PODER EXECUTIVO
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

REDESIGNING FOR EXPERIENCE - REX

An Approach for the Evaluation of User Experience and Suggestion of Improvements in Mobile Applications

Luis Jorge Enrique Rivero Cabrejos

Doctoral dissertation presented to the Informatics Post-Graduate Program (*Programa de Pós-graduação em Informática*) - PPGI, at Universidade Federal do Amazonas, as one of the requirements for achieving the PhD in Informatics degree.

Advisor: Ph.D. Tayana Uchôa Conte

Manaus
June, 2017

REDESIGNING FOR EXPERIENCE - REX

An Approach for the Evaluation of User Experience and Suggestion of Improvements in Mobile Applications

Luis Jorge Enrique Rivero Cabrejos

Doctoral dissertation presented to the Informatics Post-Graduate Program (Programa de Pós-graduação em Informática) - PPGI, at Universidade Federal do Amazonas, as one of the requirements for achieving the PhD in Informatics degree.

Dissertation evaluation on June 8th, 2017 in Manaus (Brazil), by the evaluation committee constituted by the following professors:

_____

Ph.D. Tayana Uchôa Conte

_____

Ph.D. Guilherme Horta Travassos

_____

Ph.D. José Luis de Souza Pio

_____

Ph.D. Raimundo da Silva Barreto

_____

Ph.D. Bruno Freitas Gadelha

Manaus

June, 2017

Acknowledgments

To my parents, Gina and Luis, for supporting me since the beginning and pushing me to pursue my dreams. You have always encouraged me to work harder every day, and were always happy to see me travel to show others about my research. I hope that we get to see each other more and more, and to share with you all the new adventures I am about to live. Thank you for being part of this moment.

To Davi Viana, for his constant support, friendship and the time we spent together. These past years we have grown closer and closer, and we have faced so many challenges. I hope that many years come for us to continue doing as we have. Thank you for taking care of me and being my family whenever I needed it. Because of you and what you've shown me, I feel part of your beautiful country.

To my dearest friends, Nathy, Aurea and Lipe, you've been so kind to me at this new job, and every second I spend with you has made me feel even more certain that I am doing what I want to do in life. I am grateful to have such wonderful people to call colleagues and friends. Thank you for the laughter at the end of each hard work day and at our "job meetings". You are amazing!

To my friends, Priscila Fernandes, Bruno Bonifácio, Marcos Gomes and Jacilane Rabelo. You have given me such nice memories over these years, that it is hard to pick one in particular to remember you by. Thank you for being there to discuss research, life decisions, or just stay next to each other to tell silly jokes. You have made my day so many times, that there is nothing else I could ask from you, but to keep having such type of moments over and over again. Thank you all.

To my friends Claudia and Veronica, whom I grew up with. Our friendship is certainly strong. Not even travelling to other countries has managed to stop us from calling each other or marking dates to hang out whenever we were in town. I am truly blessed to still have you in my life.

To my friends from the USES group. Thank you all for the constant support, your critics, "banquets" each USES Day, and all the jokes. I've certainly become a better researcher and person because of your support. I hope to see you all soon whenever life takes us, sharing papers and experiences.

To Guilherme Travassos, José Pio, Raimundo Barreto and Bruno Gadelha for agreeing to participate in my thesis evaluation committee. Certainly, your critics and

suggestions have helped and will help me improve my research. Thank you for taking the time to read my work and for your outstanding reviews.

To all the subjects who participated in the empirical studies. Specially, to Gabriel and Anderson, for your support when evaluating the SIGNAL application.

To the Universidade Federal do Amazonas and the Instituto de Computação, for their support during my research. Specially, Elienai, Helen, Frank and Marcio. You've made each trip and training easier with your support. Thank you very much.

To CNPq, FAPEAM and CAPES for their financial support in each accepted research project, and for my scholarship.

Finally, I have a special acknowledgement to my advisor, Tayana Conte. Dear Tayana, there are no words to describe how I feel about this moment and how I think it is only possible because of you. When I first met you, I didn't know how to write a paper in Portuguese or English. I didn't know how to give a lecture, or how to carry out a project. It took many hours of dedications, but now, I am teaching classes and writing papers, and having students who want to do research with me. I hope that they are seeing in me, what I only think is a reflection of you. Because of you, I have masters and soon a doctorate in which I learned how to carry out research. However, I've learned so much more, like: how to be responsible, ethical, critical, disciplined, creative, and so much more because of you. Surely, you didn't have to teach me those things, but you have in one way or another. All I can say is how grateful I am, and that I hope to see you in the near future as my always mentor and friend.

Abstract of the doctoral dissertation presented to PPGI/UFAM as a partial fulfillment of the requirements for the degree of Doctor in Informatics.

REDESIGNING FOR EXPERIENCE - REX

An Approach for the Evaluation of User Experience and Suggestion of Improvements in Mobile Applications

Luis Jorge Enrique Rivero Cabrejos

June/2017

Advisor: Ph.D. Tayana Uchôa Conte

Program: Informatics Post-Graduate Program (*Programa de Pós Graduação em Informática*)

User eXperience (UX) refers to a holistic perspective and an enrichment of traditional quality models with non-utilitarian concepts, such as fun, joy, pleasure, hedonic value or ludic value. In order to evaluate UX in software applications, several technologies (tools, methods, techniques) have been proposed that range from using questionnaires to employing biometrics to gather quantitative and qualitative data on users' experience. However, there is a need for research in the development of specific UX evaluation technologies that are easy and comfortable to use from the point of view of users, while supporting software engineers in the correction of the aspects that cause poor experiences. Additionally, new UX approaches should be proposed for evaluating mobile applications, as there is still a shortage of methods for this type of applications, which is rising in demand. This doctoral dissertation proposes an alternative approach for evaluating mobile applications called Redesigning for EXperience (REX), which intends to be less intrusive for users when extracting UX data, while generating reports containing design suggestions for improving the UX. We assessed the acceptance of the initial versions of the REX approach from the point of view of users and software engineers in two studies. When compared to 3E, a qualitative UX evaluation method, the results showed that REX was perceived as more fun, useful and more interactive. Additionally, software engineers considered REX useful and easy to understand, while suggesting improving its report to facilitate its understanding and increase its use. After working on the improvements opportunities from the empirical studies, we developed a tool support for the REX approach called the REX report generator. Also, we carried out an observational study to verify to which extent the REX approach could be applied in a real software development project. Thus, REX was employed by users to evaluate a mobile educational application and a discussion meeting was held with the software development team to discuss the improvement suggestions provided by REX to support

the redesign process. The findings from the observational study indicated the satisfaction of users to report their experience with the REX approach, while the members of the development team agreed with the usefulness of the REX report and its improvement suggestions. By providing design suggestions, we aim to support software engineers in improving the UX of the developed mobile applications, thus increasing their quality and acceptance in the market.

Keywords: User eXperience, Software Quality, Literature Review, Empirical Study, Evaluation Approach, User Interface, Redesign Suggestions, Mobile Applications, UX Evaluation Tools.

Table of Contents

Figures Index

Tables Index

1 Introduction

This chapter presents an introduction to this research. Besides providing a contextualization, we present the motivation, goals and the applied research methodology. Finally, we present the structure of this document regarding the development of proposals for the evaluation and suggestion of improvements in mobile applications aiming at improving their UX.

1.1 Context

As the number of mobile devices has increased, software development teams have focused on releasing mobile applications, allowing users to carry out business transactions, access information and improve their lifestyle more efficiently (Sarwar and Soomro, 2013). Nevertheless, even when providing useful means for carrying out daily tasks, users report dissatisfaction or frustration when using these applications (Hu et al., 2006). In this sense, in order to improve the adoption and acceptance of mobile applications, software developers need to improve their ease of use and the emotions that they convey (Ervasti et al., 2011).

Usability is one of the main attributes that represent software quality. According to the ISO 25010 (2011), usability is "*the capability of the software product to be understood, learned, operated, attractive to the user, and compliant to standards/guidelines, when used under specific conditions*". In that context, usability subsumes the aspects of how easy a software application can be used, such as learnability, operability, aesthetics, and others.

Usability is fundamental for achieving software acceptability. According to Mendes et al. (2006), if an application is not usable, it will be soon replaced by a more usable one or will not be used at all. Moreover, the results from an empirical study regarding the importance of software attributes pointed out that usability was considered a key attribute for the development of high-quality software (de la Vara et al., 2011).

Despite the increasing attention that usability has achieved in the software engineering community (Fernandez et al., 2011), a new term, "User eXperience" (UX), has emerged as a new phrase for new ways of understanding and studying the quality in use of interactive products (Bargas-Avila and Hornbæk, 2011). UX is defined as the "*person's perceptions and responses that result from the use and/or anticipated use of a product, system or service*" (ISO 9241, 2010).

The key focus of the UX movement is on the experience of the interaction with a product. Hassenzahl and Tractinsky (2006) state that UX is a consequence of a user's internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g., complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g., organizational/social setting, meaningfulness of the activity, voluntariness of use and others). Thus, UX research represents "*a turn to experience*" searching for new approaches for the design of interactive products, which accommodate experiential qualities of technology use rather than product qualities (Hassenzahl et al., 2010).

Vermeeren et al. (2010) point out that the relationship between usability and UX is intertwined. While usability focuses on task performance (e.g., measuring task execution time, the number of clicks or errors), UX focuses on lived experiences, analyzing peoples' emotions while interacting with software products. In that context, usability is related to UX. Therefore, as UX is subjective, objective usability measures are not sufficient for measuring UX. It is necessary to analyze how the user feels about the software application while performing tasks on it.

1.2 Defining and Evaluating User eXperience

New research has identified that functionality and usability are just not enough for customers who want products that "*dazzle their senses, touch their hearts and stimulate their minds*" (Hassenzahl, 2005). According to Bargas-Avila and Hornbæk (2011), existing usability research is too focused on task efficiency and work, and therefore, there is a need for more encompassing notions of quality. In that context, User eXperience (UX) research has emerged, also focusing on hedonic qualities of product use (Bargas-Avila and Hornbæk, 2011).

Several definitions have been given to the term User eXperience. For instance, Isomursu et al. (2004) state that UX is the totality of the subjective experience of using an application in a situation. Authors who agree with this definition state that as UX is formed in a dynamic relationship between the user and the device, the application and the usage environment, it cannot be evaluated in a vacuum. Additionally, UX is also considered going beyond the task-oriented approach of traditional Human Computer Interaction (HCI) by bringing out aspects such as beauty, fun, pleasure, and personal growth that satisfy general human needs but have little instrumental value (Kujala et al., 2011). Similar definitions are provided by other authors (Hassenzahl, 2005; Väätäjä et al., 2009; Vermereen et al., 2010; Gegner and Runonen, 2012). In this sense, the key focus of UX is considering attributes beyond the instrumental such as users' emotions before, during and after their interaction with a software application.

Hassenzahl (2005) states that when designing a product, a designer chooses and combines certain features (e.g., content, presentational style, functionality, interactional style) to convey an intended product character. When a user comes in contact with a product, first (s)he perceives the product's features. Then, (s)he constructs a personal version of the product character, leading to consequences such as: (a) the judgment about the product's appeal (e.g., "It is good/bad"), (b) emotional consequences (e.g., pleasure, satisfaction) and (c) behavioral consequences (e.g., increased time spend with the product). The consequences of the character of a particular product are not always the same, as they are moderated by the specific usage situation.

The attributes that a product can possess can be distinguished based on the functions of the product and can be categorized into pragmatic and hedonic. Pragmatic attributes are related to the manipulation of the environment and focus on providing functionality (i.e., utility) and ways to access that functionality (i.e., usability). In this sense, a software application with pragmatic attributes will provide the necessary functionalities to carry out a task, and employing it should also be intuitive in a way that a user can figure out how to use it. Hedonic attributes, on the other hand, emphasize the individuals' psychological well-being and are related to pleasure. Hedonic quality focuses on aspects of stimulation, identification, and evocation. Stimulation is related to personal development, identification addresses the expression of the self and the user's values to others, and evocation refers to the product's ability to provoke memories such as important past events or relationships. Consequently, software applications with

hedonic attributes will be perceived as “outstanding”, “impressive”, “exciting” and “interesting” (Hassenzahl, 2005).

It is necessary to evaluate user experience (considering both hedonic and pragmatic attributes) within the development process and during their use in order to obtain a complete understanding of the users’ needs, thus providing a better user experience (Bevan, 2008). In order to assist in the observation of UX, Figure 1-1 presents some factors that are related to the observation of UX. In this sense, pragmatic attributes can be observed through traditional usability (e.g., task execution time, the number of clicks or errors) (Vermeeren et al., 2010). However, hedonic attributes dealing with emotion require further investigation. Desmet (2005) indicates that emotions should be treated as multifaceted phenomena consisting of the following components: behavioral reactions (e.g., retreating), expressive reactions (e.g., smiling), physiological reactions (e.g., heart pounding), and subjective feelings (e.g., feeling amused). To gather information on the users’ emotions, two types of instruments have been proposed (Desmet, 2005): non-verbal (objective) instruments and verbal (subjective) instruments.

<table>
<tr><td>UX pragmatic do goals</td><td colspan="3">To be effective and efficient</td></tr>
<tr><td>UX hedonic be goals</td><td colspan="3">Stimulation, identification and evocation</td></tr>
<tr><td>UX: actual experience</td><td>Visceral</td><td colspan="2">Experience of interaction</td></tr>
<tr><td rowspan="2">Measures of UX consequences</td><td colspan="3">Satisfaction in use:
satisfaction with achieving pragmatic and hedonic goals</td></tr>
<tr><td>Pleasure</td><td>Likability and Comfort</td><td>Trust</td></tr>
</table>

Figure 1-1 Factors related to UX suggested by Bevan (2008)

Non-verbal instruments comprise mechanisms that allow observing either the expressive or the physiological component of emotion. An expressive reaction (e.g., smiling or frowning) is the facial, vocal, and postural expression that accompanies the emotion (Desmet, 2005). On the other hand, verbal self-report instruments typically assess the subjective feeling component of emotions. A subjective feeling (e.g., feeling happy or feeling inspired) is the conscious awareness of the emotional state one is in. Subjective feelings can only be observed through self-report. The most often used self-report instruments require respondents to report their emotions with the use of a set of rating scales or verbal protocols (Desmet, 2005).

When applying one of the instruments above, the observed period of experience details when the UX observation can be carried out in product usage (Kujala et al., 2011). If a UX instrument can be applied before using an application, we measure the expectations of the user regarding the product concept. On the other hand, UX observations performed during product usage (in momentary episodes) are more complex, as the observer must gather data while the informant is carrying tasks or exploring the hedonic/pragmatic aspects of the software application. It means that UX instruments that allow gathering data during product usage are more intrusive and, thus, special care must be taken to ensure that observing UX does not affect the experience itself. Additionally, most UX assessment technologies allow gathering data after an episodic experience (Vermeeren et al, 2010). By asking questions or making a retrospection analysis, the observers can gather information on aspects that the users reflected on regarding their UX. Finally, assessing UX over time allows gathering UX data before, during and after the different cumulative episodic experiences of the user. Although this type of assessment is richer in information on the aspects that affect product attachment, its main disadvantage is that the users may not be available to carry out an evaluation for such a long period, or the development team may not have time to carry out such long-term UX studies. Furthermore, despite the moment in which the evaluation is carried out, some users may not remember specific aspects of their interaction/expectations, while others may have difficulty in expressing their experiences.

New UX evaluation methods have been proposed in order to assist software engineers in choosing an appropriate design for software applications, to ensure that the development is on the right track, or to assess if the final product meets quality standards regarding UX (Vermeeren et al., 2010). However, besides allowing the identification of UX problems, UX evaluation methods need to be usefully employed in product development and be easy to use by both practitioners and users at a lower cost (Bargas-Avila and Hornbæk, 2011). Additionally, it is still necessary to develop evaluation technologies that consider the specific characteristics of mobile applications, to achieve better results (Vermeeren et al., 2010).

Given the current needs regarding UX evaluation methods, we are interested in developing technologies (methods, tools, processes, others) that can be applied to ensure the quality of UX of mobile applications in the development process guided by software

engineering standards. In the following subsection, we describe the problem to be solved in this research and the goals to be achieved.

1.3 Problem and Goals

According to Juristo et al. (2007), some of the benefits of including usability and UX principals within the software development process guided by Software Engineering standards, are: (a) improving the ease of use of the interfaces, (b) increasing user productivity, and (c) reducing time and costs in the development, user training, software maintenance and software documentation. In that context, this research is motivated by the need of technologies that can be employed by software engineers that are willing to improve the quality of their developed mobile applications regarding UX.

In software engineering, when using an iterative development approach, we set aside time to improve what we have developed (Cockburn, 2008). In that context, instead of integrating the software deliverables at the end of the cycle, we examine it from various standpoints (e.g., Was it the right thing to develop?; Do the users like the way it works?; Does it work fast enough?). To this end, there are two particular, specialized rework strategies:

- To develop the software application as well as possible in the thinking that if it is done sufficiently well, the changes will be relatively minor and can be incorporated quickly;
- To develop the least amount possible before sending out for evaluation, in the thinking that less work will be wasted when new information arrives.

In both strategies, the purpose is to examine the software from various standpoints in order to improve product quality, verifying if we are developing the right thing, if users like the way it works, and if it meets other necessary requirements. According to Travassos et al. (1999), by applying evaluation techniques early in development to identify problems or improvement opportunities; we can improve software quality while reducing costs regarding rework, as problems will be found before releasing the product into the market. Analogously, by applying evaluation and redesigning applications in each iteration, we can improve their quality as well by making changes that improve their UX. The second strategy is mostly applied in mobile applications, where the development of new Application Programming Interfaces

(APIs) allow software development teams to accommodate new feature requests, fix bugs, meet new standards and provide higher performance (McDonnell et al., 2013). Therefore, this research aims at employing iterative strategies to evaluate and suggest improvements in mobile applications to improve their quality regarding UX. In this sense, our research question is:

"*How can we support software engineers in the evaluation and redesign of mobile applications providing a positive UX*?"

We are interested in providing solutions applicable to software engineers in the context of mobile applications. The focus of mobile applications was selected due to the nature of these applications, which follow an iterative development lifecycle, developing features and improving them as the software is released in the market (Khalid et al., 2014). Also, researches showed that as the number of devices has increased (IDC, 2015; GWI, 2014), it is necessary to design proper applications that meet the users' expectations (Khalid et al., 2014). As a result, besides allowing software engineers to identify UX problems, we are interested in assisting these practitioners in correcting them. This will be done by employing an iterative strategy, aiming at facilitating the evaluation process, both for software engineers and end users.

By creating techniques and a process to guide software engineers during the evaluation and redesign of mobile applications, we intend to: (a) facilitate the evaluation process, by providing support; (b) reduce the time spent during the evaluation, since they will be able to know what data to gather; (c) increase the number and relevance of the identified problems aiming at identifying problems having a negative impact on the user experience; and (d) facilitate the correction of the identified problems by indicating suggestions of what could be improved. Therefore, the main goal of this research is to create, empirically evaluate and evolve a set of technologies for the UX evaluation and suggestion of improvements of mobile applications. These technologies will guide software engineers during the problem identification and correction processes. To achieve that goal, we have decomposed it in the following specific goals:

- To identify improvement opportunities in UX evaluation methods and design proposals and include them in the development of our approach.

- To create a set of technologies for the evaluation and suggestion of improvements of mobile applications aiming at achieving a positive UX.
- To evaluate the performance of our proposal regarding time spent, identified UX problems, false positives, the suggestion of improvements, to identify improvement opportunities in the developed technologies and facilitate their use.
- To present empirical evidence of the performance of the proposed technologies in the context of software development projects, aiming at verifying their feasibility in industrial settings.

Considering these goals, we proposed an alternative approach (i.e., a technique and tool) for evaluating mobile applications called Redesigning for EXperience (REX), which intends to be less intrusive for users when extracting UX data, while generating reports containing design suggestions for improving the UX. We assessed the acceptance of the initial version of the REX approach from the point of view of users and software engineers in two studies. When compared to 3E, a qualitative UX evaluation method, the results showed that REX was perceived as more fun, useful and more interactive. Additionally, software engineers considered REX useful and easy to understand, while suggesting improving its report to facilitate its understanding and increase its use.

Following the suggestions from users and software engineers from the first empirical study, we developed a tool support for viewing the UX evaluation results and improvement suggestions called the REX report generator. Also, we carried out an observational study to verify to which extent the REX approach could be applied in a real software development project. The findings from the study indicated the satisfaction of users to report their experience with the REX approach, while the members of the development team agreed with the usefulness of the REX report and its improvement suggestions. By providing design suggestions, we aim to support software engineers in improving the UX of the developed mobile applications, thus increasing their quality and acceptance in the market. In the following subsection, we explain our methodology for the development, evaluation, and improvement of the REX approach.

1.4 Methodology

In order to achieve the goals of this research, we applied a methodology based on evidence, which uses empirical studies to evaluate and evolve software technologies.

