysis of cultural dimensions into the development of new UI designs to ultimately embed cultural theory in the design process.

3 A CASE STUDY FOR CULTURALLY SENSITIVE UI DESIGN

The goal of this paper is to understand how cultural theory can be used to support a culturally sensitive design process. In line with our research question, we analyzed the theoretical background to understand the design space of culturally sensitive UI design and derived design hypotheses to create two UI prototypes that focused on different cultural aspects. We aimed to investigate if and how cultural dimensions are a suitable starting point to design culturally sensitive interfaces. Our study results show that German and Vietnamese participants varied in their evaluation of selected design elements that we considered as culturally sensitive already during the conceptualization of the UI design.

3.1 Background and Setting

At the beginning of this project we partnered with a social start-up that offers an online Q&A-website. The website allows people from around the world to share and discuss solutions for global social issues, e.g., refugee crisis, global warming, or world hunger. Users and visitors can freely access the platform to (1) browse and click through different topics, questions, and answers for different topics and (2) comment on or start a new discussion. As the platform aims to address users and contributors from various countries, the goal of our partner was to identify culturally sensitive design elements for future development. In order not to interfere with the ongoing UI design and development process we decided to make our own adapted version of this site and study the impact of culture in UI design in a well-controlled setting. In addition we decided to narrow

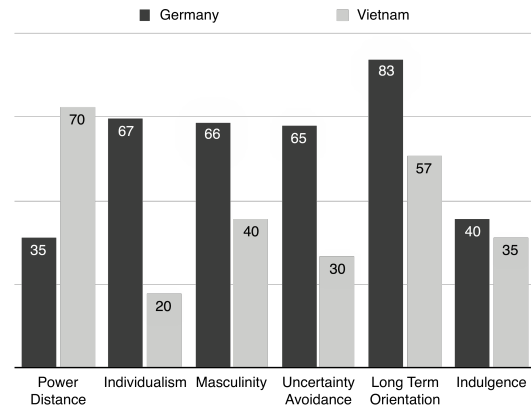


Figure 3: Values for Germany and Vietnam in the six cultural dimensions according to Hofstede et al. [34] (values can range from 0, low, to 100, high).

down the scope of this study to eventually derive concrete design insights and specific starting points for future work. Consequently, together with our partner, we decided to limit our case study to the investigation of Germany and Vietnam. Representing two empirically diverse cultures, we considered Germany and Vietnam as suitable comparison countries for this project. In addition, due to existing collaborations and the authors' personal backgrounds we were able to recruit suitable study participants from both countries for our evaluation study.

3.2 Design Hypotheses from Cultural Theory

Our case study was divided into several different steps to address our research question in a structured manner. Below, we describe the reasoning behind the initial analysis of cultural dimensions.

3.2.1 Procedure. As a first step, we analyzed established cultural dimensions (as presented above) and decided to base our case study on the dimensions according to Hofstede and colleagues [34]. In this case study, we consider cultural dimensions as a means to enhance designers' mental models that enable a better understanding of culture to ultimately support the development of a culturally sensitive UI for our partner's Q&A-website. We focused on only one cultural dimension, namely Power Distance (PD), to decrease the complexity of our hypotheses and the associated UI designs. We chose PD for our case study as the two target countries Germany and Vietnam differ significantly in their PD values and Mimouni & MacDonald [60] evidenced that PD has an influence on UI design.

We used the framework of Marcus & Gould [54] (that previous studies, e.g., Alexander et al. [4], have also used to evaluate existing websites in a cross-cultural context) as a starting point for the development of culturally sensitive UI prototypes. More precisely, we translated the design guidelines related to Hofstede's cultural dimensions according to Marcus & Gould [53, 54] into website elements for cultures with a high (e.g., Vietnam) and a low value (e.g., Germany) for PD (see Figure 3).

Table 1: Design aspects for high and low PD cultures according to Marcus & Gould [53].

	Low PD	High PD
<i>Metaphor</i>	Objects that represent free choice and equality	Objects that represent hierarchy and distance
<i>Mental Model</i>	Simple, informal, less structured, organized	Complex, highly organized and categorized
<i>Navigation</i>	Flexible paths, multiple choices	Restricted access, predefined paths
<i>Interaction</i>	Helpful error messages, keywords	Severe error message
<i>Presentation</i>	Pictures of groups, individuals, established symbols, informal language	Pictures of leaders, logos, official symbols, formal language

3.2.2 Design Hypotheses. Based on Hofstede’s cultural dimensions, Germany and Vietnam can be considered culturally diverse (see Figure 3). With higher scores for individualism, masculinity, uncertainty avoidance, and long term orientation, Germany can be seen as a country where self-actualization is strongly believed, performance is highly valued, systematic overview in thinking, presenting, and planning is preferred, and where people believe that truth is dependent on context and time. In addition, a participate communication style is common, leadership can be challenged and people tend towards cynicism and pessimism - characterized by a lower value for power distance and indulgence. In Vietnam, people similarly tend towards cynicism and pessimism (high value for indulgence) and generally base truth on context and time (high value for long term orientation). In contrast, Vietnamese accept hierarchical order and inequalities (low value for power distance), foster strong relationships (low value for individualism, i.e., collectivistic culture), value equality and well-being (low value for masculinity, i.e., feminine culture), and have a low preference for avoiding uncertainties [34]. For this case study and our focus on Power Distance, we are mainly interested in how preferences regarding hierarchical structures (i.e., PD) may impact UI design. Marcus & Gould [53, 54] analyzed how relative differences in a culture’s perception with regards to cultural dimensions can be translated into UI design elements. Based on their work, we can derive the hypothesis that low PD cultures (e.g., Germany) value informally organized and categorized data, overview, flexible navigation, helpful error messages, and graphics representing groups rather than selected leaders, whereas high PD cultures (e.g., Vietnam) prefer complex structures, pre-defined navigation paths, direct error messages, symbols, logos, and graphics that represent leaders (see Table 1). We used these guidelines as a starting point for the development of two distinct interfaces of a Q&A-website.

3.3 Benchmark with Existing Q&A-Websites

As a next step, we analyzed existing German and Vietnamese Q&A websites in order to enrich our theoretical insights and calibrate our perspective on how to translate theoretical insights into concrete design elements.

Table 2: Websites for our benchmark analysis.