Empirical studies allow researchers to answer questions that examine or identify problems in new proposals (Shull et al., 2001). Furthermore, our methodology also adopts secondary studies which identify, interpret and evaluate a determined research topic (Mafra et al., 2006), so that these results can be used in the proposal of the new technologies.

According to Shull et al. (2001), an empirical methodology is an approach that allows evolving a technology from its definition to its transfer to the software industry by carrying out empirical studies. Table 1-1 shows what type of studies can be performed, their purpose and the research question that they aim at answering. The order in which the questions are applied intends to avoid rework by: (a) identifying basic and fundamental problems with the technology that can generate drastic changes at the beginning of its definition; and (b) identifying specific problems with the technology that require smaller changes to be corrected.

Table 1-1 Questions to be answered on the empirical methodology by Shull et al. (2001)

Type of Study	Research Question	Purpose
Feasibility Study	Are the results of the process useful and is time well spent?	To determine if it is feasible to use the technology.
Observational Study	Do the steps of the process make sense?	To observe the technology in order to improve the researchers' understanding of its application and to allow its improvement.
Case Study (in the lifecycle)	Is the process adequate to the real lifecycle?	To analyze the technology in a specific context of the lifecycle and to characterize its application.
Case Study (in the industry)	Is the process adequate to the industrial environment?	To identify if there are problems in the integration of the proposed technology in an industrial environment.

Mafra et al. (2006) indicate that besides carrying out empirical studies, it is necessary to carry out secondary studies. While empirical studies allow testing hypotheses, secondary studies allow collecting relevant data on a research topic.

Therefore, Mafra et al. (2006) extended the methodology proposed by Shull et al. (2001) dividing it into two main stages: (a) the first stage, in which secondary studies would be performed; and (b) the second stage, which corresponds to Shull's et al. (2001) original methodology.

Figure 1-2 shows our research methodology. This methodology adapts the methodology by Mafra et al. (2006) considering secondary studies to gather information on needs of the software industry regarding evaluation and design technologies for meeting UX principals. Also, it applies empirical studies to evaluate our proposal in the evaluation and suggestion of improvements for mobile applications. We will now describe the activities from each stage of the methodology:

- **Literature Review:** We carried out secondary studies to gather information on evaluation and (re)design technologies in order to identify improvement opportunities.
- **Initial Evaluation of the Existing Methods:** During this stage, we carried out initial studies to evaluate existing proposals for the evaluation and (re)design of software applications aiming at achieving a high-quality UX. During this evaluation, we have selected a subset of the identified methods from the Literature Review stage in order to identify their strengths and gather a set of features to be incorporated in our proposal.
- **Development of Technologies for UX Evaluation and Suggestion of Improvements:** During this stage, we have incorporated the identified features from both the Literature Review and Initial Evaluation Stages to propose a set of technologies aiming at evaluating and suggesting improvements for mobile applications. The goal of this stage is to condense the gathered data to make a proposal that meets the current needs in the UX research area, assisting software engineers in the evaluation process of mobile applications, while suggesting what could be changed to solve the identified problems.
- **Execution of Empirical Studies:** During this stage, we have performed feasibility studies in order to determine whether the results of applying our approach are feasible.

- **Refinement of the Technologies:** Based on the results from the empirical studies, we made modifications in the proposed technologies, aiming at improving their performance and facilitate their use.

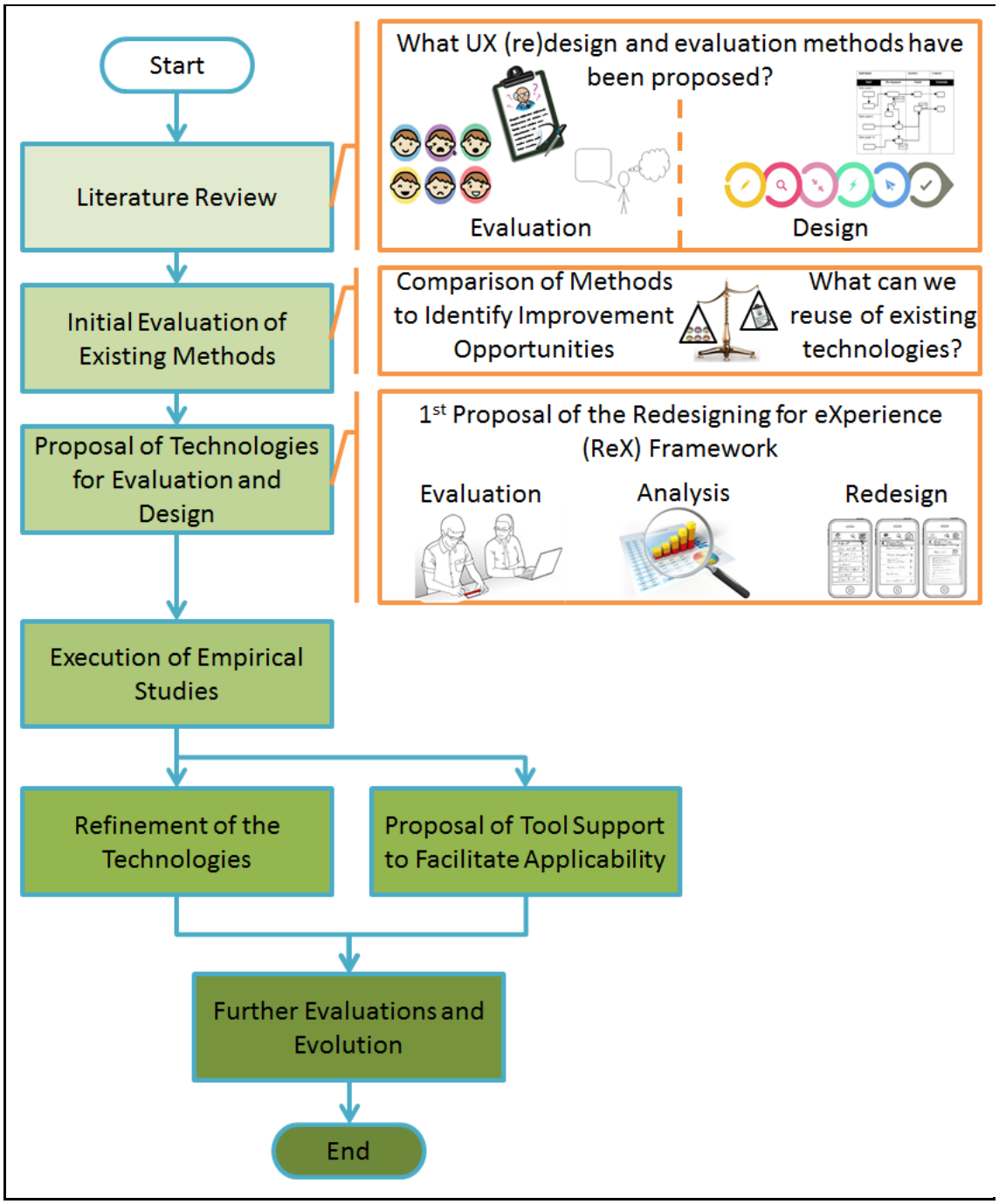


Figure 1-2 Applied research methodology for the creation, evaluation, and evolution of the proposed technologies

- **Development of Tool Support to Facilitate Applicability:** Also, simultaneously to the Refinement stage, we have gathered a set of requirements for suggesting supporting tools for facilitating the use of our approach. Such tools were proposed based on: (a) the difficulties encountered by software engineers when applying our approach when evaluating and redesigning mobile applications aiming at achieving high-quality UX; (b) requirements suggested within the literature review; and (c) results from our initial evaluation of UX evaluation and (re)design methods.
- **Further Evaluations and Evolution:** Finally, during this stage, we have performed other types of empirical studies, such as observational studies to gather further data on how to apply our proposal in order to achieve better performance or meet the needs of the software industry. These results have also been considered in the refinement of our approach, aiming at facilitating its use by software engineers in an industrial environment.

1.5 Organization

This doctoral dissertation presents our findings on methods for evaluating and (re)designing software applications for a positive UX, and how we used these and the results from initial empirical evaluations to propose the Redesigning for EXperience (REX) approach. Additionally, we present the empirical evaluations of REX and its refinements during the execution of this research. The remainder of this document is organized as follows:

Chapter 2 – Related Work discusses the results from literature reviews aimed at identifying needs in new proposals of UX evaluation methods. Besides identifying these methods, this chapter discusses requirements that need to be met by future UX evaluation proposals. Also, this chapter presents an overview of UX design methods and other existing alternatives that can be applied for gaining insights into how to improve the UX of a software application.

Chapter 3 – Initial Studies of UX Evaluation and Design Technologies presents an empirical study of specifically selected methods for evaluating UX in interactive products. Besides explaining how we applied the selected methods, we compared their performance and cite relevant features and suggestions to be considered in future proposals for UX evaluation. Furthermore, we present the results of a

qualitative study in which novice software engineers applied design suggestions in order to correct pragmatic problems in mobile applications. The results also indicate what can be improved when applying design suggestions to correct UX problems.

Chapter 4 – The Proposed Technology shows how we analyzed the data collected from our previous literature reviews and empirical studies, to propose an approach for the evaluation and suggestion of improvements in mobile applications. This chapter presents the REX approach, its development, its tools support and an application example.

Chapter 5 – The Initial Evaluations and Improvements of REX shows the initial evaluations of the REX approach, evaluating the performance of our proposal and its acceptance. In this chapter, we present the results from two empirical studies, one evaluating REX from the point of view of users, and the other one evaluating REX from the point of view of software engineers. Also, we present the changes made in our proposal, presenting the initial version of the REX Report Generator tool, which was developed to facilitate viewing the results from a UX evaluation using REX.

Chapter 6 – An Observational Study of REX shows the performance of REX in an observational study, where it was employed to evaluate the UX of a mobile educational application under development. In this chapter, we also gathered information on the applicability of REX from the point of view of users and software engineers.

Chapter 7 – Conclusions and Future Work presents the conclusions and contributions from this research. Also, it discusses future work to be carried out in order to continue this research.

2 Related Work

This chapter presents concepts regarding user experience to better ground this research. Besides, defining attributes that should be considered when observing UX, this chapter discusses the results from literature reviews by other authors that were carried out to identify UX evaluation methods and improvement opportunities.

2.1 Introduction

Identifying and analyzing current research in the field of UX evaluation is crucial to understand what works or not in specific situations. Several researchers have studied the existing proposals in terms of UX evaluations methods to provide both industry and academy with an overview of these methods and the conditions in which they can be applied.

Jordan (2003) condensed research for designing pleasurable products, indicating the need of considering further features besides usability (e.g., pleasure). In his book, the author gathered not only UX evaluation methods for identifying UX problems but also methods for designing products meeting UX criteria. Although the summary on UX evaluations was short, the book provides readers with an awareness of the issues associated with creating pleasurable products.

To investigate the different available UX evaluation approaches, Vermeeren et al., 2010 carried out a review where 96 UX evaluation methods were identified. They characterized the methods according to their origin, type of collected data, evaluated the type of applications and others, and found out that there is a need for methods for the early phases of development and that special attention should be given to proposing UX methods that are practical to use. Bargas-Avila and Hornbæk (2011), on the other hand, focused on identifying papers where empirical studies were performed in the field of UX. The authors identified a total of 51 studies and indicated the type of employed methods within the studies and the UX dimensions that were assessed. Although their

review was not specifically performed for characterizing UX evaluation methods, the authors identified 15 papers describing UX evaluation methodologies.

Another generic review was performed by Law (2011). In their review paper, the author explored UX metrics and types of methods that could provide those metrics. Such review can provide a basis for understanding UX and future needs in the field of UX evaluation. Bargas-Avila and Hornbæk (2012) also provided a ranking of UX evaluation technologies and the aspects that these methods evaluated through the analysis of the data from their previously published review (Bargas-Avila and Hornbæk, 2011). Moreover, Rajeshkumar et al. (2013) carried out an analysis of UX evaluation methods identified in other reviews, considering the review by Vermeeren et al. (2010), providing an overview of how UX methods have been applied and a categorization to assist software engineers in their application.

Other reviews focused on a specific category of UX evaluation methods. For instance, the reviews by Frey et al. (2013) and Balters and Steinert (2015) focused on the evaluation of UX by means of physiology measurement. According to Frey et al. (2013), physiological sensors and neuroimaging allow exploring concepts such as workload, attention, vigilance, fatigue, error recognition, emotions, engagement, flow and immersion, which could assist in the evaluation of UX. In their work, the authors explain the physiology of the human body and describe methods that can be applied to gather UX data based on considering the human response to a stimulus. Furthermore, Balters and Steinert (2015) provide an overview of current studies using physiology sensors in engineering and human–computer interaction settings. Although details on how the review was performed were not provided, the authors provide a list of papers describing the physiological measurement of UX.

Although there are several reviews that provide indicators of current needs in the field of UX evaluation, we still need to understand how the identified methods work, and how they have been employed. Thus, it is necessary to dig deeper, finding further information on these methods and the situations in which they prove useful for assessing the UX and identifying problems. In the next subsections, we describe the results from a literature review, where we analyzed the methods cited by Vermereen et al. (2010) and carried out an extension of their review to identify methods that have been proposed since the publication of their paper. Also, following the identification of

improvement opportunities, this chapter discusses the use of interface design suggestions as a means for correcting UX problems, and which technologies in such context have been considered for the proposal of our approach.

2.2 UX Evaluation Methods

Based on the analysis of our systematic literature review, Figure 2-1 shows a categorization considering the way in which the UX data is gathered. Below, we describe each of the categories, and we will also cite examples of technologies for each category. Additionally, when citing a paper that was considered in our review, we use the letter S (Selected) followed by the paper's sequence number (i.e., SXX). Interested readers can find the complete list and categorization of papers reporting UX evaluation technologies in Appendixes B and C; and a thorough description of our review in a technical report (Rivero and Conte, 2016).

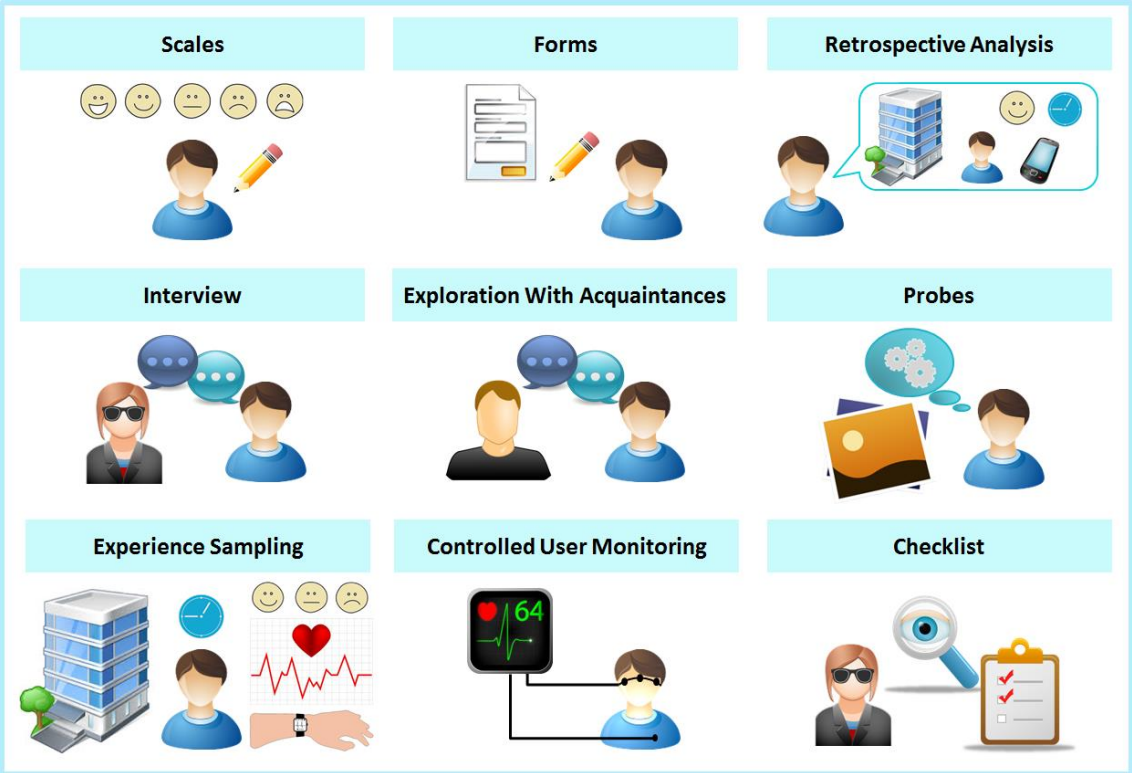


Figure 2-1 Types of identified UX evaluation methods

In essence, technologies such as scales, forms, interviews, and checklists do not provide innovative means for assessing UX. However, it is the new evaluated UX factors that make this type of technologies essential. For instance, some methods focus on the evaluation on emotions in the dimensions of valence (happiness-sadness), arousal

(aroused-sleepy) and dominance (controlled by or in control of the situation). Additionally, other approaches propose specific items for evaluating pragmatic quality (PQ), Hedonic Quality (HQ) and attractiveness (ATT). An example of such methods is the AttrakDiff instrument in the format of semantic differentials [S6]. It consists of 28 seven-step items whose poles are opposite adjectives (e.g., “confusing - clear,” “unusual - ordinary,” “good - bad”). Other forms such as the one provided in the Sentence Completion [S47] method can be employed to assess a variety of constructs, including motivations and attitudes. By providing only the beginning of a sentence, a researcher gives the topic, but respondents have the freedom to respond to it as they wish.

As scales and forms are quick and easy to use, if implemented online, they can allow gathering data from distributed users covering different user profiles and can enable gathering data both on positive and negative experiences. Currently available forms and scales allow gathering information on the degree of acceptance of software, the emotions that users are feeling when using it, and information on the degree of portrayed hedonic attributes, such as beauty, identification, evocation, and stimulation. Furthermore, UX checklists are being applied by software engineers and UX designers to fulfill specific UX attributes and check whether the application is following UX standards.

Some current UX evaluation methods, such as experience sampling, allow users to report their experience at specific moments of their day. While a device periodically senses their physiological responses to stimulus, users are asked how they are feeling, and the context they are in. The main advantage of employing these methods is having users experiencing the applications in real usage scenarios, gathering valuable information to improve UX. Authors employing these methods suggest enhancing these evaluation technologies to gather further information on the context of use of the application, without interrupting the user. This will provide designers with further information to address the issues for a poor UX and the contexts in which specific hedonic and pragmatic features are desired.

Another type of UX evaluation technology that allows gathering information on the users’ psychophysiological responses or their reaction towards a stimulus is controlled user monitoring. This type of technology, however, is applied in controlled environments, which can have an effect over how the user behaves. In that context, the

gathered information can be useful for identifying improvement opportunities or reactions towards the stimulus that the evaluator is examining, reducing the effect of undesired conditions. Furthermore, at the end of the evaluation, the evaluators can request further information from the users by applying other UX evaluation techniques, gathering further data to refine the evaluated software.

In the category of retrospective analysis techniques, users are asked to recall information. Then, several instruments can be applied to gather the recalled data. For instance, in the use of the iScale [S29], participants are asked to sketch how their opinion on a given product quality (e.g., the perceived usability of a product) has changed over a course of time. In essence, participants sketch linear segments with a certain slope (which denotes the change in the value of the reported quality) and length (which denotes the temporal span of the reported period). Participants may then associate each line segment with one or more experiences that are believed to have induced the given change in their perception of the product quality. Although it may be difficult for some users to remember specific events during their experience, a retrospection method allows users to report on the main events that affected their experience. Furthermore, these methods can be useful to confront users with previously gathered data (by means of applying experience sampling methods), which will allow software engineers to understand the key aspects that affected an experience.

Exploration with acquaintances is another way of letting users discuss the positive and negative aspects of their experience, however without the intervention of an evaluator. The main advantage of this approach is that users tend to comment based on the opinions of others and this method could encourage shy users to open to their friends and report on their experience. This type of method can also be useful for monitoring users' reactions towards an application when socializing. This information is relevant for gaining a holistic view of users' interaction with the evaluated applications and other users.

Last but not least, probes have emerged as an alternative for motivating and inspiring users in a user-centered design process. We can use materials such as multimedia and objects to engage users in the design process of an application. This approach is useful for finding means to understand customers and new ways for meeting

their needs. Having probes can allow us to communicate with users and let our creativity help them improve the quality of the developed application.

2.3 Improvement Opportunities in UX Evaluation Methods

When analyzing the results from our systematic mapping study (Rivero and Conte, 2016), we considered a total of 283 publications (See Appendixes A and B). As a result, we noticed that quantitative data is mostly analyzed when monitoring users' physiology or categorizing the overall opinion of users through scales. On the other hand, qualitative data is analyzed when digging deeper into the users' opinion, or when trying to understand the causes for a poor experience. Therefore, new UX evaluation technologies should take into account both quantitative and qualitative data; therefore, allowing the UX evaluation technologies to provide more consistent and complete evaluation reports on hedonic and pragmatic aspects.

Regarding the evaluated type of application, we identified that new types of applications are arising and more methods for applications of ubiquitous computing are gaining attention in the field of UX. We identified that there are methods that are broad enough to be applied in the evaluation of any type of interface, while others focus on Web applications and Mobile applications, yet these numbers are still low (we only identified 24 techniques – around 8% of the selected papers from our review). However, the "other types of applications" category had the highest number of papers (above 40%). The methods within these papers evaluated games, recommended systems, management tools, e-learning systems, virtual reality applications, museum guide applications and others. However, games are receiving more attention due to the importance of user experience in the engagement of users.

We also investigated how many of the current proposals provided support and what kind of support this was. In this sense, only around 6% of the papers described methods in which support was provided. From these methods, some recommend: (a) employing the identified violated UX principals and making changes in the application to reflect those rules [S68]; (b) asking the user what could be improved [S57], and (c) extracting improvement opportunities from the obtained data [S111]. Nevertheless, few methods suggest the specific steps for modifying the interface or the process for extracting the improvements opportunities. Additionally, as UX is personal, it may be

difficult for users to suggest solutions that may apply to most users without being previously tested. This result suggests that new proposals in UX evaluation should also consider indicating how to correct the identified UX problems in a way that software engineers can make improvements in the application, enhancing its UX.

Finally, regarding the availability of the methods, we noticed that many scales or questionnaires for UX evaluation do not provide the entire artifact (e.g., [S79][S87][S121][S144]). Also, other methods that propose a new process or tool do not provide its artifacts or are not available for download (e.g., [S81][S109][S116][S119]). This feature affects how these methods can be applied in the market. Furthermore, there are some methods which require specific equipment to be applied (e.g., [S128][S157]), which enhances the cost of applying the method. In this sense, there is a need for more methods that can be applied without costs or are fully described to be employed in real development scenarios.

Our findings regarding the evaluated period of experience suggest that further attention should be given to developing methods for evaluating UX before usage and during long term usage situations. Additionally, the methods that are already published should also be made available in order to allow their use in the software development market. Although some methods still require specific equipment to be applied, the others that do not should describe the process for their application and the necessary items/questions/artifacts that should be used.

Overall, our results show that there is a need for UX evaluation technologies able to provide improvement suggestions once a problem is found. This problem is recurrent in all types of applications, specifically when evaluating mobile applications. Therefore, new research should guide the development of approaches that provide assistance to software engineers willing to improve the quality of the evaluated application in terms of UX.

2.4 Design Suggestions

Considering the problems identified above, an alternative way to support the correction of identified interface problems affecting the UX of an application could be the use of design suggestions (Seffah, 2010). A design suggestion is a well-working solution to a problem. In addition to describing such best-practice solution, a design suggestion can

also contain a rather extensive discussion of when it is appropriate to use a specific solution and what the consequences (positive and negative) are if it is applied (Kircher and Völter, 2007). According to Mu and Jiang (2011), applying design suggestions enables developers to reuse them to solve a specified design issue.

User interface design suggestions have been introduced first as a medium to capture and represent solutions to users' problems (Seffah, 2010). These suggestions have also been used as a medium for transferring the expertise of HCI designers and usability practitioners to software engineers, who are usually unfamiliar with UI design and usability principles. When designing software applications that provide a positive UX, user interface design suggestions are tools that provide a means to abstract and reuse the essential details of the successful and usable design.

Overall, design suggestions provide some of the following benefits (Seffah, 2010):

- They are straightforward and readable for designers, developers, and other stakeholders, and can, therefore, be used for communication purposes;
- They represent design knowledge from different views, including social and organizational aspects, conceptual and detailed design;
- They capture essential principles of good design by telling the designer what to do and why, but are generic enough to allow for different implementations.

Several collections of design suggestions have been proposed in HCI. For instance, Tidwell (1998) proposed a set of Human-Computer Interface Design Suggestions, which is intended to support high-quality interaction between a person and a software artifact, aimed at one or more of a broad spectrum of activities, ranging from the most passive (e.g., absorbing information with little or no interactivity) to the hands-on (e.g., creation of other objects). Furthermore, Welie (2011) proposed a Library for Interaction Design which contains best practices with examples and insights on their applicability. According to the author, the solutions described in these suggestions may succeed in one context but may also fail in another. Thus, the challenge is to understand why and how it depends on elements of the context of use. Another example is the Design of Sites by Duyne et al. (2002) which provides a common language for

articulating a variety of Web designs. Such language explores best practices from the authors' consulting, research and Web development experience.

According to Dix et al. (1998), it is not necessary to follow a specific design method to use a set of design suggestions, since the value of any interface design technique is that it forces the designer to remember that someone (mainly someone else) will use the system under development. Borchers (2001) states that it is possible to apply design suggestions in different activities aiming at developing applications with a positive user experience:

Knowing the user: The development team can begin eliciting application domain concepts in the form of suggestions to give a uniform format to what needs to be captured, explicitly stating problems.

Competitive analysis: Existing products can be examined to gather information and hints for the design of the new system. Thus, user interface design solutions of successful competing systems can be applied to a system under development.

Setting UX goals: Various aspects of UX (e.g., usability) need to be prioritized. By considering HCI design suggestions, these aspects can be identified, and their conflicts can be resolved.

Parallel design: Design suggestions can serve as a common ground for software engineers designing different parts of a large software application.

Applying guidelines and heuristic analysis: Design suggestions can improve the use of standards through their structured format and contents, combinations of existing concrete examples and a general solution, and an insightful explanation not only of the solution but also of the problem context.

Testing: While design suggestions cannot help the actual evaluation process, they can be used to relate discovered problems to their corresponding suggestion to solve those problems.

Iterative design: design suggestions are an important tool to inform the designer about changes to be made in the application because they are constructive (i.e., they suggest how a problem can be solved).

Collecting feedback from field use: As design suggestions point out alternatives for solutions that need to be improved, feedback can strengthen the argument of those suggestions that created a successful solution and suggest rethinking those that led to suboptimal results.

2.5 Technologies Used in the Proposal of our Approach

In the following subsections, we present evaluation methods and redesign suggestions that have been considered as the basis for our approach. These technologies were considered as they allow evaluating hedonic and pragmatic aspects of UX and can be related to the context of our research. Further information on these technologies will be provided in Chapter 4.

2.5.1 Heuristic Evaluation

The Heuristic Evaluation (Nielsen, 1994) is an evaluation method for finding the usability problems in a user interface so that they can be attended to as part of an iterative design process. The heuristic evaluation involves having a small set of evaluators examine the interface and judge its compliance with recognized usability principles. In general, the heuristic evaluation is difficult for a single individual to do because one person will never be able to find all the usability problems in an interface (Nielsen, 1994).

The Heuristic evaluation is performed by having each individual evaluator inspect the interface alone. Only after all evaluations have been completed are the evaluators allowed to communicate and have their findings aggregated. This procedure is important to ensure independent and unbiased evaluations from each evaluator. The rules applied during the Heuristic Evaluation are shown in Table 2-1.

Table 2-1 Heuristics proposed by Nielsen (1994)

Heuristic	Description
Visibility of system status	The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

Heuristic	Description
Match between system and the real world	The system should speak the user's language, with words, phrases, and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
User control and freedom	Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.
Consistency and standards	Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.
Error prevention	Even better than a good error message is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.
Recognition rather than recall	Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for the use of the system should be visible or easily retrievable whenever appropriate.
Flexibility and efficiency of use	Accelerators - unseen by the novice user - may often speed up the interaction for the expert user such that the system can accommodate to both inexperienced and experienced users. Allow users to tailor frequent actions.
Aesthetic and minimalist design	Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.
Help users recognize, diagnose, and recover from errors	Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
Help and documentation	Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

2.5.2 Geneva Emotion Wheel

In Paper [S24], Scherer proposed the Geneva Emotion Wheel. This method is a graph that organizes the emotions in families according to their meaning and uses the

dimensions of valence and control. The arousal dimension is measured through the size of the circle that the user employs to determine his/her emotion. To evaluate their reaction to a stimulus, users employ the Geneva Wheel to depict their emotions in terms of valence (negative/positive), activity and arousal (intensity and control). When using this method, users can evaluate the appraisal dimensions (arrangement of emotion in terms of the two-dimensional space) and the intensity of the associated subjective feeling (in terms of the distance from origin). Figure 2-2 shows the different emotions that were identified by Scherer [S24] to evaluate the hedonic aspect of UX.

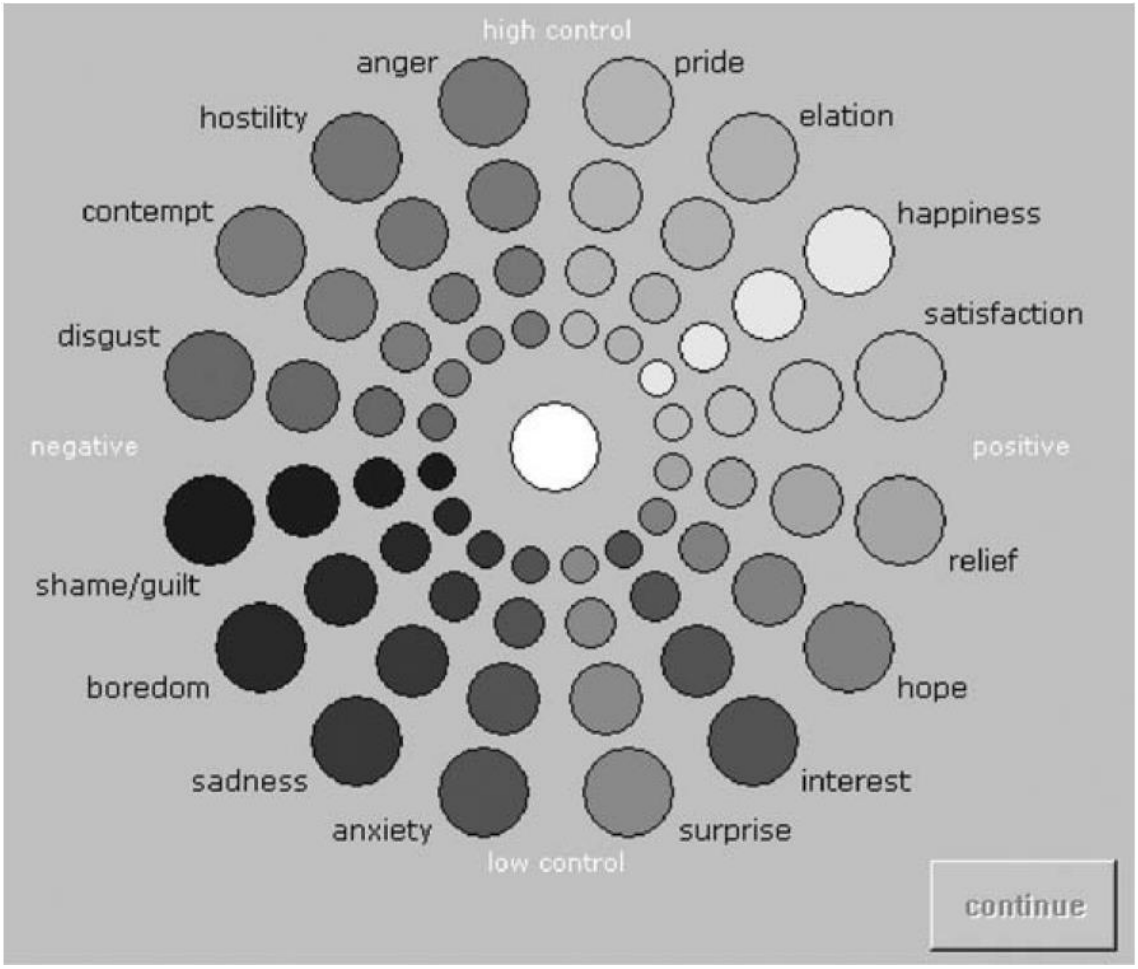


Figure 2-2 Structure of the Geneva Emotion Wheel proposed by Scherer [S24]

2.5.3 Interface Design Suggestions for Mobile Application

Neil (2014) proposed a collection of problems paired with solutions that could be applied to the design and redesign of mobile applications. The author indicates that the design suggestions library has been created considering existing problems and examples that were found useful to deal with those problems. Additionally, the experience from

previous problems was useful for deciding which suggestions could be applied in which contexts. For instance, Figure 2-3 shows an example of a design suggestion by Neil (2014). In this example, the author employed the "Error Message" suggestion, indicating what should be informed to the user when carrying out a task in an application and an error occurred. The design suggestion indicates that "*error messages should be expressed in plain language (no codes), precisely indicate the problem, and suggest a constructive solution. Best practice is to make Error Messages highly visible on screen. These approaches are preferable to modal dialogs, since they may literally cover up the issue.*" In this case, the example shows how error messages should be presented to the user and what pragmatic problems this could improve.

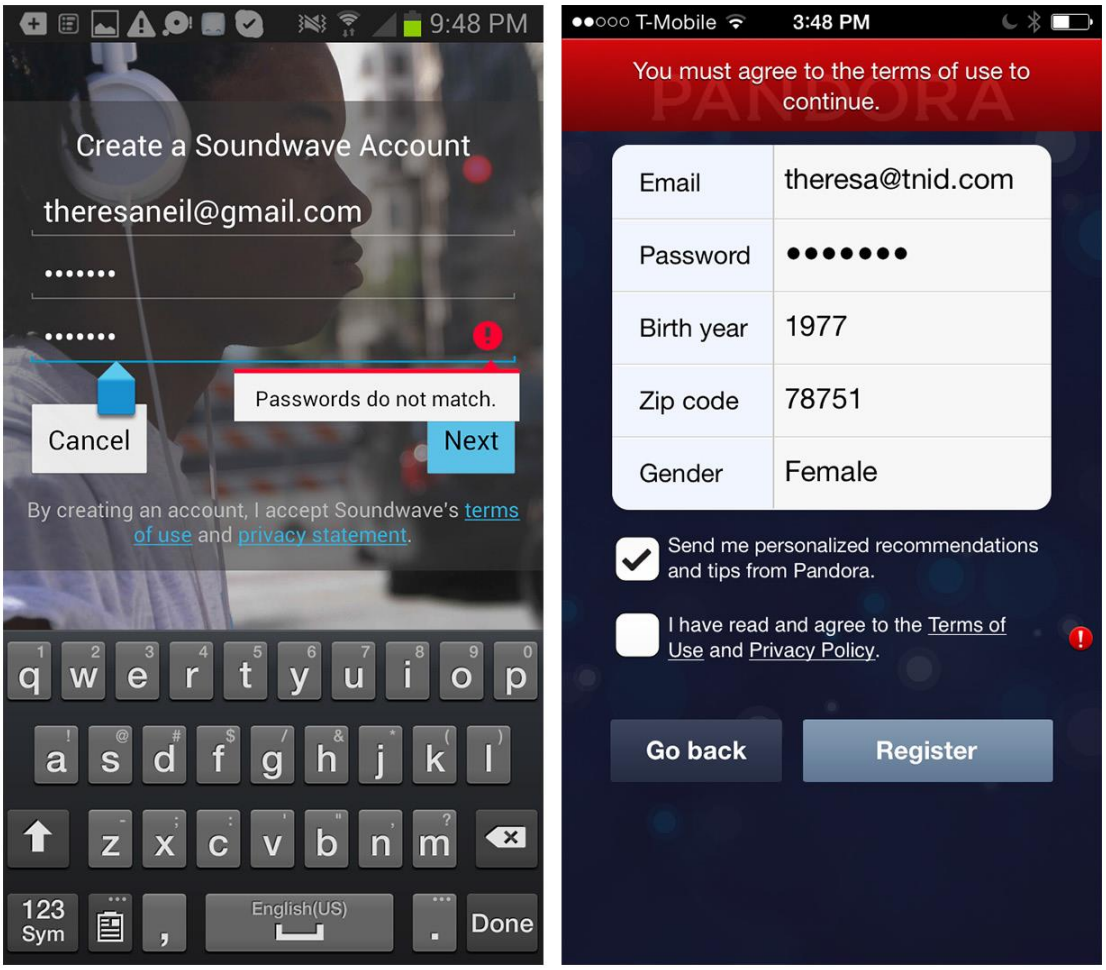


Figure 2-3 A screenshot of an application showing the 'error messages' suggestion from the design suggestions by Neil (2014)

2.6 Summary

User eXperience is about technology that fulfills more than just instrumental needs in a way that it acknowledges its use as a subjective, situated, complex and dynamic encounter. In that context, this chapter presented an analysis of UX evaluation technologies and discussed the use of design suggestions as an approach for correcting the identified problems.

The advantages and disadvantages of each type of technology have been discussed. Besides increasing the understanding of existing UX evaluation technologies and research opportunities, this review can be a basis for practitioners willing to choose a technology in a specific stage of the development lifecycle and for the evaluation of specific types of applications. For practitioners who are working in software development companies, these results could be used for integrating UX evaluation technologies into their software development process. For researchers, it would be interesting to propose (or adapt) new technologies according to the research gaps identified above.

Among the identified research opportunities in terms of UX evaluation technologies, we noticed that there is a need for methods that allow identifying more qualitative data besides rating the UX, as this can help evaluators understand what the cause of the problems or ways to solve them is. In this sense, we also noticed that few methods provide means to facilitate the correction of the identified problems. This is essential, as novice software engineers carrying out the evaluations require guidance in order to make changes in the evaluated application to improve its quality. Finally, there is a need for methods that are available for software engineers without the need of costly equipment, to enable small teams to improve and increase the acceptance of their applications.

To provide other means to identify design solutions to UX problems, we have discussed design suggestions as an alternative to allow communication in interdisciplinary design teams. These suggestions can help people from outside the development team (including users) to understand solutions. Also, they can be constructive in a way to allow both the discussion of changes and the understanding of the user feedback.

Due to the advantages that the use of both evaluation technologies and design suggestions can bring to each other in the development of a software application, in need for improvement suggestions for UX problems identified in evaluation sessions, we have considered developing a hybrid approach to facilitate the redesign process of mobile applications. To verify what could be improved in the evaluation and redesign process of a mobile application, in the next chapter, we present two empirical studies evaluating: (a) how users report their experience through UX evaluation methods; and (b) how novice software engineers apply design suggestions for improving an application.

3 Initial Studies of UX Evaluation and Design Technologies

This chapter presents the results of two empirical studies to identify research opportunities in terms of current proposals of UX evaluation methods and how novice software engineers apply design suggestions to correct pragmatic problems, respectively. In the first study, we have analyzed the features that make it easy or difficult for users to employ two UX methods: Expressing Emotions and Experiences and EmoCards. In the second study, we collected data on (a) the quality of the redesigned applications, and (b) factors affecting the applicability of the employed design suggestions. Our results suggest requirements for UX evaluation technologies and need to facilitate the redesign process.

3.1 Introduction

Considering the results from Chapter 2, there is a need for further investigation on the available technologies for evaluating the UX of interactive applications, while understanding what is needed to support software engineers in identifying improvement opportunities based on the identified UX problems. Therefore, this chapter presents two empirical studies. The first study aimed to identify research opportunities in terms of current proposals for UX evaluation, while the second study aimed to understand how software engineers apply design suggestions to improve software quality.

In the first study, we applied two UX evaluation methods: Expressing Emotions and Experiences - 3E (Tähti and Niemelä, 2006) and EmoCards (Desmet et al., 2001). We chose these methods as they intend to make users feel comfortable, reducing the effect of the evaluation of their experience. Also, we have analyzed the results of employing 3E and EmoCards in terms of identified UX problems and the perception of users on the employment of these methods for the evaluation of interactive products.

Through this study, we have identified improvement opportunities on the evaluated methods to guide future research in the evaluation of UX.

In the second study, we investigated how novice software engineers apply design suggestions and what makes it easy or difficult to employ them during the redesign of a mobile application. To this end, four teams of novice software engineers participating at a course on software quality identified usability problems through a usability inspection and, using interface design suggestions, proposed modifications in the interface of two real mobile applications in the market. Here, we present our findings regarding such use in terms of quality of the developed redesigned applications and the factors that affected applying the user interface design suggestions.

The remainder of this chapter is organized as follows. In Section 3.2, we describe the empirical study in which we show how we employed both 3E and EmoCards in the evaluation of a real Web application under development. Section 3.3 presents the planning, execution, and results from the second study, where we evaluated the use of design suggestions from the point of view of novice software engineers. Finally, in Section 3.4 we provide our conclusions for this chapter.

3.2 Study 1: Evaluating 3E and Emocards from the Point of View of Users

According to Vermeeren et al. (2011), there is a need for future research in analyzing the practicability of UX evaluation methods. UX evaluation methods need to be usefully employed in product development and be easy to use both by practitioners and users. Our initial assessment proposes to provide insights for future research and an example of the applicability of current approaches for the UX evaluation of a real application under development.