German websites	Vietnamese websites
www.gutefrage.net	www.webtretho.com/forum
www.chefkoch.de	www.tinhte.vn/forums
www.computerbase.de	www.vforum.vn
www.motor-talk.de	www.sinhvienit.net/home
www.android-hilfe.de	www.vozforum.org
www.forum.chip.de	www.otofun.net/forums
www.hifi-forum.de	www.violet.vn
www.fotocommunity.de/forum	www.hdvietsnam.com
www.board.gulli.com	www.tivnol.com
www.wer-weiss-was.de	www.lamchame.com/forum

3.3.1 Procedure. We selected frequently used online forums in Germany and Vietnam, as such websites best represent our partner’s platform. We manually searched for forums using the search engine Google and the search queries *most common/popular/visited forums/question-and-answer websites in Germany/Vietnam* and *German/Vietnamese forums/question-and-answer websites* to first identify compiled lists of suitable websites. After identifying popular Q&A-websites, we ranked them according to the number of website visits as stated at <https://www.similarweb.com/>. In this ranking, we only considered websites that had more than 60% local website visitors. Finally, we decided to analyze the top ten websites in detail (see Table 2). For this analysis, we looked at all selected websites and manually derived design characteristics of each website according to the framework of Marcus & Gould [53], i.e., metaphor, mental model, navigation, interaction, and presentation.

In general, our website analysis served as a reference, calibration of our mental model, and comparison of the design insights that we derived from our analysis of cultural dimensions according to Marcus & Gould [54]. We did not aim to holistically describe significant differences and similarities of German and Vietnamese websites in this step. Consequently, we only manually analyzed a few selected websites for each country.

3.3.2 Status Quo of German and Vietnamese Q&A-websites. Our analysis of existing websites allowed us to better understand familiar design elements for German and Vietnamese users. Table 2 gives an overview of all German and Vietnamese websites that we have analyzed in this step. We used the same framework, i.e., the analysis of metaphors, mental model, navigation, interaction, and presentation according to Marcus & Gould [53], as before. Also, we focused again on the analysis of design elements associated with the cultural dimension PD.

We recognized that, e.g., leading organizations, administrators or moderators are rarely prominently presented on German Q&A-websites. The websites are rather targeted towards the general user base (e.g., gutefrage.net), both in the structure of the websites and the visual presentation. However, contrasting with our theoretical hypotheses we saw that in our sample set most of the websites were highly structured with many sub-levels and topic categories.

Yet, the general navigation structure offered several flexible paths, including search feature(s), drop-down menus, and direct links to, e.g., related content (e.g., computerbase.de). Mimouni & MacDonald [60] describe similar insights for website navigation in relation to cultural dimensions in their analysis of American and Arabic websites. The language used on the selected websites was generally polite and error messages even provided helpful comments in many cases (e.g., wer-weiss-was.de).

Q&A-websites that are popular in Vietnam, in contrast, generally did emphasize moderators and administrators (e.g., webtretho.com highlights admins for each topic), navigation paths are less flexible and related content is rarely linked (e.g., vozforums.com). Furthermore, information is usually highly structured, e.g., into categories and sub-categories. However, we could not derive differing insights with regards to a distinct formulation of error messages or language.

3.4 Culturally Sensitive Prototype Design

Based on the derived insights from our theoretical analysis, we developed two UI prototypes that differed in navigation structure, visual presentation, and language. In addition, we particularly focused on content structure and the design of error messages as our comparison of cross-cultural design heuristics and existing Q&A-websites pointed out mixed results.

3.4.1 Procedure. We aimed to design two distinct yet similar UIs based on design elements for (1) a high and (2) a low value for PD according to Marcus & Gould [54]. We primarily focused on the translation of theoretical design insights to address our overall research question yet used our benchmark analysis of existing websites to align the overall structure and presentation of the prototype concepts to familiar Q&A-websites. Both UI prototypes were translated into a German and a Vietnamese website by native speakers resulting in four different versions (see Figure 1). The information architecture, visual design, and navigation structure were not changed for the two Vietnamese or the two German versions. All prototypes were designed using the design software Axure¹.

3.4.2 Prototypes. We used the existing design and corporate identity (e.g., colors, font, etc.) of our collaboration partner for both prototypes to avoid biases but adapted selected elements according to our previous analysis. Our partner's website focuses on different social issues and global problems. For our case study, we decided to only focus on the topic 'refugee crisis', as it was the a key topic of our partner's strategy at the time of the study, as well as a topic of global interest. Furthermore, we created and used a number of test profiles and articles to provide a suitable amount of content.

For the first version, we focused on high Power Distance elements. We aimed to prominently position the administrators (in our case the founders of the platform), used pictures and logos of organizations represented on the website, and added only a few selected interaction possibilities (e.g., search for a solution, create an article). The main topics on the home page were arranged in a grid format. However, articles for a distinct topic were only listed one after another without further filter options. The error message for a login and a search task that we used in our case study did not provide further helpful information in this design (see Figure 4).

¹<https://www.axure.com/>

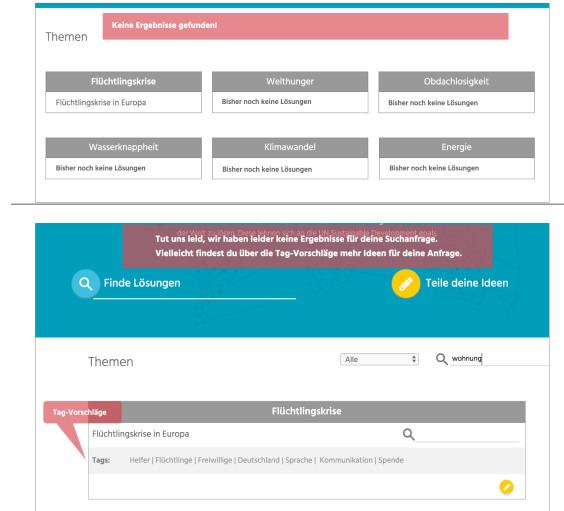


Figure 4: Different types of error messages in our two German UI prototypes, without (top) and with (bottom) further information (Vietnamese error messages were translated accordingly).

Also, the introduction about the platform at the main page was written formally and in a succinct style. We did not consider Marcus & Gould's [53] guideline for Mental Models (i.e., flat or highly structured UI) as we had defined our case study around one focus topic, and hence lacked enough content and categories that would have been necessary to be structured accordingly. In sum, the first version (high PD) was mainly defined by the following elements:

- Administrators are prominently presented
- Few alternative navigation paths
- Error messages are short and direct
- Language is formal and distanced

The second version was based on low Power Distance characteristics. In this version, we did not add photos of the administrators on the home page or logos of organizations but provided additional navigation elements, such as links to different sub-categories as well as tags, filters, and breadcrumbs. The selected topics on the home page were arranged as a list to have enough space to add additional search bars and tags. In sum, the second version (low Power Distance) was based on the following parameters:

- No focus on authority figures
- Many alternative navigation paths
- Error messages provide further information
- Language is informal and personal

3.5 Evaluation and Results

We conducted a think-aloud study to evaluate how study participants from Vietnam and Germany perceived the two different UI prototypes. Based on our research question we wanted to understand how the feedback of the study participants differed and whether we would be able to anticipate these differences through the previous analysis of cultural theory.