3.2.1 Planning and Execution

In Chapter 2, we described several methods such as SAM, Psychophysiological Techniques, the Affective Diary, the Visual Aesthetics Scale (and others), that allow the UX evaluation of interactive products. Table 3-1 shows a brief analysis of the advantages and disadvantages of these methods. These disadvantages show that there is a need for further investigation on methods that capture both the emotion and its context to explain what aspect of the interaction affected the feelings of the user (Tähti and

Niemelä, 2006). Some methods that can assist in the identification of UX problems and their cause are the Expressing Emotions and Experiences - 3E and EmoCards.

Table 3-1 Advantages and disadvantages of SAM, Psychophysiological Techniques (PT), the Affective Diary (AD), the Visual Aesthetics Scale (VAS), 3E and EmoCards (EC)

	SAM	PT	AD	VAS	3E	EC
Advantages						
Helps users identify their experiences			X		X	X
Quick and easy to use	X			X	X	X
Requires few resources	X			X	X	X
Gathers objective measures		X				
Disadvantages						
It takes time to be applied			X			
Users have difficulty in understanding the represented emotions			X			X
Does not explain the reason for the emotion	X			X		
Expensive equipment		X				
Makes users feel uncomfortable		X	X			

As shown in Table 3-1, besides being cheap, 3E and EmoCards allow gathering more profound information than just the user's feeling at the moment of interaction (Tähti and Niemelä, 2006). Moreover, users find expressing their emotions using both 3E and EmoCards a pleasant task (Desmet et al., 2001; Tähti and Niemelä, 2006).

Due to their advantages, we wanted to verify if users felt comfortable during the employment of 3E and EmoCards when carrying out a UX evaluation. Thus, the independent variables of this study are: (a) applying the 3E and EmoCards for evaluating UX and (b) the evaluated Web application. Moreover, the dependent variables that were considered to measure the users' perception of the UX evaluation methods are: (a) preference of the methods or, in other words, which UX evaluation method the users would choose if given the chance; (b) the users' opinion on quality factors of evaluation methods (i.e., features that make the method easy to use and make

users feel comfortable when employing it) (Vermeeren et al., 2010); and (c) rate of UX problems found with each method, which indicates how many UX problems each method allowed researchers to identify in the evaluated application.

We chose to evaluate the "Similar Products" Web application since it was being developed for the common user (people willing to buy a product and interested in finding further products). "Similar Products" is a Web application that is currently under development by Universidade Federal do Amazonas as part of a project on information recovery. The main goal of the project is to assist citizens when buying a product, so they can easily find out which other products are available that are similar in characteristics when compared to the desired product. Therefore, a positive UX was strongly desired in order to enhance its acceptance (Hassenzahl, 2005). Also, the development team of "Similar Products" was interested in improving the quality of the application earlier in the development process, while its mockups were being developed.

To evaluate how users felt when using 3E and EmoCards, we have employed the Smileyometer (Read and MacFarlane, 2006) which is a discrete Likert-type scale to apply a judgment score after an experience (in this case, employing each method). Our goal when applying the smileyometer is to have an idea of how users rate their experience of applying each of the UX methods. Also, to verify other factors (such as difficulties, how easy it was to employ the method, others), we have applied a questionnaire asking for the subjects' opinion regarding 3E and EmoCards. Finally, to measure the rate of UX problems found with each method, we have counted the number of problems that each method allowed the researchers to identify.

We carried out the study in 2014, with potential users from the "Similar Products" Web application, residing in the city of Manaus (Brazil). The people participating in this study were students/workers approached at Federal University of Amazonas while sitting in a lobby during lunch/break time. In order for a person to participate in the study, first (s)he was asked if (s)he had free time and would like to participate in a user interface design activity. If the person agreed, we would ask questions to see if (s)he met the user profile of a user from the "Similar Products" application. Thus, (s)he was asked questions regarding: (a) if (s)he regularly accessed the internet; and (b) if (s)he bought or browsed products on the internet. All subjects

who answered “Yes” to these questions were asked to participate in the UX evaluation. At all, 10 people (4 male and 6 female) with ages ranging from 18 to 37 years old (median 26,5 and SD = 6,67) meeting the selection criteria were asked and agreed to participate in the study. All subjects signed a consent form explaining the main activities of the study and the confidentiality of the results.

There were other participants in this empirical study: (a) the development team, who was responsible for preparing the mockups and creating a navigable PDF file containing the mockups using the Balsamiq Mockups tool (http://balsamiq.com/); (b) the moderator, one of the researchers from our research team, who guided the UX evaluation sessions; and (c) the analysis team, which was responsible for gathering all responses to the questionnaires and preparing a summarizing report with the overall results of the UX evaluation and the perceptions of users regarding the employed UX evaluation methods.

Figure 3-1 shows the procedures we applied for performing the UX evaluation using 3E and EmoCards. In this empirical study, the subjects acted as users of the “Similar Products” Web application.

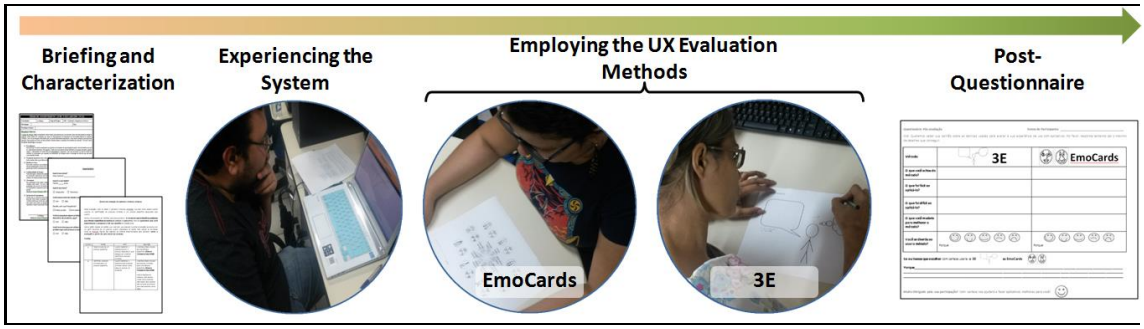


Figure 3-1 The process for carrying out the empirical study evaluating 3E and EmoCards

After agreeing to participate in the study and signing a consent form, users were asked to enter a lab room, where all equipment and materials necessary for performing the UX evaluation were prepared. Thus, the questionnaire asked for further information regarding the subjects’ age, gender, the habit of buying or browsing products on the internet and willingness to use an application such as “Similar Products.” Then, for each subject, the moderator explained him/her that (s)he would view a proposal for the “Similar Products” application and that the goal of the evaluation was not to evaluate the user, but how (s)he felt when using the application and identifying improvement

opportunities. While experiencing the application, the users performed tasks using the PDF file which mapped the mockups (see some mockups of the "Similar Products" application in Figure 3-2). Such file contained previously added links that had been created using the Balsamiq Mockups tool and allowed simulating interaction and navigation between the mockups. The tasks were selected due to their importance for achieving the main goal of the "Similar Products" application: (a) search for a specific product for which users wished to view similar products, and (b) view similar products for a selected product and rate their similarity. After experiencing the "Similar Products" application, the users would employ both UX evaluation methods, 3E and EmoCards to evaluate their user experience. We highlight that the order in which these methods were employed was randomly assigned, guaranteeing that 5 subjects employed 3E and then EmoCards (in that order) and then 5 subjects employed EmoCards and then 3E (in that order) to avoid bias. Finally, after employing the methods to evaluate their experience with the application, the users were given a follow-up questionnaire which contained the Smileyometer (Read and MacFarlane, 2006) and open questions regarding their opinion on the methods for evaluating UX.

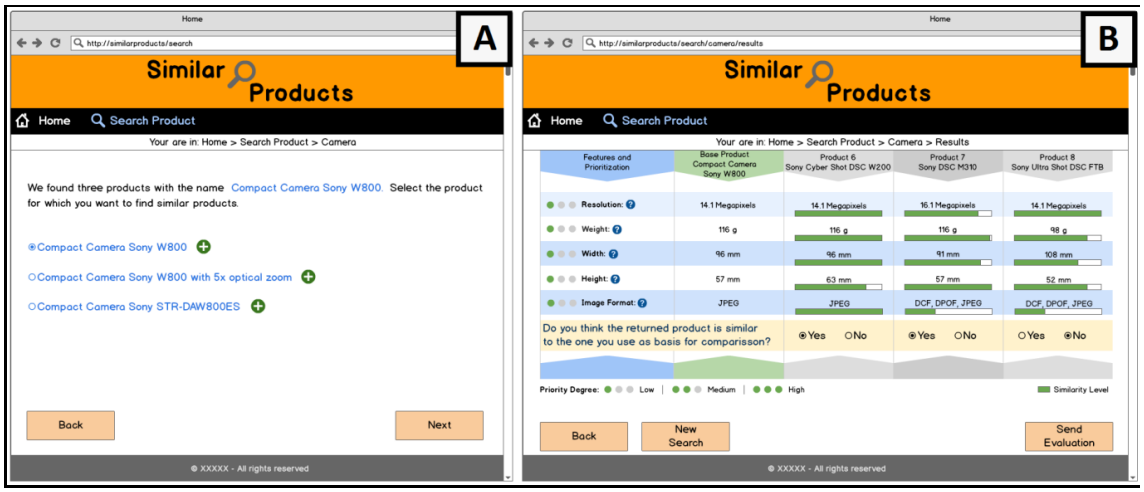


Figure 3-2 Mockups from the "Similar Products" Web application: (A) search results for the base product screen and (B) results for the similar products screen

We wanted to collect data from users who had experienced both methods (3E and EmoCards), so they would provide feedback regarding which method they preferred. Since there were no dropouts and all users filled the characterization form and follow-up questionnaires, none of the users' data were discarded.

The PhD candidate of this research acted as the moderator of the evaluation, checking the responses to the follow-up questionnaires regarding: the users' judgment score after their experience with the UX evaluation methods, their preference, and the answers to the open questions regarding difficulties, how easy it was to employ the UX evaluation method and improvement opportunities. In the following section, we describe our findings regarding 3E and EmoCards from the point of view of users.

3.2.2 Results

The material collected using EmoCards and 3E were analyzed by our research team. We give here examples of the results we extracted from the materials. The examples have been selected as they show how each UX evaluation method can support the identification of causes for the expressed emotions. Also, to facilitate the identification of the users, we will use the code UXX, where XX represents the number of the user (ranging from 01 to 10).

In this UX evaluation, we followed the suggestions by Desmet et al. (2001) who employed the EmoCards to help users express their emotional responses. Thus, through an unstructured interview, we asked users for the reasons for selecting a specific card, and how it had affected their experience. Although the questions from the unstructured interview was similar, in order to encourage the user to report his/her experience, we had to ask further questions. This was done for situations in which the user provided information on a UX problem, but further details were required. Below, we show an extract from the interview with user U07, showing both positive and negative aspects that influenced her experience:

- **Interviewer:** *If you had to choose a card to represent your experience with the application, which one would you choose?*

- **U07:** *I would choose this one. (U07 selects the exciting/neutral card)*

- **Interviewer:** *What made you feel that way?*

- **U07:** *I think that the desire to find the product I was looking for, and then the similar ones (...)*

- **Interviewer:** *And would you choose any other card besides that one?*

- **U07:** *Yes, I would choose this one too. (U07 selects the average/unpleasant card)*

- **Interviewer:** *Why?*

- **U07:** *Well. Because it was very difficult.*

- **Interviewer:** *What was difficult?*

- **U07:** *(...) There was only text, and I wanted to see images of the products as well. It would have made it easier to compare the products. (...) Also, the background color is too bright.*

- **Interviewer:** *What color?*

- **U07:** *(...) The orange color. It just distracts me from what I want to read.*

Regarding the results obtained from 3E, in general, the users produced very communicative drawings and explanations. Figure 3-3 shows the 3E template filled by user U09. We can see that user U09 had mixed feelings about the application and therefore, drew two faces at the same time: one happy and surprised and another angry one. In the oral expression balloon the user indicated that she: (a) felt confused, (b) that the application had a lot of disorganized information and (c) that many steps were required to perform the tasks. Similar information was provided in the inner thoughts balloon. However, there was a feature described by the user that differed from what she really believed. In the oral expression balloon, the user indicated that she thought that the color orange was "nice." Nonetheless, when analyzing the inner thoughts balloon, the user stated that she "hated orange." This would suggest that some users might not reveal their real opinions while being interviewed and that 3E might be able to capture those inner thoughts.

At all, we managed to identify 14 problems that affected the overall experience of the users by carrying out 10 UX evaluation sessions of about 30 minutes each. Since each user carried out an evaluation of both methods separately and there was only one moderator to guide the evaluation session, the total execution of the study took around 5 hours in total. The time needed to evaluate the UX of the application would have been shorter if we had not applied a follow-up questionnaire on their perception of the

employed UX evaluation methods. Also, there was no need for training in the use of these techniques, only a short explanation on how to fill in the template and the range of emotions that could be selected in the Emocards method. In the following subsections, we refer to the number of identified problems and the overall perception of the users towards the methods to draw conclusions about the applicability of the methods.

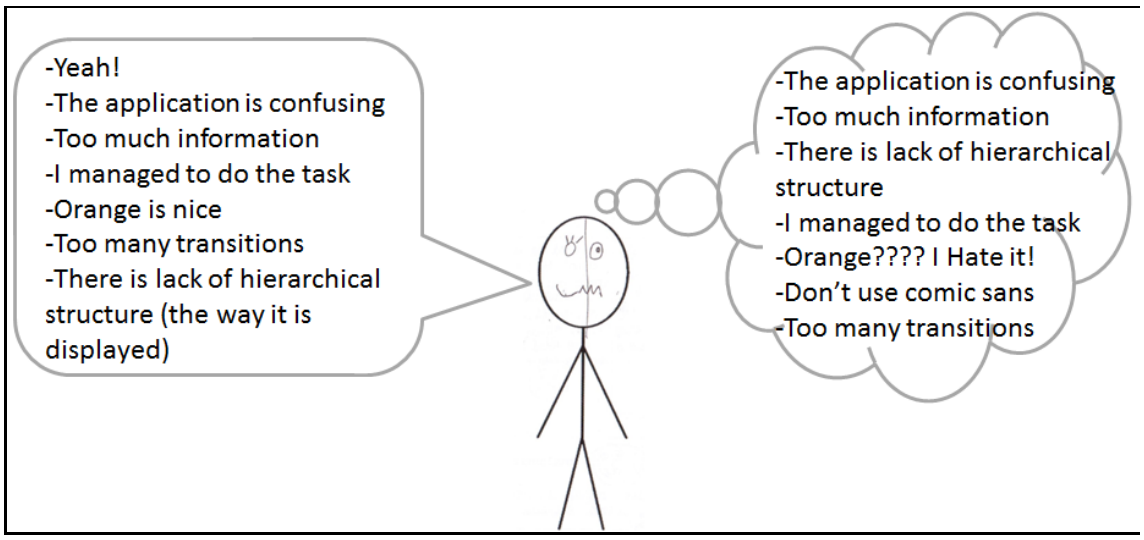


Figure 3-3 A 3E template filled (translated from Portuguese) by one of the users with mixed feelings

The smileyometer measured the users' judgment on a Likert-type scale (Read and MacFarlane, 2006): (1) Awful, (2) Not Very Good, (3) Good, (4) Really Good, and (5) Brilliant. Thus, the maximum judgment on a method is 5. The descriptive statistics showed that the median score (4) for the 3E method was higher than the median score (3,5) for the EmoCards method, in general, users gave similar scores for both methods, meaning that they enjoyed applying both of them. However, when asked to choose one of the methods, 3E achieved higher preference (6 out of 10). The explanation for such result will we describe below.

The users gave both positive and negative feedback regarding the 3E method for providing affective information about using a system. Comments regarding what made the 3E method easy to use include:

"It was easy to describe the emotions by writing."

"I like feeling free when I'm doing things, and this method allows me to express myself, describing what I am feeling and thinking. It also allows me to draw."

"Very nice, it uses graphs."

"It is easy to draw my emotion."

"It is not always easy to say what you think, so this method makes me feel at ease."

Some users also pointed difficulties in using the 3E method. However, some of the difficulties were contradictory to the features that made the method easy to use. For instance, some users stated that they liked drawing the faces, while other stated that they found doing that difficult. Moreover, some users indicated that the balloons were not that easy to identify and that it was difficult to think of what to write. This would suggest that perhaps 3E might be more suitable for users who prefer a free environment to describe their experience. Comments on difficulties on using the 3E method included:

"It is very difficult to draw an expression."

"It is not that easy to list what you are thinking. I think it would be necessary to have an analytical view of your thoughts in order to use this method."

"It is difficult to differentiate the balloons."

Regarding the use of the EmoCards method, some users stated that it was easier to think of the emotion as they had a predefined set of cards. Also, such approach allowed them to think of what caused the emotion and describe it. Comments regarding what made the EmoCards easy to use include:

"It already has emotions from which you can choose."

"It is easy because you just have to choose and talk."

"It is easier to express yourself based on the emotion."

"It helps me define my emotions."

When the users answered the questions regarding the difficulties of using the EmoCards method, most of them indicated that choosing the representative emotion was the hardest thing to do. Sometimes, the faces did not represent what the users wanted to express and therefore, having a limited set of cards made it difficult to provide an exact

evaluation of what they wanted to portray. Comments on difficulties on using the EmoCards included:

> *"It is difficult to choose the cards for each exact situation."*
>
> *"Some faces are similar, and it is difficult to find one face that actually fits what I want to say about the interface."*
>
> *"It is not as simple to understand how the method works. The other method (3E) is much easier to understand."*
>
> *"The order of the faces was confusing; perhaps they should be grouped to make it easier to find the right emotion?"*

Finally, when asked about which method they would choose, six users preferred 3E while four would choose EmoCards. Among the reasons for choosing 3E, users listed some of the advantages regarding its ease of use. For instance, users indicated that they felt freer in applying the method and that it allowed providing a thorough explanation of why they felt those emotions. On the other hand, the users who chose EmoCards indicated that the method was more dynamic and that it was much more visually appealing than the 3E method. Also, these users indicated that EmoCards would be more objective since they would be able to directly tell what they wanted to say.

3.3 Study 2: Evaluating Design Suggestions from the Point of View of Software Engineers

Even though there are several design suggestions that indicate alternatives for improving the UX of applications in general, few have focused on both suggestions for mobile applications interface design and their applicability by novice software engineers. For instance, in the study by Chung et al. (2004), the participants were experienced software engineers with knowledge in UI and Web design. Also, Lanzilotti et al. (2011) did not study the effect of applying design suggestions for (re)designing an application. Furthermore, although Koukouletsos et al. (2006) analyzed the quality of design proposals created by novice software engineers, their evaluation did not gather information on the application process of design suggestions and was not focused on

mobile applications. Consequently, there is still need to gather information on how novice software engineers apply mobile suggestions in the (re)design of mobile applications. Such information will be useful for identifying improvement opportunities in the application of design suggestions by novice software engineers in the industry and improving the UX of the developed applications. In the following section, we describe an empirical study that aims to collect such data.

3.3.1 Planning and Execution

We employed a qualitative research methodology aiming to understand which factors contribute or make it difficult to apply design suggestions in the redesign of a software interface by novice software engineers. In that context, we conducted the study at a university in the city of Manaus (Brazil) during a class on software quality at a Computer Science course focusing on methods for software quality assurance. The study took place during 6 weeks out of the 4 month duration class.

Some of the topics from the course were usability/UX, ways of identifying problems related to these aspects and how to correct them. Thus, training regarding usability and UX evaluations was prepared as part of the course. Also, this training contained examples of mobile applications in which design suggestions had been applied to correct usability and UX problems.

At all, 12 software engineering students in their last semester of college (at Manaus-Brazil) were enrolled in the class and agreed to participate in the study. These software engineers had a technical background (more than 3 years of experience studying and/or practicing in the area) on software development methodologies, but low or no experience in interface design. Consequently, the students had attended classes related to software engineering, such as software analysis, software engineering, software process, while attending a class on software quality related to human-computer interaction. These students were divided into four groups of three novice software engineers each. The groups were formed by the convenience of the students, as all students had the same degree of experience and none of them had practical experience in the industry. Each group participated in all of the following stages:

- **Stage 1 - Lectures and training:** In this stage, the novice software engineers participated in regular classes on the methods that were to be

applied (i.e., usability/UX evaluation methods and examples of redesign through design suggestions).