3.5.1 Study Participants. In total, we recruited 14 study participants from Germany and 14 from Vietnam for our think-aloud study through the network and mailing lists of two collaborating universities in Germany and Vietnam. We decided to pursue a qualitative approach as it allows to understand the reasoning behind users' experiences and related design issues. This represents a substantial aspect to understand if we properly applied cultural theory for our UI design, and hence suitably addressed our research question (see Law et al. [50]). For our study, we only accepted study participants that have not yet lived abroad (either privately or work/study related) for longer than for 6 months. In total, we had 60% male and 40% female study participants with an average age of 23.6 years. Table 3 shows an overview of all study participants as well as the average age and gender distribution per country.

3.5.2 Think-Aloud Procedure. We conducted a within-subjects think-aloud study with a 2 UI design (high PD vs. low PD) \times 2 nationality (German vs. Vietnamese) design. To avoid language effects, we presented both prototype versions (i.e., high and low PD versions) to all study participants in their native language. Also, we randomized the order in which the participants evaluated both prototypes to avoid biases. We were able to run the study with Vietnamese participants on-site in the facilities of the Vietnamese university that supported us for participant recruiting. The Vietnamese study participants were invited to a study room and used one of the authors' laptop to interact with the prototypes. To stay within our time schedule we were forced to interview the German participants remotely using the IP telephone and screen sharing service Skype as none of the authors was in Germany at the time the study was conducted. However, we wanted to ensure and hence prioritized that all study participants were situated in a familiar environment, i.e., their home country. For this remote study, we sent all HTML files of the German prototypes to the participants and let them interact with the prototype on their own laptop or computer. In addition, to cope with an initially low response rate for our first call for participation, German participants took part in a \$60 shopping voucher lottery. Vietnamese study participants did not receive a compensation as we had many returns within a short time for our first call for participation hence we had conducted the interviews already before the second call for participation for German users.

During the study, we asked all participants to think out loud while they interacted with the prototypes and performed several tasks. First, they were asked to browse through the start screen and then look for an answer for a given topic. Second, they were told to log in into a default user account and, third, enter a search term in the prototype's search bar. In these two cases, an error message appeared that differed for both versions, i.e., for the low and high PD version. Fourth, they were asked to describe which version they preferred in general after evaluating both the low PD and high PD version before they filled out a questionnaire about their personal data. Once again, we conducted the think-aloud study with every study participant in their mother tongue. However, four Vietnamese study participants wished to conduct the study in English. The Vietnamese think-aloud studies lasted between 40 and 55 minutes with an average of 45 minutes, the studies with German participants lasted between 30 and 45 minutes with an average of 34 minutes.

Table 3: Demographic data of study participants.

	Vietnam	Germany
<i>n</i>	14	14
Gender	(m) 64% (f) 36%	(m) 57% (f) 43%
Age Range	19 - 24 yrs	20 - 31 yrs
Average	21 yrs	26 yrs

3.5.3 Data Analysis. All think-aloud sessions were recorded, transcribed, analyzed through grounded theory, and finally translated from German and Vietnamese to English for this paper. The two authors who conducted the think-aloud study individually coded all statements using the coding categories *Metaphor*, *Mental Model*, *Navigation*, *Interaction*, *Presentation*, and *General Comment* to ensure inter-rater reliability (IRR). The coding categories were derived from the framework of Marcus & Gould [54] that had already been used for the development of the prototypes. Based on these categories, we were able to cluster and identify common statements through inductive category formation according to Mayring & Fenzl [56] to finally derive differences and similarities in the answers of the German and Vietnamese participants. Our IRR analysis resulted in a value for $\alpha = .8963$, 95% in a CI of (0.8224, 0.9120). According to Krippendorff [46], values for α that are higher than .8 can be seen as satisfactory.

3.5.4 Results. The overall goal of our think-aloud study was to evaluate (1) whether participants from Germany preferred different UI aspects compared to the Vietnamese participants and (2) whether we were able to consider these differences already during the design of the UI through the analysis of cultural dimensions. We did not specifically ask our study participants how much they liked distinct design and interaction aspects but rather motivated them to think out loud while they performed our tasks. In addition, we were interested in both usability-focused pragmatic aspects and experience-focused hedonic aspects and inquired reasons for the participants' opinions and statements.

All participants naturally commented on the features of interest related to our design hypotheses due to the nature of the case study tasks (e.g., search for a specific answer or topic, login attempt that prompted an error message). In particular, when they were confronted with the respective second prototype (either version 1 or version 2, depending on order) the participants were able to easily verbalize differences and their preferred concepts. In general, all participants from Vietnam (PV) and Germany (PG) gave feedback about the visual presentation (e.g., the photos of the social start-up founders) the structure and navigation (e.g., the search bar, tags, and filters), as well as the perceived trustworthiness of the website designs. Both concepts were generally well perceived and we had only few comments related to the overall design concept, i.e., our partner's corporate identity, or the fidelity of the prototype (e.g., "the font is too small" (PV13), "I like the colors" (PG3, PV3), or "the wording 'urgent issues' sounds weird" (PG3)).

A differentiated look at our results, however, revealed perceptible differences between the Vietnamese and German participants.

Table 4: Exemplary statements from the study participants per country for the derived insight and number of participants who preferred version 1 or 2 (*numbers do not add up to 28 as 3 participants did not state a clear preference).

	Vietnam		Germany	
	Version 1 (high PD)	Version 2 (low PD)	Version 1 (high PD)	Version 2 (low PD)
	Preferred emotional language style		Preferred factual language style	
Content, language, and communication	"I was surprised. A more friendly sentence would have been better, e.g., I am sorry, we did not find [...]" (PV11)	"It is friendly and the users understands the problem." (PV8) "I think this information is better." (PV12)	"I would have liked to see that the website helps me [...] and suggests additional links to other websites." (PG2)	"I think it is normal that [websites] say if the user name of password is wrong. This states actually the same." (PG6)
	Restricted information density		Guided information density	
Information retrieval	"I have no problem [with this design] as I have the same feeling as I have on Google or Wikipedia." (PV7)	"There are more search functions [...] and it is not easy to use." (PV9)	"I would have preferred to have a navigation menu, [the website] becomes easily confusing with more text." (PG7)	"I would tend to version 2 because I can add tags. [...] I have a better overview when I do not use the search bar." (PG6)
	Trust through emotions and visuals		Trust through content and validity	
Trust	"The photos of the founders make the website look more professional. I like the logos of the organizations." (PV1)	"Information about the founders of the website is necessary [...] to increase the level of trust." (PV6)	"I think it is good that there is the source [of the information]. It makes it trustworthy." (PG7)	"[The photos] can be interesting in general but when I go to such a website I don't think it is important." (PG6)
preferred version*	8	4	4	9

In line with our design hypotheses and as indicated in Table 4, the majority of the Vietnamese participants preferred the design version 1 (focusing on high PD) and the majority of the German participants the design version 2 (focusing on low PD). An in-depth analysis of all coded think-aloud protocols allowed us to derive three main topics that summarize differing tendencies between the Vietnamese and the German study participants (Table 4 provides an overview of all three insights including exemplary comments from participants):

- (1) Emotional vs. factual language style
- (2) Restricted vs. guided information density
- (3) Trust through emotions vs. trust through content

Emotional vs. factual language style: During the think-aloud study, all participants were faced with two planned error message, one as the result of an intentionally failed login attempt and another one after using the search bar on the home page. Based on our theoretical analysis we expected Vietnamese participants to prefer the short and direct error message in version 1 and German participants message with additional information about the error type in version 2. However, in contrast to our expectations, the error message in version 2 (low PD) was generally perceived familiar or even better (PV2, PV5-8, PV11-14, PG2-4, PG6-7, PG10-13), as *"it is friendly"* (PV8), *"the [additional] information is helpful to detect the error"* (PV10) and it *"makes it clear that [the participant] has to review his input"* (PG11).

Nevertheless, the study participants from Vietnam and Germany mentioned different reasons why they were shocked or disappointed (PV1, PV4, PV8, PV11, PG2, PG4, PG7) when they saw the error messages in version 1. German participants did primarily ask for factual information, e.g., *"links to other websites"* (PG2) as *"related content"* (PG4) and *"alternatives [are] missing"* (PG6) or even *"a prompt to add a new solution"* (PG11). Vietnamese participants,

in contrast, rather noticed a lack of emotions and suggested to *"use different words"* (PV9) or more precisely a *"more friendly sentence"* (PV11). PV6 even suggested to use emoticons or stickers. However, the error messages did not strongly affect the overall evaluation of the Vietnamese study participants from which the majority still favoured version 1 (see Table 4).

Restricted vs. guided information density: The main tasks of our think-aloud study included the search process to find a specific question and related answers on the websites. Consequently, all participants had to familiarize with the platform and its navigation structure. We saw that the navigation had a stronger impact on the final evaluation of the designs. Vietnamese participants generally valued the *"clear, concrete, and direct"* (PV2) structure of version 1 that makes it *"comfortable [and] easy to use"* (PV6). They also highlighted to prefer using the search bar (PV1, PV2, PV4, PV7, PV9, PV12). PV7 even mentioned that the general concept reminds her of Google or Wikipedia. PV10, however, still concluded that it *"would be good to have more sub-categories"* as it would make the search process easier. In contrast, version 2 was perceived as *"too long and difficult"* (PV12) and *"not easy to use"* (PV9). PV9 suggests, among other things, that the read-on button of article should not link to another page but open as a flip-out menu instead, PV13 highlights that the list of categories in version 2 is too long. In sum, our Vietnamese study group preferred a restricted information density.

German participants rarely used the search functions as, e.g., they *"prefer to click through the topics on the website and browse through the articles"* (PG1). In contrast, they generally preferred an independent yet guided step-by-step information search. Consequently, several German participants positively valued the tag and filter possibilities in version 2 (PG1-4, PG6) as it allows to *"get fast to the respective topic"* (PV8) and one does not have to scroll through unimportant information (PG7, PG14). In addition, PG14

highlights that it is important to have an overview about all the data. Also, PG2 appreciated the breadcrumbs in version 2, while PG12 suggests that an additional navigation bar would have made the search process even easier. PG 13 summarizes that he just *"like on Amazon [...] first looked for sub-categories"* and then fine-tuned his search.

Trust through emotions vs. trust through content: The main aspect of cultures with different perceptions of PD is a varying acceptance of hierarchy in society. According to Marcus & Gould [53] this is reflected in images of hierarchically higher people. In our case study, we received divergent feedback from Vietnamese and Germany participants yet in line with the theoretical hypotheses.

The majority of our Vietnamese participant group appreciated the photos of the platform founders in version 1 as they *"make the website look more professional"* (PV1). PV6 highlights that even more information about the founders should be provided to *"increase the level of trust"*. Additionally, the logos of organizations for each article increased the credibility of the platform (PV1, PV3). In contrast, the lack of the photos in version 2 was frequently stated as a negative aspect of the version that was based on low PD design aspects (PV6, PV9, PV10, PV13).

From the German participants, only a few stated that the photos increased the reliability of the website yet take up too much space on the home page (PG3). In fact, many participants claimed that they do not necessarily favour the photos of the founders on the home page (PG6-9, PG13) or even think they are too dominant (PG1). However, the sources of the articles (e.g., links or names of the organizations) increased trust and validity of the content on the platform (PG1, PG2, PG7).

4 DISCUSSION AND LIMITATIONS

In this work, we showcased the applicability of cultural dimensions in early stages of a UI design process. Our think-aloud study revealed that Vietnamese study participants tended to favour UI elements that we anticipated as suitable for high PD cultures, such as Vietnam, and German participants, vice versa, design elements for low PD cultures. Throughout the whole case study, we felt that the analysis of cultural dimensions represented a helpful framework to consider critical UI elements for a culturally sensitive design process. At the same time, we acknowledge that our decisions to limit the scope and complexity of our case study invariably brings other limitations and open questions.

4.1 Reflection About the Study Setting

Study setting and generalization of our insights? Due to our collaboration with a social start-up we defined our case study based on their Q&A-website. Different use cases or websites might require further analysis or even lead to difficulties in applying cultural dimensions. For future studies, we recommend and will continuously consider a validation of design insights, similar to our analysis of existing Q&A-websites. Likewise, as we focused our analysis on only one cultural dimension, further research is required to draw a conclusion on the applicability of other cultural dimensions. We decided to focus on only one dimensions to derive distinct design insights and to decrease and cope with the complexity of our cross-cultural case study. Also, we have chosen to focus our case study on the

analysis of Vietnamese and German study participants. Although the feedback from participants from other countries might lead to more in-depth insights, the controlled setting with selected target countries allowed us to specifically interpret our results as well as the value of both cultural dimensions and our benchmark analysis. Furthermore, we do not want to argue that websites should be adjusted for every single country but rather for regions with comparable cultural backgrounds. Consequently, we see our comparison of Germany and Vietnam as an initial use case and will further investigate how cultural dimensions are applicable for different culturally similar regions. Similarly, a quantitative analysis (e.g., a log analysis as conducted by Lachner et al. [47]) will provide additional insights and help to validate our derived design insights. In our study, however, we decided to conduct a qualitative study with at least 12 participants per participant group for data saturation in interview studies (as suggested by Guest et al. [26]) to better understand why certain design elements are preferred. Finally, we needed to conduct the study with German participants remotely as no author was able to interview our German participants on-site at the time we had scheduled the interviews. Although a remote setting might affect the implementation of the study, we designed all study tasks and questions in a way that participants were able to focus merely on the design of the websites.

Implementation of cultural theory in design processes? Our overall research question was guided by the motivation of implementing cultural theory in the UI design process. On the one hand we were able to see that cultural dimensions represent a helpful tool to introduce cultural consideration in the design process, on the other hand we still need to investigate how such theoretical constructs can be best combined with further methods. According to Pettersson et al. [69], particularly questionnaires and observations represent commonly used methods that, from our perspective, should be evaluated in this context. In our case study, we saw that our qualitative think-aloud study benefited from the initial analysis of cultural dimensions.