- **Stage 2 - Usability evaluation on mobile applications in the market:** In this stage, each group of novice software engineers chose between two mobile applications for Android devices and carried out a usability inspection using the Heuristic Evaluation (Nielsen, 1994) presented in Chapter 2. The evaluated applications were: (a) Manaus Bus[1], an application that supports users in finding the right bus to their destination; and (b) Manaus in Theaters[2], which helps users find out about the different movies on display in the theaters of the city. We focused on these applications as they are entertainment and information mobile apps[3], which require a high degree of usability to facilitate their use and enhance the UX of the users. Therefore, they could be a good example of how redesigning an application could improve its perceived quality in terms of UX (Ervasti et al., 2011). After collecting all inspection reports, the identified problems were verified by a high-experienced analyst with more than 5 years of experience in usability and UX evaluations to check which problems were real problems and which false positives. Thus, there was a meeting with each team discussing the inspection results.
- **Stage 3 - Correction of the identified problems using design suggestions:** In this stage, the novice software engineers employed design suggestions for mobile applications (Neil, 2014; Google, 2016) and proposed changes in the user interface. We made sure that the software engineers knew that they could apply the design suggestions as they felt appropriate. This was done to understand what process they would apply to redesign the application using the design suggestions and the difficulties they faced during the redesign process.
- **Stage 4 - Usability and UX evaluations to verify quality improvement:** In this stage, each team carried out a usability and UX evaluation (with between four to six users) to verify to what extent the redesigned application

[1]https://play.google.com/store/apps/details?id=com.manausemcartaz

[2]https://play.google.com/store/apps/details?id=com.onibus.manaus

[3]https://support.google.com/googleplay/android-developer/answer/113475

improved. Thus, each team prepared prototypes, linked navigable PDF files, for both versions of the application (original and redesigned) and they asked users from the streets to interact with them, taking notes on problems faced by users in both applications and their satisfaction degree. In this sense, data on usability issues on each version of the application was gathered through observations of the users' actions while interacting with the prototypes. Additionally, UX data was gathered through a questionnaire containing the Expressing Emotions and Experiences (3E) method (Tähti and Niemelä, 2006). Finally, users were asked to rate their overall experience with the app on a scale from 1 to 10 and indicate which of the experienced versions they preferred.

In this study, we applied a questionnaire with open questions as we aimed to: (a) assessing the acceptance of the use of design suggestions for the improvement of user interfaces by novice software engineers; and (b) identifying advantages, constraints and improvement opportunities for improving the adoption of design suggestions in the software industry by novice software engineers. Table 3-2 shows each of the asked questions and their purpose.

For instance, Q2 asks the software engineer which suggestions (s)he thought were the easiest and hardest to apply, which aims at making him/her think of these suggestions and internally think of the reasons for choosing the easiest and most difficult suggestion, preparing him/her for the next questions (Q3-Q4). This questionnaire was answered by the software engineers, right after finishing stage 3, so they would report their experiences with using the user interface design suggestions for redesigning an application.

Table 3-2 Questions from the follow-up questionnaire and their purpose within our study

ID	Open Question	Purpose
Q1	Which were the steps that you followed in order to apply the design suggestions?	To identify how novice software engineers apply the design suggestions.
Q2	Which was the most difficult design suggestion to apply? And which was the easiest to apply?	To make the novice software engineer think of the aspects that make it easy or difficult to apply a design suggestion.

Q3	Please, describe the difficulties that you faced when searching/understanding/applying the design suggestions.	To identify what makes it difficult to apply design suggestions for novice software engineers.
Q4	Please, describe the aspects that made it easy to search/understand/apply the design suggestions.	To identify what makes it easy to apply design suggestions for novice software engineers.
Q5	If you had to redesign an application again, would you consider using the design suggestions? Why or Why not?	To understand the overall opinion of novice software engineers regarding the use of design suggestions for the correction of UX problems.
Q6	What would you do to facilitate the redesign using design suggestions?	To identify improvement opportunities in the use of design suggestions for improving the UX of a software application.

3.3.2 Results

At all, four redesigns were proposed. As an example of how the novice software engineers applied the design suggestions, Figure 3-4 shows the original and redesigned versions of the Manaus Bus application for one of the teams participating in this study. Here, the software engineers applied two interface design suggestions: (A) Superior Bar (Tabs) (Google, 2016) and (B) Auto-Complete (Neil, 2014). In a report, the software engineers had to explain the reason for choosing a specific suggestion and what problem they were trying to fix. We have used this report to point out the problems in the original version of the application and how the novice software engineers applied the design suggestions in the redesigned version (see Figure 3-4).

Regarding the first identified problem in Figure 3-4 (see Item 1-A), during the evaluation, the participants identified that the icon was not clear as it did not specify what type of search was being performed. Also, it was necessary to go back and forward in the application to select the desired search option. Therefore, the software engineers within this group decided to use the Superior Bar design suggestion, including each option in a tab and providing labels to facilitate understanding (see Item 2-B). Additionally, regarding Item 2-A, the participants noticed that whenever it was necessary to search for the place, neighborhood or street within the route of a bus, the user had to go through an extensive list of names by browsing. Therefore, the participants decided to add a search field to facilitate identifying the desired option by the user (see Item 2-B).

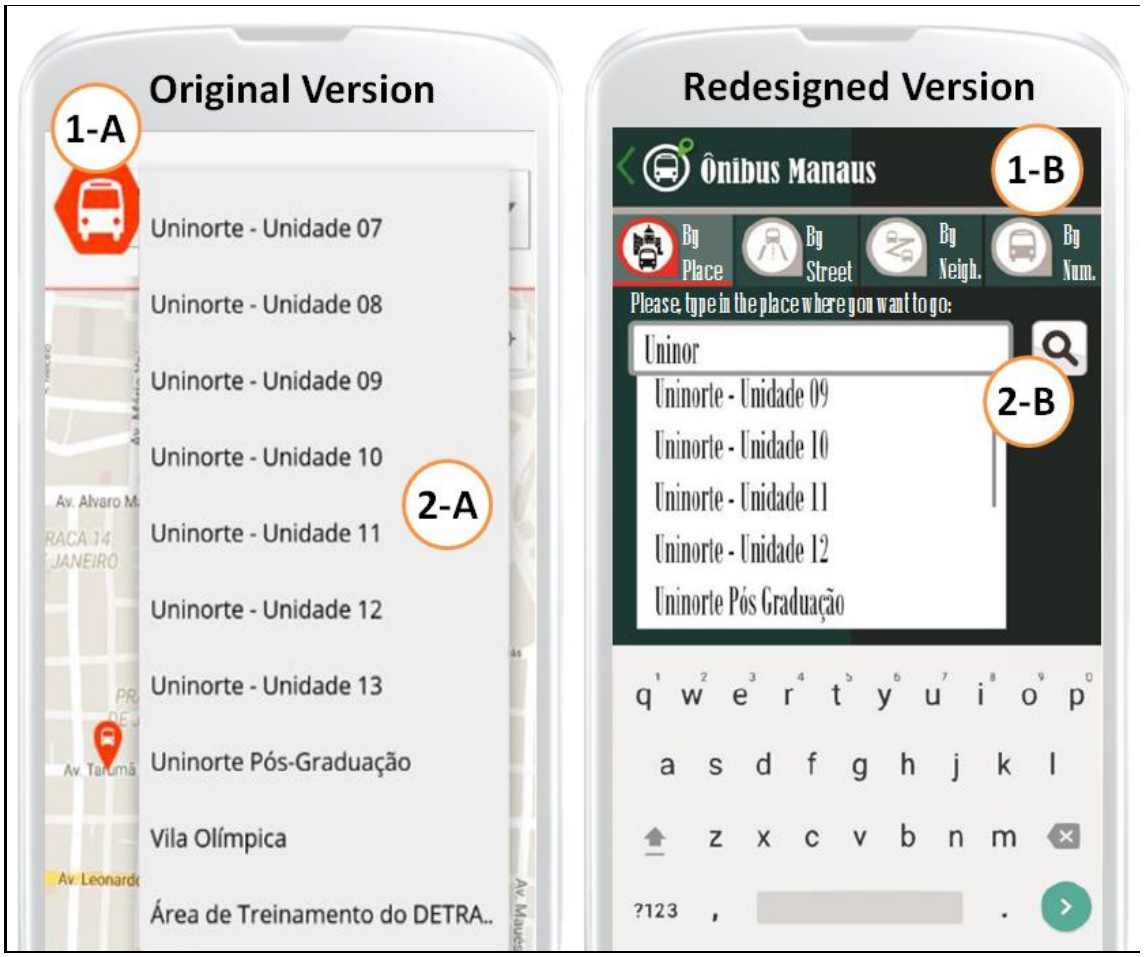
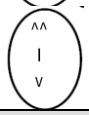


Figure 3-4 Example of redesigned application using design suggestions

In Stage 4 each team carried out usability and UX evaluations to verify whether the changes in the user interface had any impact in the UX of the users. Table 3-3 shows overall results of such evaluations. In this sense, users who tested both versions of one of the evaluated applications were asked to: rate the versions (original and redesigned) by giving a grade from 0 to 10, indicate which version they preferred and fill the 3E template. The examples of answers to the 3E method in Table 3-3 are from four different users, each experiencing the two versions of the app from each team. In this sense, the redesigned versions of each app obtained higher median scores and achieved a higher preference from users. Also, the experienced depicted by the users through the 3E method showed that users were more pleased after using the redesigned versions of the apps. These results suggest that by applying design suggestions, novice software engineers were able to improve the quality of the mobile applications in terms of UX.

Table 3-3 Overall results of the UX evaluations comparing original and redesigned applications

Version	AS	TP	Example of 3E Method Answer	
Team 1 (Manaus Bus)				
Original	5,3	0		*"I am stressed and very angry with this app."*
Redesigned	9,5	6		*"I am very happy and excited to test it."*
Team 2 (Manaus Bus)				
Original	5,5	1		*"Everything looks the same."*
Redesigned	6,7	5		*"This app is really simple."*
Team 3 (Manaus in Theaters)				
Original	5,8	0		*"It could be better organized with dates and times."*
Redesigned	7,8	4		*"This app is more complete with very interesting information."*
Team 4 (Manaus in Theaters)				
Original	8,2	0		*"I hate it."*
Redesigned	9,0	6		*"The app is really cool, and it could be very useful."*

AS = Average Overall Score, TP = Total Preference

In our study, we analyzed the data obtained from the answers to the questionnaires (see Table 3-2) using procedures from the Grounded Theory (GT) method. Grounded Theory (i.e., Data Grounded Theory) uses a set of systematic data collection and analysis procedures to generate, prepare, and validate substantive theories on essentially social phenomena, or on wide social processes (Glaser and Strauss, 2009). Although the purpose of GT is the construction of substantive theories, a researcher may use only some of its procedures to analyze qualitative data. According to Glaser and Strauss (2009), the coding process, in which concepts (or codes) and categories are identified, can be divided into three stages: open, axial, and selective coding. During the open coding, the researcher carries out the breakdown, analysis, comparison, conceptualization, and categorization of the data. After the identification of conceptual categories by the open coding, the axial coding examines the relations between the categories. The relations between codes can be defined by the very

researcher. Finally, the selective coding realizes all the process refinements by identifying the core category with which all others are related.

In our analysis, we have carried out the open and axial coding, but have not elected a core category yet because GT suggests the circularity between the collection and analysis stages until the theoretical saturation is reached (Glaser and Strauss, 2009). Therefore, we decided to postpone the selective coding phase and we do not claim that we applied GT, but some of its procedures.

After the novice software engineers had delivered their filled questionnaire, we started the codification process. While we analyzed the data contained within the questionnaires, we created codes related to citations within them. These codes were reviewed with an experienced researcher in qualitative analysis, who verified them and the categories in order to audit the coding process.

After performing the open coding, we initiated the axial coding, in which we created the relationship codes. We identified codes related to the process and steps that novice software engineers employed for applying the design suggestions, difficulties and facilitators, improvement opportunities and the reasons that would make them use design suggestions again. Below, we present the results of the qualitative analysis, referring to the novice software engineers as subjects using the code SXX, where XX is the number of the software engineers from 1 to 10. Note that we only received the questionnaires from 10 participants even though all participated in the stages.

Regarding the Process and Steps categories, some subjects described how they applied the design suggestions and how they included further stages to facilitate their applications. In this sense, some subjects thoroughly described the exact process (see a quote from Subject S06 on Process) while others were more direct (see a quote from subject S08 on Process). Despite the varying levels of detail in the description of the processes, we identified some specific steps that were incorporated by the subjects but were not suggested during the training. For instance, some subjects indicated that they searched for other information sources; not only other design suggestion proposals but examples of how the design suggestions had been applied in real applications to verify their suitability (see quotes from subjects S03 and S09 on Steps). Also, some subjects indicated that they approached the way in which the problem to be corrected was

selected, as well as the selection of the suggestion to correct it. In this sense, some of the criteria for prioritizing a problem was based on the difficulty on correcting a problem (see a quote from subject S10 on Criteria) or based on the main functionalities of the app (see a quote from subject S10 on Criteria). Regarding the selection of a suggestion, the criteria were: suggestions that solve the problem (see quote from subject S06 on Steps), suggestions that do not make the app difficult to use (see quote from subject S05 on Criteria) and suggestions that are more related to the app's logic (see quote from subject S04 on Criteria).

> *"We had to identify the problems in the application ... and, after that; we searched for other sources containing suggestions and possible solutions to the identified problems. We applied the suggestions that better solved the problem, following its guidance and we made changes in the interface."* – Process - Subject S06

> *"According to the application we were going to redesign, we chose as a team the best option that would fit that application..."* – Process - Subject S08

> *"... I also looked up additional content, besides the one provided in the android suggestions."* – Steps - Subject S03

> *"... I verified example in other applications."* – Steps - Subject S09

> *"...we sought what would be easier to correct in the app..."* – Steps - Subject S10

> *"...we worked on the main functionalities of the Manaus Bus application..."* – Criteria - Subject S10

> *"I applied the suggestion that better solved the problem"* - Criteria - Subject S06

> *"...I also verified which suggestion would be simple enough and would not bother the user or make it difficult to use the app"* - Criteria - Subject S05

> *"Understanding the logic of the app (how it navigates, shows content, others) seems to be the best way to choose a suggestion..."* - Criteria – Subject S04

Regarding the Difficulties and Facilitators, the subjects listed several factors that affected the use of the applied suggestions. For instance, some subjects indicated that applying the user interface design suggestions was difficult when: (a) it was difficult to adapt the solution, (b) more than one suggestion was available, and one had to decide which was better, (c) it was not possible to find a suggestion that fit the app, (d) the software engineer had no experience, (e) the software engineer did not know how to apply the suggestion, (f) the software engineer did not understand the suggestion, and (g) there were few examples of how the suggestion is applied. Additionally, subjects indicated that applying the design suggestions was easy when: (a) the explanation of the suggestion was simple, (b) there is training, (c) there are more examples, (d) the software engineers knows what (s)he wants to achieve, (e) there are images, videos and others explaining how to apply them, (f) the name of the suggestion is the same as the problem that needs to be addressed, (g) you can ask an expert about them, (h) there are tools which allow supporting the application process (e.g., tools for searching suggestions), and (j) there are heuristics that help you/guide you to choose a suggestion. Some quotes that illustrate these aspects are shown below.

> *"There was some doubt when applying the navigation suggestion, where we would associate the menu option with the help icon since both were associated with the action bar"* – Difficulties - Subject S03

> *"We based on the suggestions for mobile applications, and we did not find a suggestion for that problem"* – Difficulties - Subject S10

> *"Yes, because the source where we found the suggestions was very simple and direct."* – Facilitators - Subject S04

> *"It is, I believe the easiest suggestion was the 'links' since there was a suggestion with a name that was specific for that."* – Facilitators - Subject S02

At all, all subjects indicated that they would apply the design suggestions given a chance (see a quote from subject S07 on Intention to Use). Among the Reasons for Employing Design Suggestions, the subjects indicated that they thought the design suggestions provided support and they would guarantee that the application would not be redesigned without following standards (see quote from subject S05 on Reasons for Employing). Also, the subjects indicated that using the design suggestions would make

the application more visually appealing and organized (see quote from subject S07 on Reasons for Employing). Other aspects that motivated the subjects to employ the design suggestions were: providing confidence in the results of the redesign; that the suggestions were structured, which made them easy to apply; that they provide guidance; and that they provide solutions to common problems of applications.

> *"Yes, I would definitely use them, as they (Android design suggestions) follow the standards from the company…"* – Intention to Use - Subject S07

> *"it helps the software engineer a lot because there is a rule that makes sure that the application is not redesigned without guidance"* – Reason for Employing - Subject S05

> *"Yes, with the redesign we noticed that it helped a lot in the new layout and made the application much nicer"* – Reason for Employing - Subject S07

Finally, regarding Improvement Opportunities, the subjects indicated that some suggestions could be more specific, so they do not become ambiguous (see a quote from subject S01 on Improvement). Also, they suggested studying the steps for applying the design suggestions (see a quote from subject S02 on Improvement). This indicates that perhaps, they thought that the training was not enough (or could be improved) for applying the design suggestions. Nevertheless, we highlight that we intended to let the subjects apply the design suggestions freely, so we could verify what difficulties they encountered and what they included (or would include) in their application process to make it more effective. An interesting finding was regarding the use of the documents we prepared for delivering the activities. In order to collect the data on the redesign process and its results, in Stage 3, we provided the subjects with a template of a report, which asked the reason for choosing a specific suggestion. That report made some of the subjects become more selective of the design suggestion that they would apply, as they had to describe the reasons towards that choice (see a quote from Subject S05). Also, having identified the problems that they had to correct based on an inspection method facilitated identifying the suggestions that could correct them, as these suggestions had to allow meeting the heuristic's principles (see a quote from Subject S06). Such results may indicate that having a document guiding the redesign process and a set of rules that need to be followed when deciding which design suggestion to

use could be introduced into the application process, facilitating the use of the design suggestions by novice software engineers. Nevertheless, new studies would be necessary to test these hypotheses and evaluate the extent of support that these documents and guidelines could provide for novice software engineers.

> *"Some suggestions were too repetitive, so they could be made less similar."* – Improvement - Subject S01

> *"I would study more so I could understand them and apply them."* – Improvement - Subject S02

> *"The template of the document helped when carrying out the redesign… I wondered if choosing the specific suggestion would correct the problem."* – Subject S05

> *"The heuristics [from the heuristic evaluation] were important for identifying the problems. Thus they were also the reference for choosing suggestions that would not violate them."* – Subject S06

3.4 Summary

In the first study, we compared two UX evaluation methods in terms of the number of identified problems and preference from the point of view of users. Our results showed that users managed to point out more problems when applying the EmoCards. However, a combination of both methods is strongly suggested, as each one of them has its strengths. For instance, the 3E method is more suitable for users who prefer freedom when reporting their experience. Nonetheless, this is not always a feature that turns into an advantage for the UX evaluation team. As shown in the qualitative analysis, users who employed EmoCards felt more encouraged to report their problems (portrayed by the cards) and, by having predefined answers, they were able to think more easily of the different aspects of the application that affected their experience. This was not possible with the 3E method since users were trying to report their experience on their own. On the other hand, the 3E method allowed for capturing opinions that users might hide from the evaluators.

We can see that the main advantage is that both 3E and EmoCards allow capturing both problems and their causes while being as cheap, easy to use and quick as SAM and the Visual Aesthetics Scale. Furthermore, these methods allow users to reflect on their experiences as the Affective Diary without getting them tired. Finally, although 3E and EmoCards are enjoyable, they do not allow gathering objective measures the same way as Psychophysiological Techniques. Perhaps this is a research opportunity for these methods, allowing capturing objective data besides the users' opinions.

In our first study, besides describing the features that made the 3E and EmoCards easier or difficult to use, users made suggestions for improvement. For instance, regarding the 3E method, users who had difficulty in drawing suggested providing predefined faces that they could paste over the human body to express their emotions. Also, users suggested providing descriptions of the balloons in the template itself, so they would not forget what to write on them. Moreover, regarding the EmoCards, users suggested arranging the cards in a sequence or subgroups to facilitate the choosing process. Finally, other users suggested improving the cards' faces by making them less exaggerated.

In the second study, we verified to what extent novice software engineers can apply design suggestions in the redesign of user interfaces of mobile applications. After identifying UX problems, novice software engineers applied design suggestions to correct them and managed to improve the UX of users of the application. Also, in their answers to the questionnaire of their experience applying the user interface design suggestions, the novice software engineers indicated that they would apply the suggestions again if given a chance and indicated specific activities such as selecting the problems to be corrected and suggestions to address them. Nevertheless, they indicated that the lack of experience and knowledge on how to apply the design suggestions would make it difficult to do so. As improvement opportunities, we identified that document templates and a set of rules could be useful for guiding novice software engineers in the redesign process, making it easier.

Given that each qualitative study provides evidence and hypotheses that can be later tested using quantitative methods, we intend to evaluate how including the improvement opportunities identified in this study can facilitate the application process of design suggestions by novice software engineers. Our intention is to provide means

of evaluating UX features and facilitate the redesign process through the use of design suggestions, so novice software engineers can improve the quality of their applications. In the following chapter, we use the results from these studies to propose a set of technologies aiming at improving the results of UX evaluations and redesigns of user interfaces by software engineers.