The role of the user? In addition, we see potential for further research on the impact of culture on user studies in general. In our case study, we realized that Vietnamese participants were less talkative during the think-aloud study. German study participants, however, were more talkative. Lewis [51] describes an indication for this observation as Vietnam is listed as a reactive culture (i.e., cultures that prefer to focus on respect, listen quietly, and react carefully) and Germany a linear-active culture (i.e., cultures that plan, systematize, and follow correct procedures) in his model for cultural considerations. In addition, Hall [28] describes Vietnam as a higher-context culture, Germany as a lower-context culture indicating that communication in Vietnam includes more implicit information than in Germany. We suggest to investigate implications on design evaluations in the future.

4.2 Reflection About Cultural Dimensions

Implications for HCI in further contexts? In our analysis of related work in the context of culturally sensitive design, we realized that cross-cultural considerations based on theoretical frameworks (e.g., cultural dimensions) show a tendency to investigate differences in Asia, Europe, and America (see, e.g., Calabrese et al. [13] who focus

on Brazil, Portugal, Angola, and Macau, Karacay-Aydin et al. [42] who study USA and Turkey, Singh et al. [86] who investigate USA and China, Singh, Zhao, and Hu [87] who compare China Japan USA, or Sachau & Hutchinson [81] with their study of USA and Mexico). Research in rural or developing regions, particularly in Africa, is rather based on ethnographic studies or remote analyses as previously discussed. However, we see high potential for culturally sensitive HCI research based on cultural theory in such settings. First, from our perspective, the analysis of cultural dimensions in the context of HCI can lead to a better understanding of good design in a globalized world, particularly as, e.g., in South Africa (see Pretorius et al. [73]), the field of UX and design still lacks appropriate knowledge and inadequate training. Second, a better understanding of cultural dimensions and cultural preferences will be beneficial for cross-collaboration as international and globally acting design teams will better understand potential biases and culturally diverse mental models.

Limitations of cultural dimensions? The application of cultural dimensions comes along with inherent limitations of Hofstede's work [91]. McSweeney [57] highlights that cultural dimensions are based on a concept of national culture, but Hofstede's data was gathered through a survey that was only sent to IBM employees. In addition, researchers argue that culture is a dynamic construct, whereas cultural dimensions describe a static taxonomy [12, 57, 85]. To cope with the dynamics of the term culture, Irani et al. [38] introduce the term "postcolonial computing", referring to an approach that is based on engagement, articulation, and translation. The concept is mainly driven by the fact that further aspects, such as, e.g., gender, ethnicity, race, or subculture may influence the overall construct of culture [39]. Chandra et al. [14], Jack & Jackson [40], and Wyche et al. [99], for instance, apply the concept of postcolonial computing and derive an in-depth ethnographic research approach for their studies. However, we do not see the concept of postcolonial computing as a contradicting approach but rather as an overall "tactic" [71]. Based on our results, we see cultural dimensions as a suitable starting point rather than a standardized framework for culturally sensitive considerations that are followed by further in-depth investigations, e.g., ethnographic research (similar to the approach of Schneider et al. [82]).

5 CONCLUSION AND FUTURE WORK

In this paper, we have presented a case study to investigate a culturally sensitive UI design process based on the analysis of cultural theory. In collaboration with a social start-up we developed two distinct UI prototypes for a Q&A-website and conducted a think-aloud study with 14 German and 14 Vietnamese study participants. The differing design elements in our prototypes were derived from the analysis of cultural dimensions. We found out that our study participants differed in their evaluation of information density, trust, and error handling. Our overall goal was to investigate how cultural theory can be applied during the UI design process.

In sum, we learned that the analysis of cultural dimensions helped us to anticipate differing feedback and, in particular, foresee critical design elements for a culturally sensitive design process in general. Consequently, we feel encouraged about the general value of cultural dimensions for a culturally sensitive design process.

However, in accordance with previous study results, such as the work of Winschiers [98], we would like to raise concern if cultural dimensions are applied blindly for the development of a new design. Although we were able to identify and foresee crucial design elements that were assessed differently by our culturally diverse study participants, we also found slightly different yet culturally sensitive design insights. In this context, our additional benchmark of existing websites helped us to calibrate our mental model and shape our design hypotheses. In general, we suggest to use cultural dimensions to develop a common language in interdisciplinary design teams, calibrate the mental models, and to inspire culturally sensitive design solutions. However, due to the dynamic nature of culture and values (see Irani et al. [38]), we want to motivate researchers, designers, and developers to conduct additional research in cross-cultural projects to fully understand the needs of culturally diverse user groups.

In the future, we see the potential to derive more insights from further studies in other culturally diverse countries as well as the investigation of more cultural dimensions or varying study settings. In addition, a research questions that we did not raise yet but that needs to be addressed is how technologies, such as machine learning that is more and more discussed in relation to UX and UI Design (see, e.g., Dove et al. [16] or Yang et al. [101]), can support a culturally sensitive design process, e.g., through the automated detection of personal and cultural traits based on behavioral data (comparable to the work of Epp et al. [17] in the field of emotion detection). Further topics that will be relevant for culturally sensitive design are a suitable balance of globally implemented hence corporate identity conform design elements and locally adapted aspects.

Reflecting on our approach, we feel confident that cultural dimensions allowed us to anticipate crucial UI elements with minimum costs and expenditure of time compared to more in-depth ethnographic approaches. Similar to the case studies presented by Yaaqoubi & Reinecke [100] we argue that the analysis of cultural dimensions is helpful in early design stages, particularly if additionally validated during the design process. Overall, we see our current work as complementary to previous results in the complex landscape of culturally sensitive design and as a fruitful starting point for future work in different contexts, continents, and based on more cultural theories to ultimately define the role of culture in HCI.

ACKNOWLEDGMENTS

We would like to thank our interview partners involved in the evaluation process, our collaboration partner as well as publications support and staff, who wrote and provided helpful comments on previous versions of this document. Also, we would like to thank our reviewers who provided helpful feedback to shape the final version of this work.

REFERENCES

- [1] Chadia Abras, Diane Maloney-Krichmar, and Jenny Preece. 2004. User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction*. Thousand Oaks: Sage Publications 37, 4 (2004), 445–456.
- [2] Aino Ahtinen, Shruti Ramiah, Jan Blom, and Minna Isomursu. 2008. Design of mobile wellness applications: identifying cross-cultural factors. In *Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat*. ACM, 164–171.