4 The Proposed Technology

Based on our literature reviews and empirical studies, this chapter presents our proposal for an approach to support both the evaluation and suggestion of improvements in mobile applications. Here we present the concepts of Redesigning for EXperience (REX), which aims at collecting UX data from users while attempting to be less intrusive than other UX methods. Also, we describe our strategy for supporting the correction process of the identified UX problems, by suggesting improvement opportunities based on the identified problems.

4.1 Introduction

In Chapters 2 and 3, we analyzed evaluation methods for identifying UX problems and improving the UX in software applications and carried out two exploratory studies to verify what could be improved in the application of those methods. Based on the results from our review of existing methods and empirical studies, we were able to understand their advantages and disadvantages. Besides the lack of UX evaluation methods for mobile devices, we identified the following difficulties in using current UX evaluation methods:

(DI) Difficulty in Information Externalization: One of the problems in UX evaluation methods relying on users reporting their emotions is the users' difficulty in externalizing their feelings and opinions. This problem affects the effectiveness of the method, as users who are not able to easily express themselves may have difficulty in reporting problems that affected their experience. If a method is not able to gather information on the aspects that affected the users' experience, it will not be able to provide useful information for the development team to improve the evaluated software application.

(DII) Negative Influence of a Moderator: The problem of externalization that users face when reporting their experience increases when users are asked to express

their opinions next to a moderator. This can make users (especially those who are shy) feel uncomfortable or conceal information to please the moderator. Also, acting next to a moderator can cause users to behave differently than how they would behave in a real usage scenario.

(DIII) Negative UX during the Evaluation: According to Isbister et al. (2006), the experience of the feedback giving should be pleasant in and of itself, so users feel comfortable when employing the UX evaluation method. This is a problem in methods that use scales or report instruments that make the evaluation seem like a task.

(DIV) Lack of Redesign Suggestions: Another important issue with current UX evaluation methods is that most of them do not provide guidelines to correct the identified problems. According to Hornbæk and Frøkjær (2005), development teams require more than the list of problems to support the redesign process. In most cases, software engineers also value having solution proposals along with the identified problems to assist them in the correction process. These results were corroborated by our review and empirical study on the use of design suggestions (see Chapters 2 and 3).

Difficulty DI, DII, and DIII are related to the identification process of UX problems, while Difficulty DIV is related to the solution of those problems, i.e., the redesign of the software application to provide a positive UX. Considering the context described above, in this chapter, we aim at answering the following research question:

"*How can we easily evaluate the UX from the point of view of users and provide developers with means of identifying improvement opportunities, without having a boring or intrusive UX evaluation?*"

To provide a solution to the problems defined above, we have proposed a set of features that can be useful for future research in the development of UX evaluation methods meeting the current needs of the software industry:

(FI) To Facilitate Feedback Giving: UX evaluation methods should focus on facilitating the evaluation process from the point of view of users, by providing means to report their experience easily. This feature aims at improving the effectiveness of the method for gathering UX data, which will be useful in the next steps of the redesign process.

(FII) To Reduce the Influence of a Moderator: UX evaluation methods should focus on reducing the influence of the moderator by training practitioners in the evaluation process, or providing other means to collect UX data. The goal of this feature is to reduce the influence of the moderator in the answers of the user regarding his/her experience with the software application.

(FIII) To Make the Evaluation Pleasant: UX evaluation methods should focus on making the evaluation a pleasant activity and not seem like a boring and forced task. The purpose of this feature is to avoid users from providing poor information and wanting to finish the evaluation as soon as possible, or even drop the evaluation.

(FIV) To Provide Suggestions to Solve the Identified Problems: This final feature is related to assisting software development teams in the redesign process of a software application, after having identified a set of UX problems. The goal of this feature is to make the redesign process easier and provide meaningful and useful ideas that can be put into practice to improve the UX and the quality of the developed software application.

We will now discuss how the Redesigning for EXperience (REX) approach has incorporated the above features, in order to become a complete UX evaluation method and a means for identifying improvement opportunities. The remainder of this chapter presents the REX technique and its tool support, its architecture and an application example.

4.2 The REX Technique

The REX technique was developed using Hassenzahl's (2005) UX model which provides a set of concepts regarding how users experience products in general. In that model, Hassenzahl (2005) suggests evaluating UX from two perspectives: Hedonic, related to the emotions evoked by the product aiming at the users' well-being; and Pragmatic, related to the capability of achieving goals with the product with effectiveness and ease of use. Considering these perspectives, we organized the evaluation considering possible users' emotions, and possible pragmatic problems within mobile applications that could cause negative emotions.

To provide a set of basic emotions, from which the user could choose from, we considered the work by Scherer (2005). Scherer (2005) suggested categorizing emotions according to their meaning in the dimensions of valence, control, and arousal. However, instead of allowing the user to provide a value for each of these dimensions, Scherer provided a set of emotions grouped within families. We saw this approach as an appropriate fit to our proposal of an initial set of emotions since users would not have to categorize what they were feeling in different dimensions and also because Scherer (2005) carried out a study selecting specific emotions that were more common to regular people.

To suggest an initial set of pragmatic problems from which users could choose from to explain what caused negative emotions, we decided to start off with the design suggestions by Neil (2014) and the Heuristic Evaluation by Nielsen (1994). Neil (2014) describes design problems that are specific to mobile applications and suggests means for addressing them, while the Heuristic Evaluation (Nielsen, 1994) provides generic attributes for software applications in order to be usable. We chose to use the work by Neil (2014) as a basis, since the author gathered specific problems that users faced when carrying out tasks in the context of mobile applications, explaining the cause of the problem on real applications and how it affected their use. Additionally, each of the listed problems is accompanied by design suggestions on how to correct it, based on applications that have been accepted by end users in which the suggestion was applied, and previous experience of software development projects. Although some of the problems described by Neil (2014) may occur in applications that are not necessarily mobile, the way in which these problems are handled consider screen size and the mobile context, which may be useful for software development teams.

To complement our initial set of problems, we also consider the generic usability rules from the Heuristic Evaluation (Nielsen, 1994). The Heuristic Evaluation was considered as a basis since it has been widely adopted due to its positive results (Squires and Preece, 1999; Hvannberg et al., 2007; Hearst et al., 2016). Also, since it can be applied to different application contexts (mobile, web, desktop, others), we could use it to suggest further problems to complement the original list derived from the work by Neil (2014). Once we identified which problems could be faced by users, our idea was to rewrite these set of problems and simplify their description, developing items describing features and problems of mobile applications that could be easily understood

and chosen by users to report their experience. As a result of these aspects, we developed the following organization for the technique:

1) **Guiding Question:** A guiding question indicates the user what is being evaluated and organizes the evaluation in stages.

2) **Emotional Response:** An emotional response indicates an answer to a guiding question, and allows the users to report the hedonic aspect of their UX. These emotional responses were based on the items proposed by Scherer (2005).

3) **UX Problems:** A UX problem reports a pragmatic aspect that caused a negative emotional response. Thus, they indicate the aspects reported by Neil (2014) and Nielsen (1994).

These guiding questions, emotional responses, and UX problems are grouped in a logical order to facilitate the evaluation process. This decision was made to meet feature **FI (To Facilitate Feedback Giving)**. Therefore, we organized the evaluation according to the components of mobile applications that could be experienced by the users. The following categories of components were identified from the design suggestions by Neil (2014):

1) **Tutorial:** The user perceives it through the presentation of a tutorial or introduction or may start the application without an introduction.

2) **User Identification:** The user perceives it through the user registration and user login.

3) **Generic Functionalities:** The user perceives it through navigating within the application, carrying out searches and viewing information (text, figures, graphs, and tables).

4) **Specific Functionalities:** The user perceives it through the employment of specific functionalities of the application to achieve goals.

5) **Feedback:** The user perceives it through the information regarding successfully performed tasks, information regarding unsuccessfully performed tasks, and if the application does not provide feedback.

6) **Help:** The user perceives it by accessing system help or through the lack of system help.

The complete set of questions, emotional responses, and UX problems can be found in Appendix C. Also, the order in which the questions must be asked according to the aspects that are being evaluated can be found in Appendix D. Interested readers can refer to these documents in order to apply the technique.

4.3 The REX Evaluation Tool

One of the problems with UX evaluation instruments is that, in most cases, they are not interactive. Mostly, UX evaluation methods provide forms which the users have to fill in to report their experience (Vermeeren et al., 2010). To help improve this feature, we have developed a REX Evaluation Tool as a mobile application with an engaging design. Also, the results from other empirical studies involving evaluation methods showed that evaluation tools that automate part of the evaluation process could improve the acceptance of the proposed method (Rivero et al., 2014; Rivero et al., 2015).

In order to avoid bias by introducing the role of a moderator into the evaluation process, but still being able to guide users, we decided to create a moderator avatar, also called REX which would be embedded into the evaluation tool. The REX avatar has two main responsibilities: (1) to explain what is asked of the user at every stage of the evaluation process, and (2) to answer questions about any of the stages if the user requires it.

Figure 4-1 shows some of the expressions from the REX moderator avatar (see Introduction and Answering Users' Questions) and some screens from the REX Evaluation Tool. The appearance of the REX avatar was obtained using the free resources of Animaker[4], a Website that provides online tools to create videos and animations for different purposes. The avatar changes his expression according to the stage of the evaluation process. For instance, the avatar is presenting himself using common terms to users instead of technical language, in order to facilitate understanding what is asked of the user at each step of the evaluation. Also, Rex provides hints at each step of the evaluation, such as explaining how the user's personal

[4]http://www.animaker.com/

data (e.g., name, age, experience) will be handled, informing that the data will only be used for trying to improve the evaluated mobile application.

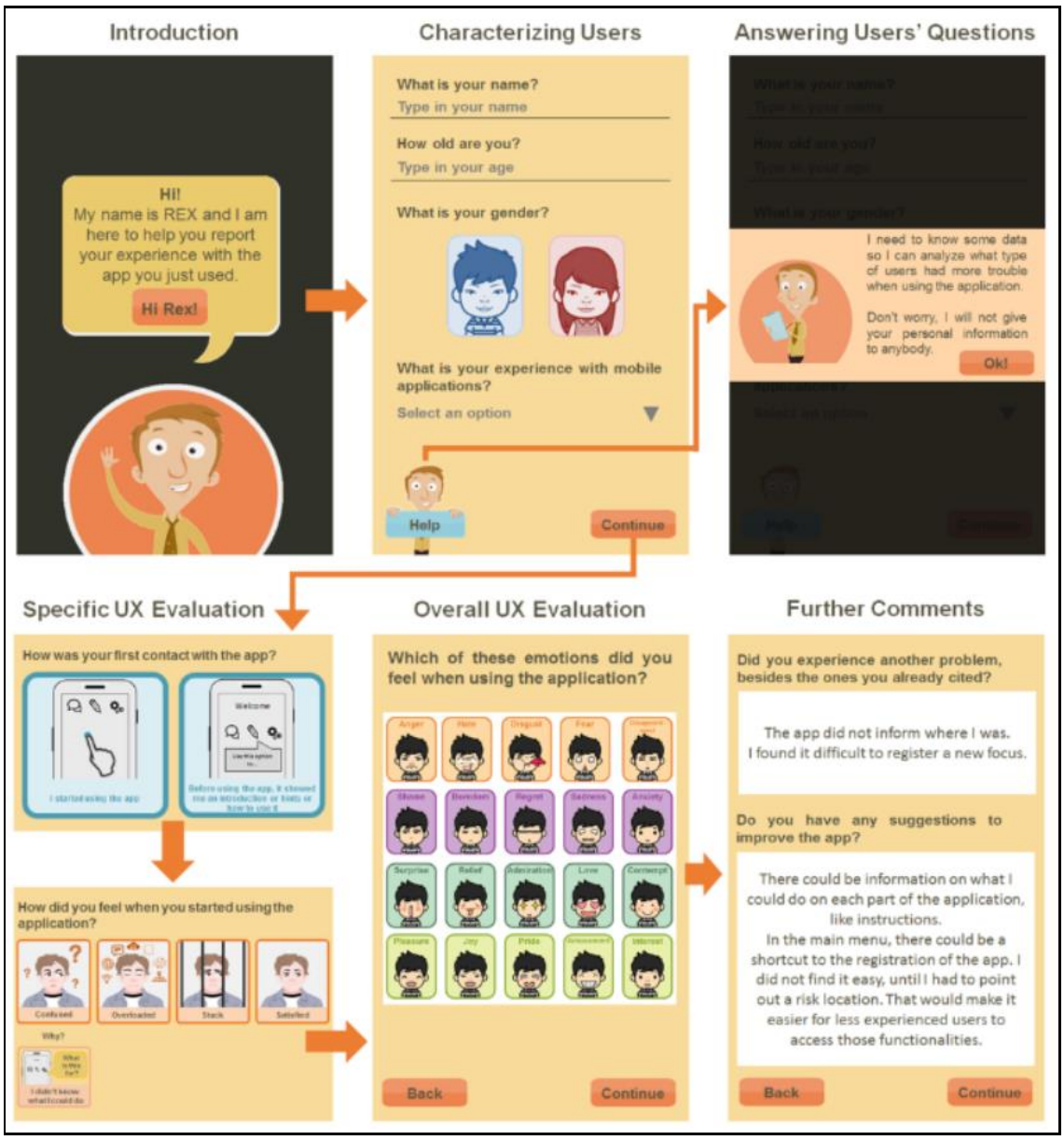


Figure 4-1 Screens from the REX Evaluation Tool

The choice of using an avatar was made to meet feature **FII (To Reduce the Influence of a Moderator)**, as this can cause discomfort to users during the UX evaluation (Rivero and Conte, 2015). However, although the use of an avatar can facilitate the interaction in several contexts and provide guidance (Lin et al., 2004; Chen et al., 2008), we do not claim that such choice would be better than other alternatives. We saw it suitable for the context of representing the moderator while guiding users through the evaluation process.

Furthermore, to meet feature **FIII (To Make the Evaluation Pleasant)**, we included animations between one screen and another, and the colors, text font and images of the application were chosen given the advice of designers and following examples of interfaces in games. The purpose of changing the appearance of the REX application was for it to be perceived as informal, so users would not feel that participating in the evaluation was a forced task. Additionally, the questions are presented according to what the user answers in previous stages of the evaluation, making the app react to the user's choices.

4.4 The REX Architecture and Evaluation/Redesign Procedure

Figure 4-2 shows the architecture from the REX approach, considering the technique and its evaluation tool. In this figure, the REX technique is composed of questions, emotional responses, and items. These are incorporated into the tool, which automates them through the moderator avatar and guides the user through the evaluation. Based on the individual evaluations, a technical report to support the redesign process can be created. The procedure for carrying out an evaluation using the REX approach will be explained below.

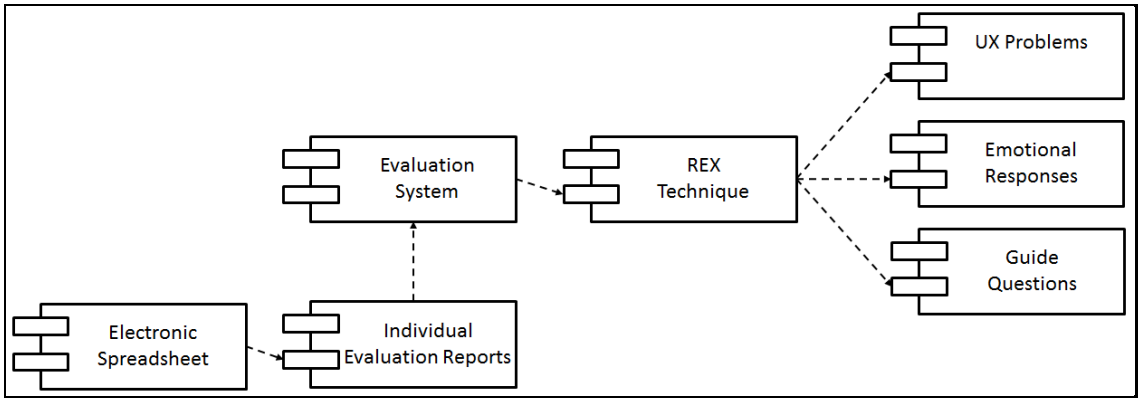


Figure 4-2 The architecture from the REX approach

Figure 4-1 (see Specific UX Evaluation) shows an example of how the items from the REX approach are organized and presented to the user as she progresses in the evaluation. In this example, the user is presented with a question regarding the "Tutorial" component asking how she first experienced the application. For this question, the user can indicate if: (1) (s)he started using the application without assistance; or (2) a tutorial was presented. Once the user selects an option, the avatar asks questions on how the user felt regarding that component of the application, and for

each possible emotion, it provides reasons for explaining them through the items reporting problems that we developed. The components, emotions, and items for explaining them are presented gradually to the user as she progresses in the evaluation. If the user has not experienced or seen the suggested component, she leaves it unmarked, and the avatar will continue asking questions regarding the rest of the components without wasting the users' time (as it will not be necessary to evaluate unrelated aspects).

In that context, we have specified three main stages to carry out a UX evaluation:

1) **Presentation:** In this stage, Rex is presented to the user, while it also asks simple demographic questions to tailor the rest of the evaluation (e.g., change the emotion cards according to the user's sex).

2) **Reporting of UX problems according to specific features of the application:** During this stage, the user reports on his specific emotional response and the different problems she faced while using the application. In that context, the user will be guided by means of features of the application, questions, experienced emotions and items to explain these emotions.

3) **Overall Emotional Evaluation:** During this stage, the user summarizes her experience in terms of emotions (see all possible emotions in Figure 4-1 – Overall UX Evaluation). Also, if she believes something has not been reported, she can use specific fields to provide a description of UX problems and/or suggestions for improving the application according to her needs (see an example of comments given by a user in Figure 4-1 – Further Comments). We chose to include this option to avoid limiting the user feedback to our initial set of problems and/or emotions.

Once a user finishes evaluating an application through REX, we generate a report containing the user's selection of emotions and items towards the evaluated application. Also, this report contains the user's demographic information (age, gender, experience with mobile technology, others) and generic opinion and improvement suggestions regarding the evaluated application. Software engineers can use this report as a basis for identifying improvement opportunities. The report is created based on a spreadsheet which links each of the identified problems to an appropriate design

suggestion. This was done in order to meet feature **FIV (To Provide Suggestions to Solve the Identified Problems).**

The report indicates the set of emotions and items that were selected by the users regarding a component from the evaluated application, and the corresponding percentage of users who indicated them. Additionally, each of the items indicating a UX problem is paired with an improvement suggestion based on the design suggestions by Neil (2014) and the generic usability rules from the Heuristic Evaluation (Nielsen, 1994). The software development team can analyze this report to modify the user interface in order to improve the UX of the evaluated application.

Table 4-1 shows an extract from the redesign report considering the feedback from several users. In this report, users' feedback is organized according to the component from the application they were evaluating, the responses that such component evoked, and the reasons that caused such responses. For each emotional response and reason, there is a number indicating the percentage of users who felt that way or experienced that problem. The development team can use this information to make changes in the application. Additionally, the report can contain further problems and improvement suggestions proposed by the users themselves during the overall emotional evaluation.

Considering the evaluation and redesigned process described above, Figure 4-3 shows an activity diagram of the evaluation and redesign process of the REX approach, considering: (a) the moderator, which can be a member of the development team deciding what to correct in the application; (b) the evaluator, which can be a user testing the mobile application; and (c) the REX evaluation tool, which automates the guiding process of the REX technique. In the following subsection, we better explain how the process works through an example.

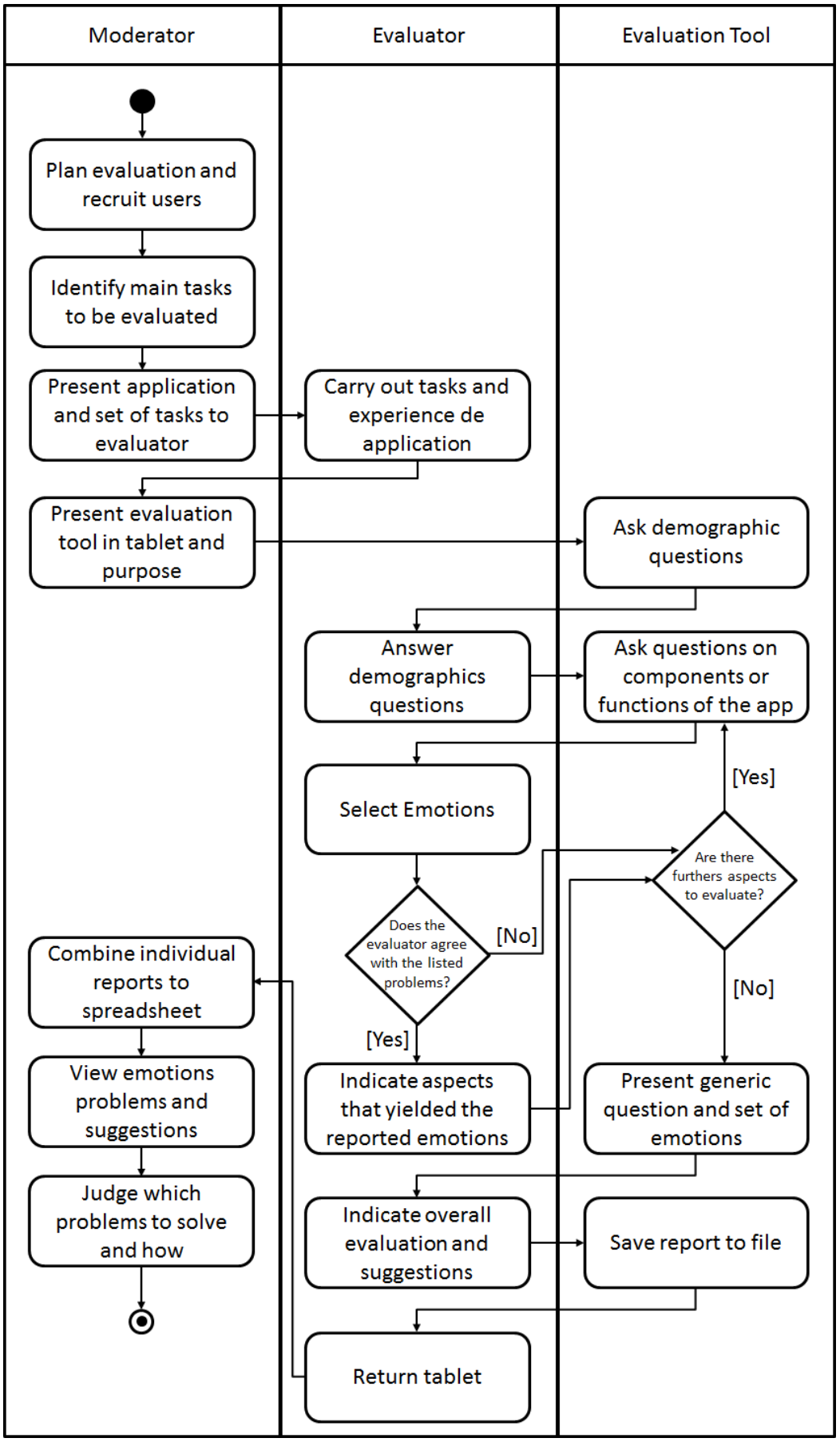


Figure 4-3 The evaluation and redesign process from the REX approach

Table 4-1 Extraction from a UX report generated by the REX approach

Component/Emotion/Item	Percentage	Improvement Suggestion
Tutorial		
Confused	45%	
(I1) The goal of the application is not clear.	17,5%	Apply the "Tutorials and Invitations" design suggestions from Neil (2014).
Navigation		
Confused	22,50%	
(I2) The user does not understand the current state of the application, once (s)he reaches a new screen.	12,5%	Include "Visibility of system status" from Nielsen (1994).
Overloaded	10%	
(I3) The user needs to access the menu all the time and to do so, (s)he has to activate it.	5%	Apply the "Persistent Navigation" design suggestions from Neil (2014).

4.5 An Application Example of the REX Approach

To illustrate how REX could be used for both evaluating the UX of a mobile application and suggesting improvement opportunities, we will show its application process through an example. In this example, we will evaluate and redesign an application under development called "ZikaZero", which aims to reduce the spread of the Zika virus. The Zika virus has spread throughout Brazil, becoming a risk to the health of the population, causing pain and mental and physical exhaustion (Campos et al., 2015). ZikaZero intends to allow people to indicate the locations in which the mosquito that spreads the disease can breed to the government authorities. By indicating these locations and providing information to its users on how to avoid the disease, the ZikaZero application can help reduce the occurrences of the virus and its spreading.

Figure 4-4 shows some screens from the ZikaZero application in its original version. In this figure, one can see the splash screen of the application (see Part A), the main screen showing some navigation options (see Part B) and the posts in which users indicated locations where the virus could spread (see Part C). The part of the report shown previously in Table 4-1 is also part of the report showing the results from

evaluating the ZikaZero application using REX. In this sense, users felt confused when first opening the app and then when navigating through it. For instance, no information was provided on what the purpose of the app was (see Item I1). Additionally, users felt confused when navigating through the application due to the lack of information on the current state of the application (see Item I2). Finally, as the users had to access the menu all the time and they had to switch among screens, the users felt overloaded with information (see Item I3).

Figure 4-4 Screens from the original version of the ZikaZero app

To deal with these problems, the development team applied the suggestions from the report in Table 4-1 for each of the identified problems. In this sense, Figure 4-5 shows screens of the redesigned version of the ZikaZero application. To deal with the lack of information for first-time users (see Table 4-1 Item I1), a tutorial explaining the purpose of the application and what users could do with it was provided when first opening the application (see Figure 4-5 Part A). Additionally, as users were not aware of their location in the application (see Table 4-1 Item I2), for each screen, instead of indicating the name of the app in the top bar of the screen, the development team included the name of the current screen in which the user is in (see Figure 4-5 Part A). Finally, to avoid having to go to a totally different screen to navigate through the application (see Table 4-1 Item I3), the development team suggested including a menu

that would show on the side of the screen if the user activates it, still remaining in the current task (see Figure 4-5 Part B).

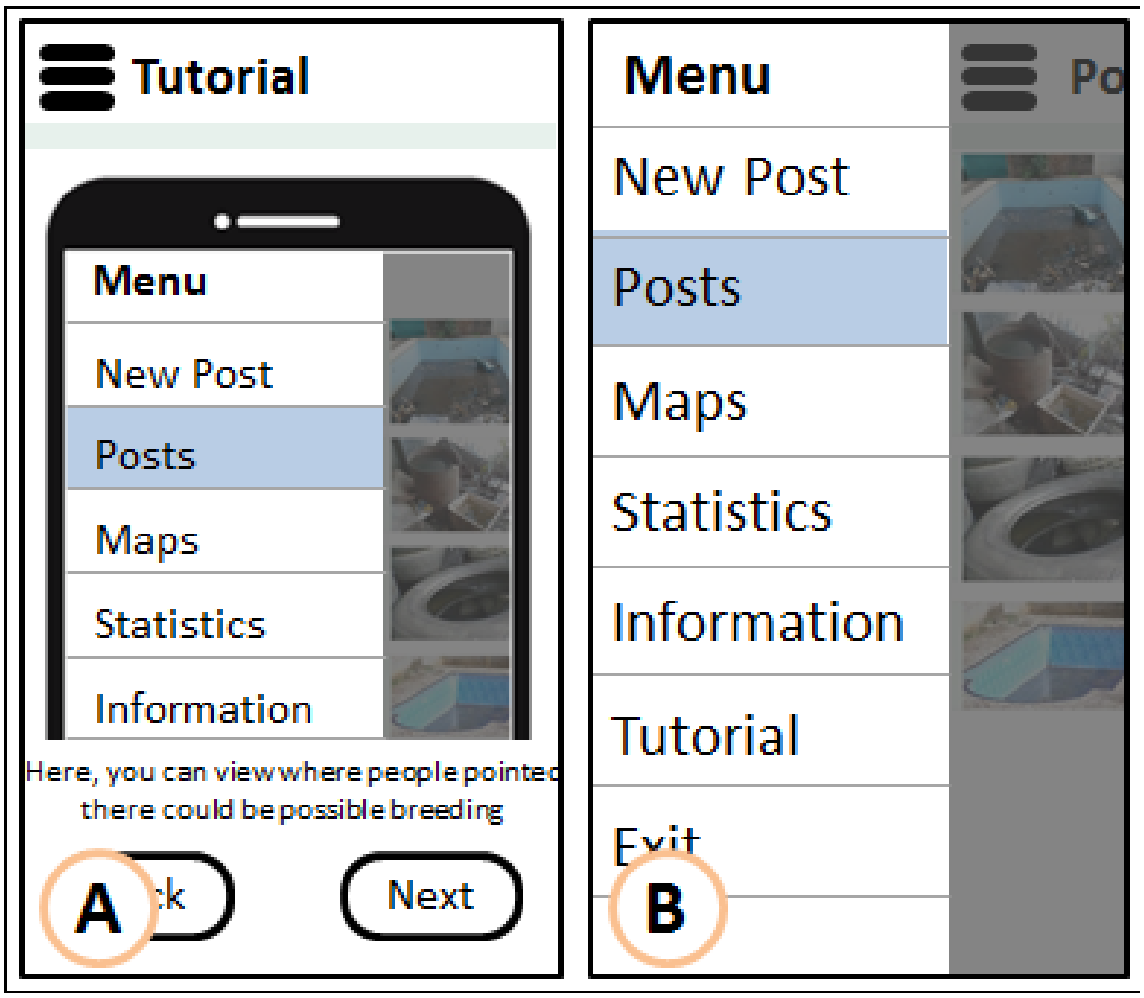


Figure 4-5 Screens from the redesigned version of the ZikaZero app

When applying the REX approach, the development team can prioritize the correction of the identified UX problems based on the percentage of users who felt a specific emotion and/or marked an item as a reason for having a negative UX towards the evaluated application. In this example, we have shown three problems that were corrected by the development team.

4.6 Summary

This chapter discussed the results from our reviews and empirical studies on UX evaluation methods and what can be improved in the evaluation and redesigned of software applications aiming at achieving a positive UX. Such analysis aimed at proposing a set of features to support future research in UX evaluation. As a result, we identified four main features: (FI) Facilitating Feedback Giving, (FII) Reducing the Influence of a Moderator, (FIII) Making the Evaluation Pleasant and (FIV) Providing Suggestions to Solve the Identified Problems.

The features defined above have been used to propose the REX approach which guides users through the evaluation process and suggests possible solutions for the identified UX evaluation problems. Furthermore, REX focuses on the evaluation of mobile applications as there are few methods specifically developed for this purpose and due to the rising popularity of this type of applications.

Redesigning for EXperience (REX) is an approach that aims to support users and software engineers in the UX evaluation and redesign of mobile applications respectively. We chose to focus on mobile applications due to their increasing number and importance in the past few years (Sarwar and Soomro, 2013), and the still shortage of UX methods that are specific to this category of applications (based on the results of our review). After trying to motivate and guide users during the reporting of their experiences, REX also provides a process and suggestions for software engineers to deal with the identified UX problems.

Finally, we highlight that REX is not a new approach for the UX evaluation of mobile applications, but an adaptation of existing approaches cited in Chapter 2. In essence, the REX approach is a “Form” type method which also provides an evaluation tool. Additionally, it uses probes (in our case, cards of emotions and items) to guide the user through the evaluation. In this sense, it allows identifying both quantitative and qualitative data, while providing improvement suggestions during the evaluation. Furthermore, it must be applied by users in a controlled environment after a momentary experience of a product or prototype (which can be functional or show the interaction with the system). In the next chapters, we will describe empirical studies aiming at

evaluating the feasibility of the REX approach and the opinion of users and software engineers on employing it.

5 The Initial Evaluations and Improvements of REX

This chapter presents the assessment of the acceptance of the REX approach from the point of view of users and software engineers in two studies. When compared to 3E, a qualitative UX evaluation method, the results show that REX was perceived as more fun, useful and more interactive. Additionally, software engineers considered REX useful and easy to understand, while suggesting providing graphs and visual highlights to facilitate interpreting the report. The improvement opportunities identified in the REX approach have been incorporated to REX generating its second version.

5.1 Introduction

We proposed a new UX evaluation method called Redesigning for EXperience (REX). In this chapter, we carried out two studies to verify the feasibility of REX as an approach for identifying UX problems and suggesting improvement opportunities in a mobile application. In Study 1, we compared the results found with REX with the results from 3E and evaluated the acceptance of the REX approach from the point of view of users. In Study 2, we analyzed the acceptance of REX by software engineers, both novice, and senior, when employing the report to improve the quality of an application in terms of UX. By presenting the results of these evaluations, we intend to encourage the adoption of the REX approach and improve the experience portrayed by the evaluated and redesigned applications.

The remainder of this chapter is organized into five more sections. In Section 5.2, we describe the execution and results of the empirical study from the point of view of users. Then, in Section 5.3, we describe the execution and results of the empirical study from the point of view of software engineers. Section 5.4 discusses our findings and lessons learned from the empirical studies in order to suggest improvements in the

REX approach, while Section 5.5 shows its new version. Finally, Section 5.6 concludes this chapter.

5.2 Study 1 - Evaluation from the Point of View of Users

In Study 1, we evaluated the feasibility of REX for the evaluation of a mobile application from the point of view of users. Therefore, we assessed the perceived usefulness and perceived ease of use of applying REX to report the emotions portrayed by an application and the causes for those emotions. We chose to evaluate the ZikaZero app (which is the same application used for showing how to use REX in Subsection 4.5) since it was an application being developed at the time, and the development team was interested in increasing its quality in terms of UX.

In addition to the use of REX, we applied the 3E method by Tähti and Niemelä (2006) as it is a cost-effective method that is also easy to use for users to report their experiences with mobile applications. Although REX and 3E are two different approaches, we wanted to compare the results in terms of time spent, a number of identified problems, false positives, and suggestions, in order to see how REX provides information on users' UX and if it is perceived as engaging and easy to use. Additionally, based on empirical evaluations (Isomursu et al., 2007), 3E had properties that match the contexts in which REX is supposed to be used. For instance, it does not require the presence of the moderator; it also allows gathering qualitative data on emotional responses from users and their causes; and users are free to provide suggestions or make drawings on aspects that negatively affected their experience.

5.2.1 Method

At all, 84 students from a local university in the city of Manaus (Brazil) were recruited for the study as part of the practical activities to be performed in class. The students were enrolled in classes related to computer science such as software quality, software analysis, and others. Before the study, all subjects who agreed to participate filled out a consent form and a characterization form. The consent form requested the participants' agreement in participating in the study and explained the main activities of the study and the confidentiality of the results. Additionally, the characterization form aimed at determining if the subjects met the expected user profile (if they were aware of the zika virus and its threats). Also, as the subjects could have background on subjects related to

Human Computer Interaction or Software Development in general, we categorized them according to the number of years in which they had worked in such activities. The categorization was performed as follows: (a) None: No experience; (b) Low: less than 1 year and with knowledge based on books and classes; (c) Medium: 1 to 3 years of practical experience; and (d) High: more than 3 years of practical experience. Around 50% of the subjects had none or low experience, while the rest had at least medium experience in one of the two mentioned topics.

All evaluations followed the same procedure shown in Figure 5-1. Before the study, the subjects were characterized. Then, they were scheduled for participation, and in their participation day, they entered a lab room where a moderator greeted them to provide the study materials. After explaining the purpose of the ZikaZero mobile application, the subjects received a mobile phone with an embedded navigable PDF file. Such file contained the screens of the ZikaZero application which had been mapped with links to simulate the main interactions with it. Users could tap on the PDF file, and they would be directed to the according screen to carry out tasks such as posting a potential risk location where the mosquito could breed and finding information on how to avoid being contaminated with the virus. After interacting with the application, the subjects were presented with one of the two UX evaluation approaches (either REX or 3E) and were asked to use them to report their experiences. As mentioned before, 3E is a UX method for evaluating mobile applications without the interference of a moderator. In that context, note that each subject employed only one of the techniques. To avoid having more experienced subjects in one group using REX than the other one using 3E, we balanced them according to the experience of the subjects in the fields of Human Computer Interaction and Software Development following their categorization. During the UX evaluation, the moderator of the evaluation and an assistant were responsible for counting the time spent per user to finish reporting his/her experience.

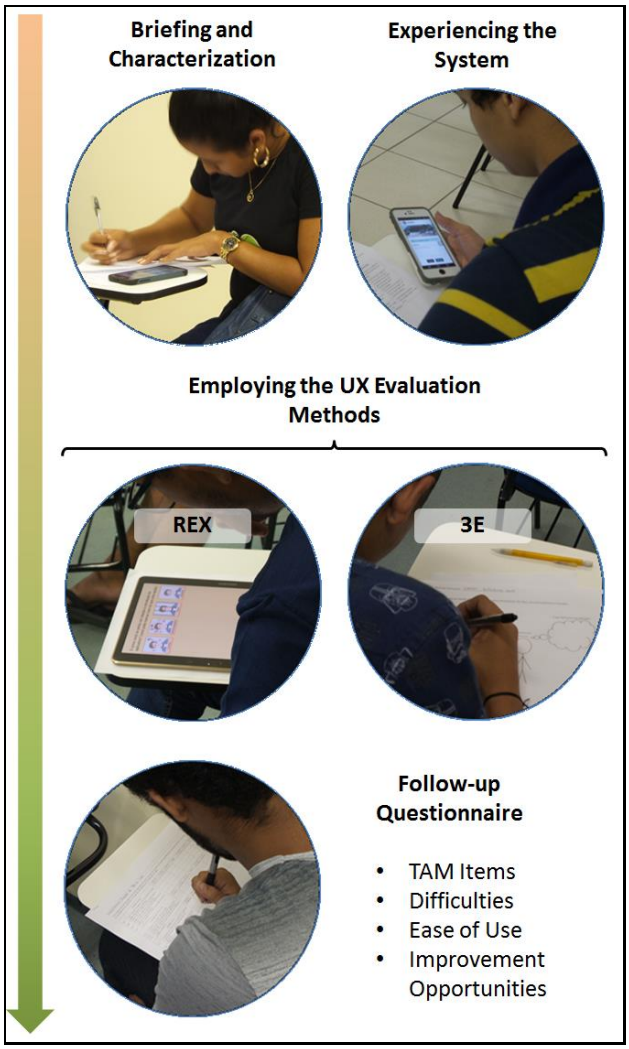


Figure 5-1 Steps for the execution of Study 1

Based on each UX report (from REX or 3E), we collected the features the users reported that negatively affected their UX or improvement suggestions they thought could be implemented in the application. For the 3E method, we extracted the information from the written passages within the filled 3E form. To extract the features from the REX reports, we used the REX spreadsheet (see Subsection 4.5), which was specifically tailored for viewing the collected UX data (negative aspects and suggestions). Such spreadsheet was created to integrate the data collected from each user applying the REX approach and allows to combine the individual REX reports generated after an evaluation. From the 84 subjects, 2 subjects that would employ REX dropped from the study for personal reasons. Therefore, a total of 82 UX reports were delivered (40 automatically generated reports from the REX approach and 42 filled forms from the 3E method). The results of the evaluation were analyzed in the following steps:

1) **Collection:** Based on each UX report (both REX and 3E), we collected the features that the users reported negatively affected their UX or improvement suggestions that they thought could be implemented in the application. These aspects and suggestions were incorporated into a single report.

2) **Removal of Duplicated Problems:** We highlighted duplicated features (negative aspects or suggestions) that were pointed out by more than one user. Then, we generated a new complete report which contained all negative features and improvement suggestions found without showing the duplicated ones. Note that for each duplicated aspects/suggestion the better-described item was retained.

3) **Discrimination:** We carried out a meeting with three UX specialists to classify the negative aspects and suggestions as appropriate to the development of the ZikaZero application. To avoid classification bias, the three UX specialists were researchers outside the study who had more than 4 years of experience in carrying out usability and user experience evaluations, while also having high knowledge on interaction modeling, and design methods. These experts reviewed the list of negative aspects and classified them as problems or false positives. Note that by false positives, we refer to issues that were pointed out by the users, but were judged by the specialists as attributes that were not related to the use of the application, or were not within the scope of the developed application (e.g., the user indicated that there was a problem but did not describe it, the user complained about the mockups, others). Thus, when discussing a negative aspect and it was not deemed to be a problem, such aspect was classified as a false positive.

The number of negative aspects pointed by the users as well as the improvement suggestions and time spent in the UX evaluation has been employed to compare the performance of the UX techniques.

To gather data for evaluating the acceptance of REX, we applied a questionnaire based on the indicators from the Technology Acceptance Model 3 (TAM 3). TAM 3 (Venkatesh and Bala, 2008) aims at assessing the users' beliefs about the usefulness, ease of use and intention of using a technology that is expected to support them in

performing a task. According to the authors of the TAM model, perceived usefulness defines the degree to which a person believes that a technology could improve his/her performance at work; perceived ease of use defines the degree to which a person believes that using a specific technology would be effortless; and intention to use defines the degree to which the user believes that (s)he would adopt a technology. The reason for focusing on these indicators is that these aspects are strongly correlated to user acceptance of a given technology (Venkatesh and Bala, 2008). Considering that REX is a new approach to evaluate UX, it is interesting to evaluate its acceptance from the point of view of users.