- [3] Ons Al-Shamaileh and Alistair Sutcliffe. 2012. Investigating a multi-faceted view of user experience. In *Proceedings of the 24th Australian Computer-Human Interaction Conference*. ACM, 9–18.
- [4] Rukshan Alexander, Nik Thompson, and David Murray. 2017. Towards cultural translation of websites: a large-scale study of Australian, Chinese, and Saudi Arabian design preferences. *Behaviour & Information Technology* 36, 4 (2017), 351–363.
- [5] Rui Alves, Pedro Valente, and Nuno Jardim Nunes. 2014. The state of user experience evaluation practice. In *Proc. of the 8th Nordic Conference on Human-Computer Interaction. NordiCHI '14*. ACM, 93–102.
- [6] W Barber and A Badre. 1998. The Merging of Culture and Usability. In *4th Conference on Human Factors and the Web, Basking Ridge, New Jersey*.
- [7] Javier A Bargas-Avila and Kasper Hornbæk. 2011. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In *Proc. of CHI 2011*. ACM, 2689–2698.
- [8] Katja Battarbee and Ilpo Koskinen. 2005. Co-experience: user experience as interaction. *CoDesign* 1, 1 (2005), 5–18.
- [9] Frøy Birte Bjørneseth, Mark D Dunlop, and Jann Peter Strand. 2008. Dynamic positioning systems: usability and interaction styles. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*. ACM, 43–52.
- [10] Keld Bødker and Jesper Strandgaard Pedersen. 1991. Workplace cultures: Looking at artifacts, symbols and practices. In *Design at work*. Lawrence Erlbaum Associates, Incorporated, 121–136.
- [11] Christine L Borgman. 1986. The user's mental model of an information retrieval system: an experiment on a prototype online catalog. *International Journal of man-machine studies* 24, 1 (1986), 47–64.
- [12] Paul Brewer and Sunil Venaiik. 2014. The ecological fallacy in national culture research. *Organization Studies* 35, 7 (2014), 1063–1086.
- [13] Armando Calabrese, Guendalina Capece, Francesca Di Pillo, and Federico Martino. 2014. Cultural adaptation of web design services as critical success factor for business excellence: A cross-cultural study of Portuguese, Brazilian, Angolan and Macanese web sites. *Cross Cultural Management* 21, 2 (2014), 172–190.
- [14] Priyank Chandra, Syed Ishtiaque Ahmed, and Joyojeet Pal. 2017. Market practices and the bazaar: Technology consumption in ICT markets in the global south. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. ACM, 4741–4752.
- [15] Alan Cooper, Robert Reimann, and David Cronin. 2007. *About face 3: the essentials of interaction design*. John Wiley & Sons.
- [16] Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX Design Innovation: Challenges for Working with Machine Learning As a Design Material. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, 278–288.
- [17] Clayton Epp, Michael Lippold, and Regan L. Mandryk. 2011. Identifying Emotional States Using Keystroke Dynamics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, 715–724.
- [18] Vanessa Evers and Donald Day. 1997. The role of culture in interface acceptance. In *Human-Computer Interaction INTERACT'97*. Springer, 260–267.
- [19] Gabrielle Ford and Helene Gelderblom. 2003. The effects of culture on performance achieved through the use of human computer interaction. In *Proc. of the 2003 annual research conference of the South African institute of computer scientists and information technologists on Enablement through technology*. South African Institute for Computer Scientists and Information Technologists, 218–230.
- [20] Gabrielle Ford and Paula Kotzé. 2005. Designing usable interfaces with cultural dimensions. In *IFIP Conf. on Human-Computer Interaction*. Springer, 713–726.
- [21] Batya Friedman, Peter H Kahn, Alan Borning, and Alina Hultdtgren. 2013. Value sensitive design and information systems. In *Early engagement and new technologies: Opening up the laboratory*. Springer, 55–95.
- [22] Reece George, Keith Nesbitt, Michael Donovan, and John Maynard. 2012. Evaluating indigenous design features using cultural dimensions. In *Proceedings of the Thirteenth Australasian User Interface Conference-Volume 126*. Australian Computer Society, Inc., 49–58.
- [23] Reece George, Keith Nesbitt, Patricia Gillard, and Michael Donovan. 2010. Identifying cultural design requirements for an Australian indigenous website. In *Proceedings of the Eleventh Australasian Conference on User Interface-Volume 106*. Australian Computer Society, Inc., 89–97.
- [24] Gennadi Gevorgyan and Lance V Porter. 2008. One size does not fit all: Culture and perceived importance of web design features. *Journal of Website Promotion* 3, 1-2 (2008), 25–38.
- [25] Emilie W Gould, Norhayati Zakaria, and Shafiz Affendi Mohd Yusof. 2000. Applying culture to website design: a comparison of Malaysian and US websites. In *Proc. of the 18th annual ACM international conference on Computer documentation: technology & teamwork*. IEEE Educational Activities Department, 161–171.
- [26] Greg Guest, Arwen Bunce, and Laura Johnson. 2006. How many interviews are enough? An experiment with data saturation and variability. *Field methods* 18, 1 (2006), 59–82.
- [27] Edward Twitchell Hall. 1966. *The hidden dimension*. Doubleday & Co.
- [28] Edward Twitchell Hall. 1989. *Beyond culture*. Anchor.
- [29] Edward Twitchell Hall. 1989. *The dance of life: The other dimension of time*. Anchor.
- [30] Marc Hassenzahl. 2008. User experience (UX): towards an experiential perspective on product quality. In *Proceedings of the 20th Conference on Interaction Homme-Machine*. ACM, 11–15.
- [31] Marc Hassenzahl and Noam Tractinsky. 2006. User experience - a research agenda. *Behaviour & information technology* 25, 2 (2006), 91–97.
- [32] Geert Hofstede. 2003. *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage publications.
- [33] Geert Hofstede, Gert Jan Hofstede, and Michael Minkov. 2010. *Cultures and Organizations. Software of the Mind. Intercultural Cooperation and Its Importance for Survival*. McGraw-Hill.
- [34] Hofstede Insights 2018. Solve Intercultural and Organisational Culture Challenges. Retrieved July 1, 2018 from <https://www.hofstede-insights.com/>
- [35] Nancy L Hoff. 1995. *International technical communication: How to export information about high technology*. John Wiley & Sons, Inc.
- [36] Nancy L. Hoff. 1996. International Users Interface. John Wiley & Sons, Inc., New York, NY, USA, Chapter Developing a Cultural Model, 41–73. <http://dl.acm.org/citation.cfm?id=241952.241954>
- [37] Kristina Höök. 2000. Steps to take before intelligent user interfaces become real. *Interacting with computers* 12, 4 (2000), 409–426.
- [38] Lilly Irami, Janet Vertesi, Paul Dourish, Kavita Philip, and Rebecca E Grinter. 2010. Postcolonial computing: a lens on design and development. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1311–1320.
- [39] Lilly C Irani and Paul Dourish. 2009. Postcolonial interculturality. In *Proceedings of the 2009 international workshop on Intercultural collaboration*. ACM, 249–252.
- [40] Margaret Jack and Steven J Jackson. 2016. Logistics as care and control: An investigation into the unicef supply division. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2209–2219.
- [41] Zayra Jaramillo-Bernal, Cesar A Collazos, Karla Arosemena, and Jaime Muñoz Arteaga. 2013. Methodological framework for design and evaluation of interactive systems from a multicultural and emotional perspective. In *Proceedings of the 2013 Chilean Conference on Human-Computer Interaction*. ACM, 60–65.
- [42] Gaye Karacay-Aydin, Elif Akben-Selçuk, and Asli Elif Aydin-Altınoklar. 2009. Cultural Variability in Web Content: A Comparative Analysis of American and Turkish Websites. *International Business Research* 3, 1 (2009), 97.
- [43] Tawhid Khan, Matthew Pitts, and Mark A Williams. 2016. Cross-cultural differences in automotive hmi design: a comparative study between uk and indian users' design preferences. *Journal of Usability Studies* 11, 2 (2016), 45–65.
- [44] Nouf Khashman and Andrew Large. 2011. Measuring cultural markers in Arabic government websites using Hofstede's cultural dimensions. In *International Conference of Design, User Experience, and Usability*. Springer, 431–439.
- [45] FR Kluckhohn and Fred L Strodtbeck. [n. d.]. Variations in Value Orientation. 1961. *Evanston, IL: Row, Peterson and Co* (n. d.).
- [46] Klaus Krippendorff. 2004. *Content Analysis: An Introduction to Its Methodology* (second ed.). Sage.
- [47] Florian Lachner, Florian Fincke, and Andreas Butz. 2017. UX Metrics: Deriving Country-Specific Usage Patterns of a Website Plug-In from Web Analytics. In *IFIP Conference on Human-Computer Interaction*. Springer, 142–159.
- [48] Florian Lachner, Philipp Naegelein, Robert Kowalski, Martin Spann, and Andreas Butz. 2016. Quantified UX: Towards a Common Organizational Understanding of User Experience. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI '16)*. ACM, Article 56, 10 pages.
- [49] Ilda Ladeira and Gary Marsden. 2014. Interactive personal storytelling: an ethnographic study and simulation of apartheid-era narratives. In *Proceedings of the 2014 conference on Designing interactive systems*. ACM, 249–258.
- [50] Effie Lai-Chong Law, Paul van Schaik, and Virpi Roto. 2014. Attitudes towards user experience (UX) measurement. *International Journal of Human-Computer Studies* 72, 6 (2014), 526–541.
- [51] Richard D Lewis. 2006. *When Cultures Collide: Leading Across Cultures*. Nicholas Brealey International.
- [52] Sanna Malinen and Piia Nurkka. 2013. The role of community in exercise: Cross-cultural study of online exercise diary users. In *Proceedings of the 6th International Conference on Communities and Technologies*. ACM, 55–63.
- [53] Aaron Marcus. 2005. Chapter 3: User Interface Design and Culture. In *Usability and internationalization of information technology*, Nuray Aykin (Ed.). CRC Press, 51–78.
- [54] Aaron Marcus and Emilie West Gould. 2000. Crosscurrents: cultural dimensions and global Web user-interface design. *interactions* 7, 4 (2000), 32–46.
- [55] Aaron Marcus and Sundus Hamoodi. 2009. The impact of culture on the design of Arabic websites. In *International Conference on Internationalization, Design and Global Development*. Springer, 386–394.
- [56] Philipp Mayring and Thomas Fenzl. 2014. *Qualitative Content Analysis*. Springer.
- [57] Brendan McSweeney. 2002. Hofstede's model of national cultural differences and their consequences: A triumph of faith—a failure of analysis. *Human relations* 55, 1 (2002), 89–118.
- [58] Indrani Medhi, Aishwarya Ratan, and Kentaro Toyama. 2009. Mobile-banking adoption and usage by low-literate, low-income users in the developing world. In

- International Conference on Internationalization, Design and Global Development*. Springer, 485–494.
- [59] Shafie Mehad, Wan Abdul Rahim Wan Mohd Isa, Nor Laila Md Noor, and Mohd Shukri Husin. 2010. Muslim user interface evaluation framework (Muslim-UI) for Islamic genre website: A quantitative approach. In *Information and Communication Technology for the Muslim World (ICTAM), 2010 International Conference on*. IEEE, H–1.
- [60] Houda El mimouni and Craig M MacDonald. 2015. Culture and information architecture: A study of American and Arab academic websites. *Proceedings of the Association for Information Science and Technology* 52, 1 (2015), 1–4.
- [61] Jakob Nielsen and Rolf Molich. 1990. Heuristic evaluation of user interfaces. In *Proc. of the SIGCHI conference on Human factors in computing systems (CHI)*. ACM, 249–256.
- [62] J Nielson, Elisa M. Galdo, Robert C. Sprung, and Piyawadee Sukaviriya. 1990. Designing for International Use. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI) (1990)*, 291–294.
- [63] Manuel Nordhoff, Tal August, Nigini A Oliveira, and Katharina Reinecke. 2018. A Case for Design Localization: Diversity of Website Aesthetics in 44 Countries. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 337.
- [64] Don Norman. 1988. *The Psychology of Everyday Things*. Basic Books, New York.
- [65] ChukwuNonso H Nwokoye, Virginia E Ejiolor, Rita Orji, and Njideka N Mbeledogu. 2016. The Topicality of Agent-Based Modeling/Multi Agent Systems in Human Computer Interaction Research: An African Perspective. In *Proceedings of the First African Conference on Human Computer Interaction*. ACM, 80–91.
- [66] Marianna Obrist, Virpi Roto, and Kaisa Väänänen-Vainio-Mattila. 2009. User experience evaluation: do you know which method to use?. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2763–2766.
- [67] Nigini Oliveira, Nazareno Andrade, and Katharina Reinecke. 2016. Participation differences in Q&A sites across countries: opportunities for cultural adaptation. In *Proc. of the 9th Nordic Conference on Human-Computer Interaction*. ACM, 6.
- [68] Cecilia Oyugi, Lynne Duncley, and Andy Smith. 2008. Evaluation methods and cultural differences: studies across three continents. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*. ACM, 318–325.
- [69] Ingrid Pettersson, Florian Lachner, Anna-Katharina Frison, Andreas Riemer, and Andreas Butz. 2018. A Bermuda Triangle?. In *Proc. of CHI 2018*. ACM.
- [70] Ingrid Pettersson, Florian Lachner, Anna-Katharina Frison, Andreas Riemer, and Andreas Butz. 2018. A Bermuda Triangle?: A Review of Method Application and Triangulation in User Experience Evaluation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM.
- [71] Kavita Philip, Lilly Irani, and Paul Dourish. 2012. Postcolonial computing: A tactical survey. *Science, Technology, & Human Values* 37, 1 (2012), 3–29.
- [72] B Joseph Pine, James H Gilmore, et al. 1998. Welcome to the experience economy. *Harvard Business Review* 76 (1998), 97–105.
- [73] Marco Pretorius, Jason Hobbs, and Terence Fenn. 2015. The user experience landscape of South Africa. In *Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists*. ACM, 32.
- [74] Dorina Rajanen, Torkil Clemmensen, Netta Iivari, Yavuz Inal, Kerem Rizvanoğlu, Ashok Sivaji, and Amélie Roche. 2017. UX Professionals' Definitions of Usability and UX—A Comparison Between Turkey, Finland, Denmark, France and Malaysia. In *IFIP Conference on Human-Computer Interaction*. Springer, 218–239.
- [75] Katharina Reinecke and Abraham Bernstein. 2011. Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)* 18, 2 (2011), 8.
- [76] Katharina Reinecke and Abraham Bernstein. 2013. Knowing What a User Likes: A Design Science Approach to Interfaces that Automatically Adapt to Culture. *MIS Quarterly* 37, 2 (2013).
- [77] Katharina Reinecke and Krzysztof Z Gajos. 2014. Quantifying visual preferences around the world. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 11–20.
- [78] Katharina Reinecke, Minh Khoa Nguyen, Abraham Bernstein, Michael Näf, and Krzysztof Z. Gajos. 2013. Doodle Around the World: Online Scheduling Behavior Reflects Cultural Differences in Time Perception and Group Decision-making. In *Proc. CSCW 2013 (CSCW '13)*. ACM, New York, NY, USA, 45–54. <https://doi.org/10.1145/2441776.2441784>
- [79] Virpi Roto, Effie Law, Arnold P.O.S. Vermeeren, and Jettie Hoonhout. 2011. User experience white paper. Bringing clarity to the concept of user experience. Result from Dagstuhl Seminar on Demarcating User Experience. *Disponible en ligne le* 22 (2011), 06–15.
- [80] Virpi Roto, Kaisa Väänänen-Vainio-Mattila, Effie Law, and Arnold Vermeeren. 2009. User experience evaluation methods in product development (UXEM'09). In *IFIP Conf. on Human-Computer Interaction*. Springer, 981–982.
- [81] Lori L Sachau and Susan R Hutchinson. 2012. Trends in culturally relevant interface design features for Latino Web site users. *Educational Technology Research and Development* 60, 6 (2012), 1033–1050.
- [82] Hanna Schneider, Florian Lachner, Malin Eiband, Ceenu George, Purvish Shah, Chinmay Parab, Anjali Kukreja, Heinrich Hussmann, and Andreas Butz. 2018. Privacy and personalization: the story of a cross-cultural field study. *interactions* 25, 3 (2018), 52–55.
- [83] Shalom H Schwartz. 1999. A theory of cultural values and some implications for work. *Applied psychology* 48, 1 (1999), 23–47.
- [84] Kennon M Sheldon, Andrew J Elliot, Youngmee Kim, and Tim Kasser. 2001. What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of personality and social psychology* 80, 2 (2001), 325.
- [85] Paola Signorini, Rolf Wiesemes, and Roger Murphy. 2009. Developing alternative frameworks for exploring intercultural learning: a critique of Hofstede's cultural difference model. *Teaching in Higher Education* 14, 3 (2009), 253–264.
- [86] Nitish Singh, Hongxin Zhao, and Xiaorui Hu. 2003. Cultural adaptation on the web: A study of American companies' domestic and Chinese websites. *Journal of Global Information Management (JGIM)* 11, 3 (2003), 63–80.
- [87] Nitish Singh, Hongxin Zhao, and Xiaorui Hu. 2005. Analyzing the cultural content of web sites: A cross-national comparison of China, India, Japan, and US. *International Marketing Review* 22, 2 (2005), 129–146.
- [88] Victoria Schwanda Sosik and Rajiv Arjan. 2016. Harnessing the Power of Digital Mobile Maps in Africa. In *Proceedings of the First African Conference on Human Computer Interaction (AfriCHI'16)*. ACM, 271–275.
- [89] Rein Suadamara, Stefan Werner, and Axel Hunger. 2010. Cultural influence on user preference on groupware application for intercultural collaboration. In *Proceedings of the 3rd international conference on Intercultural collaboration*. ACM, 215–218.
- [90] Fons Trompenaars and Charles Hampden-Turner. 2011. *Riding the waves of culture: Understanding diversity in global business*. Nicholas Brealey Publishing.
- [91] Rosalie L Tung and Alain Verbeke. 2010. Beyond Hofstede and GLOBE: Improving the quality of cross-cultural research.
- [92] Jari Varsaluoma, Heli Väättäjä, and Tanja Walsh. 2016. Exploring motivational aspects and user experience of mobile mathematics learning service in south africa. In *Proceedings of the 20th International Academic Mindtrek Conference*. ACM, 159–168.
- [93] Arnold P.O.S. Vermeeren, Effie Lai-Chong Law, Virpi Roto, Marianna Obrist, Jettie Hoonhout, and Kaisa Väänänen-Vainio-Mattila. 2010. User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*. ACM, 521–530.
- [94] David A Victor. 1992. *International Business Communication*. (1992).
- [95] Tanja Walsh and Piia Nurkka. 2012. Approaches to cross-cultural design: two case studies with UX web-surveys. In *Proceedings of the 24th Australian Computer-Human Interaction Conference*. ACM, 633–642.
- [96] Tanja Walsh, Piia Nurkka, and Rod Walsh. 2010. Cultural differences in smartphone user experience evaluation. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*. ACM, 24.
- [97] Tanja Walsh and Teija Vainio. 2011. Cross-cultural design for mhealth applications. In *Extended Abstracts of the Third International Workshop on Smart Healthcare Applications*. 1.
- [98] Heike Winschiers. 2006. The challenges of participatory design in an intercultural context: designing for usability in namibia. In *PDC*. 73–76.
- [99] Susan Wyche, Tawanna R Dillahunt, Nightingale Simiyu, and Sharon Alaka. 2015. If god gives me the chance I will design my own phone: Exploring Mobile Phone Repair and Postcolonial Approaches to Design in Rural Kenya. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 463–473.
- [100] Judith Yaaqoubi and Katharina Reinecke. 2018. The Use and Usefulness of Cultural Dimensions in Product Development. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, CS18.
- [101] Qian Yang, Nikola Banovic, and John Zimmerman. 2018. Mapping Machine Learning Advances from HCI Research to Reveal Starting Places for Design Innovation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 130, 11 pages.
- [102] Alvin Yeo. 1996. Cultural user interfaces: a silver lining in cultural diversity. *ACM SIGCHI Bulletin* 28, 3 (1996), 4–7.
- [103] Alvin W Yeo. 2001. Global-software development lifecycle: An exploratory study. In *Proc. of the SIGCHI 2011*. ACM, 104–111.

Eidesstattliche Versicherung

(Siehe Promotionsordnung vom 12.07.11, § 8, Abs. 2 Pkt. 5)

Hiermit erkläre ich an Eidesstatt, dass die Dissertation von mir selbstständig und ohne unerlaubte Beihilfe angefertigt wurde.

München, den 31.10.2018

Florian Lachner