Table 5-1 shows the questionnaire we applied for evaluating perceived usefulness, perceived ease of use and intention to use regarding REX and 3E. In order to apply the questionnaire, we:

1) Replaced the investigated "technology" in the questionnaire with the terms "REX" or "3E" according to the technology we were evaluating.
2) Replaced the process investigated in the questionnaire with "reporting of the UX" with a focus on mobile applications.
3) Employed a seven-point scale asking for the degree of agreement with the statements from the point of view of users as suggested by Venkatesh and Bala (2008): (1) strongly disagree, (2) moderately disagree, (3) somewhat disagree, (4) neutral, (5) somewhat agree, (6) moderately agree and (7) strongly agree.
4) After the statements within the questionnaire, we included open questions, to better understand the features that made the REX or 3E useful (or useless), easy (or difficult) to use and suitable (unsuitable) for reporting the users' experience.

After each subject had delivered his/her questionnaire, we counted the degrees of agreement per each of the items from the TAM questionnaire in order to provide an overview of the acceptance of the REX and 3E approaches. Additionally, for the REX approach, we analyzed the data obtained from the open questions investigating the aspects that affected its usage and acceptance by applying qualitative analysis procedures (Glaser and Strauss 2009). While we analyzed the data contained within the questionnaires, we created codes related to the citations within them. These codes were reviewed by an experienced researcher in qualitative analysis, who verified them and the categories in order to audit the coding process. After that, we created the

relationship codes. Consequently, we identified codes related to the difficulties and facilitators of the REX approach, the reasons that would make users apply it again for performing a UX evaluation, and its improvement opportunities.

Table 5-1 Applied TAM adapted questionnaire for evaluating the applied UX technologies

Statements regarding "Perceived Usefulness" (PU):	
PU1	Using the "technology" improves my performance in reporting my experience with a mobile application.
PU2	Using the "technology" in reporting my experience with a mobile application increases my productivity.
PU3	Using the "technology" enhances my effectiveness in reporting my experience with a mobile application.
PU4	I find the "technology" to be useful in reporting my experience with a mobile application.
Statements regarding perceived "Ease of Use" (EoU):	
EoU1	My interaction with the "technology" is clear and understandable when reporting my experience with a mobile application.
EoU2	Interacting with the "technology" does not require a lot of my mental effort when reporting my experience with a mobile application.
EoU3	I find the "technology" to be easy to use when reporting my experience with a mobile application.
EoU4	I find it easy to get the "technology" to do what I want it to do when reporting my experience with a mobile application.
Statements regarding "Behavioral Intention" (BI):	
BI1	Assuming I had access to the "technology," I intend to use it.
BI2	Given that I had access to the "technology," I predict that I would use it.
BI3	I plan to use the "technology" in the next <n> months.

5.2.2 Results

Figure 5-2 shows a form that was filled by one of the subjects who employed the 3E method to report his experience. In this figure, the user reports an overall positive experience. Although he reports having issues with understanding the application (mainly, in the menu), he indicates that the idea of the application is good and that it could be very useful. Furthermore, the user stated that he got lost when registering in order to use the services of the application. These problems were also stated in the REX

reports. An extract of the complete report of the negative emotions that the users reported was shown in Table 4-1 along with the features that caused those emotions.

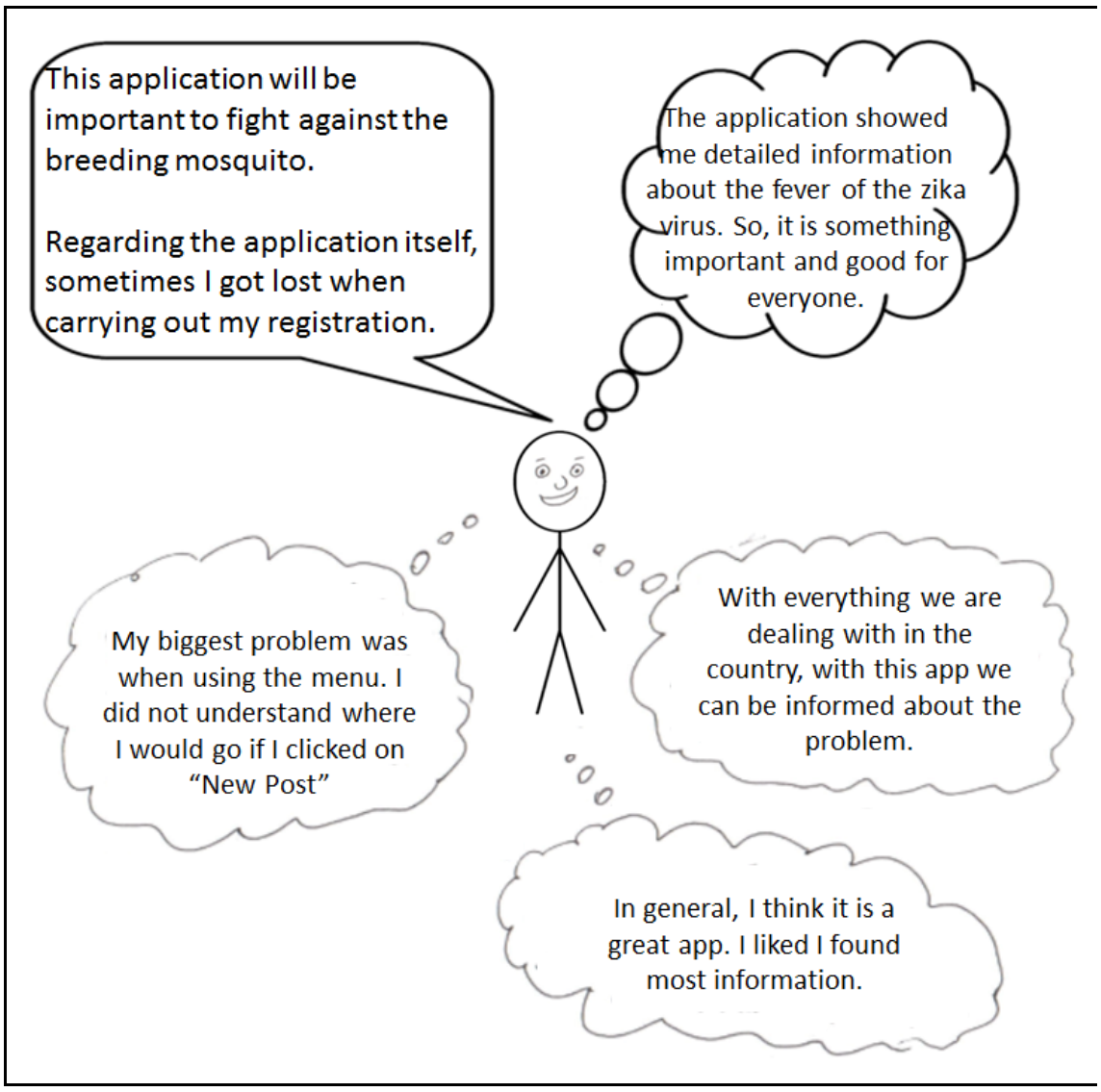


Figure 5-2 3E form filled in by one of the subjects in the study

To analyze the performance of each technique, we considered the following indicators: (a) the time spent for a user to report his/her experience; (b) the number of UX problems a user reported using the technique; (c) the number of false positives a user reported using the technique (notes that the user pointed as a problem but were not real pragmatic problems, e.g., part of the application that is not shown because it is still under development); and (d) the number of suggestions a user reported using the technique.

Table 5-2 shows the descriptive statistics from the results for each indicator per technique. We discarded 3 outliers from the REX approach and 8 outliers from the 3E approach. These outliers were considered due to their results in terms of identified UX problems, false positives and suggestions. In case a subject did not identify any UX problem, suggestion or problem that turned out to be a false positive (i.e., (s)he did not try to indicate a problem in the application or ways to improve it), (s)he was considered an outlier. Figure 5-3 and Figure 5-4 show the boxplot graphs with the distribution in terms of time spent per technique and overall results, respectively.

Table 5-2 Descriptive statistics of the quantitative results regarding time spent, total identified problems, false positives and suggestions per technique

		3E	REX
Time Spent	**Min**	5	6
	Max	23	31
	Mean	12,9	13,1
	Median	12	13
	Std. Dev.	5,1	4,7
Total Problems	**Min**	0	0
	Max	5	14
	Mean	1,0	2,9
	Median	1	3
	Std. Dev.	1,2	2,8
Total False Positives	**Min**	0	0
	Max	0	4
	Mean	0,1	0,4
	Median	0	0
	Std. Dev.	0,3	0,8
Total Suggestions	**Min**	0	0
	Max	9	5
	Mean	1,1	1,3
	Median	1	1
	Std. Dev.	1,5	1,0
Overall Results	**Min**	1	1
	Max	9	37
	Mean	2,2	7,6
	Median	1	6
	Std. Dev.	1,9	6,9

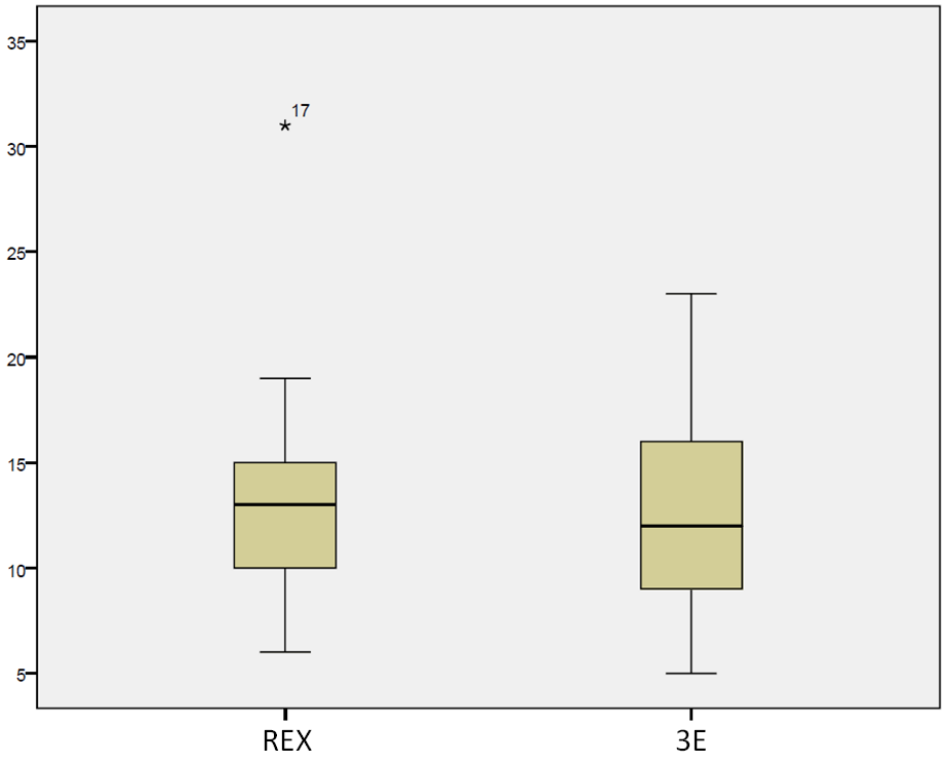


Figure 5-3 Boxplot graphs comparing the time spent per technique

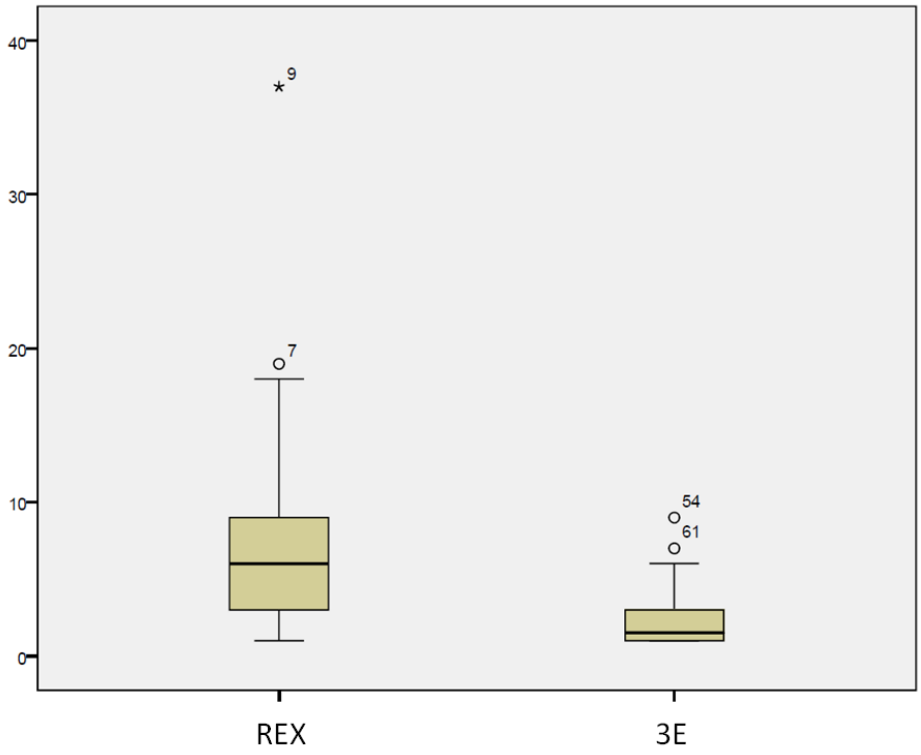


Figure 5-4 Boxplot graphs comparing the overall results per technique combining: overall identified problems, false positives and improvement suggestions

To compare the techniques, we applied the Mann-Whitney non-parametrical statistic method (Birnbaum, 1956). The choice for such test was due to the non-normality of our samples (in all cases, the Shapiro-Wilk showed $p < 0,05$) (Shapiro and Wilk, 1965) and because we also applied a between-subjects study design. The results from the statistical analysis ($\alpha = 0,05$) suggest that overall (when considering UX problems, false positives, and suggestions) there was a statistical difference ($p = 0,000$).

Furthermore, Table 5-3 shows the results of chi-square tests (Corder and Foreman, 2014) verifying the effect of the use of the techniques in the observed variables. The results suggest that when considering the UX problems found using the specific items provided by REX, using one technique or another may result in different amount of UX problems. A possible reason for this improvement may be due to the guidance offered by REX when a user is explaining the reasons for a poor UX.

Table 5-3 Summary of the statistical analysis using chi-square in each of the indicators

Indicator	P-Value	Conclusion
Time Spent	0,325	There is no significant association between the time spent and the applied technique.
Identified UX Problems	0,014	There is a significant association between the number of identified UX problems and the applied technique.
False Positives	0,227	There is no significant association between the number of false positives spent and the applied technique.
Suggestions	0,109	There is no significant association between the number of suggestions spent and the applied technique.

We also compared the emotional responses of the users regarding the evaluated application with both techniques. To analyze the reported emotions through REX, we counted how many times a user indicated (s)he felt positive or negative emotions towards the application when summarizing his/her experience at the end of the test (see Figure 4-1 - Overall Emotional Evaluation) and then, we verified the tendency of the experience as negative, neutral or positive. Regarding 3E, we also categorized the experiences as positive, negative or neutral. However, the analysis was performed over the faces that the users drew in their reports (see Figure 5-2 for an example of a 3E report), categorizing the faces. Table 5-4 shows the overall emotional responses of the

users towards the ZikaZero app. Here, we can see that users employing 3E reported having more positive experiences than users who applied REX. This can be explained by the focus of the technique. In this sense, as 3E is a reporting technique that suggests indicating what the user is thinking and what (s)he would say about the evaluated product, users are free to report what they want. On the other hand, REX makes users think of different aspects of the application, which can affect their perception of the application, making them be more critical of it. This is important, as being able to think of what is not working properly in the application may be useful for software development teams of mobile applications.

Table 5-4 Summary of the number of users per techniques who had negative, neutral or positive overall emotional responses

	REX	3E
Negative	6	1
Neutral	5	5
Positive	29	36

Regarding the reporting of the users' emotional responses through REX, we also compared if there were any inconsistencies between how users reported their experiences during the use of REX and at the end of the evaluation. Therefore, we counted the reported emotional responses of the users when evaluating each of the components of the application and we categorized them as negative, neutral and positive following the same procedures as before. Table 5-5 shows the summary of the consistency of users when reporting their experience gradually through the evaluation of the components of the application and after summarizing their emotional response at the end of the evaluation using REX. In this table, the numbers with a "=" symbol represent the number of users who, throughout the evaluation, indicated a type of emotional response and at the end of the evaluation summarized their emotional response in the same way (e.g., 22 users indicated that they had a positive experience with the application throughout the evaluation of its components and indicated the same at the end of the evaluation). Additionally, numbers with a "-" and a "+" indicate that the users reported inferior and superior overall emotional experiences respectively when describing their overall emotional response. When analyzing these numbers, we can see that only 26 out of 40 users (65%) maintained their detailed evaluation consistency with their overall evaluation using REX. This result suggests that although the overall

evaluation may reflect what users think, some aspects can be left behind if a UX evaluation method focuses on the overall evaluation and not the specifics.

Table 5-5 Analysis of the consistency of users applying REX for evaluating their experience with the ZikaZero app

		Overall Evaluation		
		Negative	**Neutral**	**Positive**
During the Evaluation	**Negative**	**4 (=)**	0 (+)	5 (+)
	Neutral	1 (-)	**0 (=)**	2 (+)
	Positive	1 (-)	5 (-)	**22 (=)**

Regarding the answers from the TAM-based questionnaire, Table 5-6 shows the descriptive statistics for the Perceived Usefulness statements (PU1 to PU4), Perceived Ease of Use statements (PEoU1 to PEoU4) and the Behavioral Intention statements (BI1 to BI3). The median for all answers in both techniques was equal or above 5 (agreement with the statements), which suggests that users thought that both techniques are useful, easy to use and that they would use them if given a chance. Nevertheless, REX achieved higher median agreement scores in terms of effectiveness for reporting an experience, required mental effort, intention to use and future usage.

Despite the cautiously positive results, some subjects were not convinced about the usefulness and ease of use of REX approach and suggested some improvements. To better understand the reasons that made the users answer positively or negatively towards the use of REX for reporting their experience, we have analyzed the answers to the open questions, looking for the justification of their answers. Below, we present our results, referring to the users as subjects using the code SXX, where XX is the number of the user from 1 to 40. Note that we only present qualitative data regarding the REX approach as we are evaluating its feasibility from the point of view of users.

Table 5-6 Median and standard deviation for perceived usefulness, perceived ease of use and behavioral intention to use indicators for the REX and 3E techniques in a UX evaluation

	REX		3E	
Item	**Median**	**Std. Dev.**	**Median**	**Std. Dev.**
Perceived Usefulness				
PU1 - Performance	6	1,30	6	0,99
PU2 - Productivity	6	1,37	6	1,20
PU3 - Effectiveness	**5 and 6**	**1,71**	**5**	**1,36**
PU4 - Useful	6	1,22	6	1,17
Perceived Ease of Use				
EoU1 - Clear	7	0,63	7	1,07
EoU2 – Mental Effort	**7**	**0,75**	**6**	**1,23**
EoU3 – Easy to use	7	0,52	7	0,98
EoU4 - Controllable	6	1,07	6	1,04
Behavioral Intention				
BI1 - Intention to use	**6**	**0,98**	**5 and 6**	**1,32**
BI2 - Prediction to use	6	0,90	6	1,23
BI3 - Usage	**6**	**1,43**	**5**	**1,62**

Regarding positive aspects of applying REX for evaluating the UX of a mobile application, we identified several features. For instance, some users mentioned that REX was intuitive, effortless, engaging, interactive, easy to use, not intrusive, dynamic, direct, friendly, and others (see quotes from subjects S03, S17, and S20). These results suggest that our initial requirements with REX were met. Additionally, several users stated that one of the main advantages of REX was providing a fixed set of emotions and items (see a quote from subject S05). Other users indicated that the images in the items and emotions made them easier to understand them (see a quote from subject S29). Furthermore, regarding the questions and guidance provided by REX, some users stated that REX made them think of features they could have easily forgotten during the evaluation (see quote from subject S10), while others indicated that the approach was useful for guiding the user due to its avatar (see quote from subject S19). These statements suggest that users felt guided during the evaluation process and that REX facilitated the evaluation by providing users with items with clear descriptions of what they could report regarding their experience.

"The REX environment was very interactive, and it presented a set of emotions from which I could choose from." Generic Positive Aspects - Subject S03.

"I felt at ease when using REX because it interacts with the user, and it makes questions and explains things during the evaluation." Generic Positive Aspects - Subject S17.

"The app has an interface that is super engaging and friendly, which allows me to express myself and my ideas." Generic Positive Aspects - Subject S20.

"The way of answering is easier because it uses predefined answers, which makes it faster to provide an answer." Items Positive Aspects - Subject S05.

"The images it provides are very interesting. It helps the user express himself as it is difficult to express that in words." Items Positive Aspects - Subject S29.

"It made me think of things I could have just left behind." Guidance Positive Aspects - Subject S10.

"I didn't have to think a lot because the questions were objective and clear. Also, it had an assistant guiding me whenever I started a new phase." Guidance Positive Aspects - Subject S10.

Regarding negative aspects of applying REX for evaluating the UX of a mobile application, we also found some features. In most cases, users also pointed out what could be done to mitigate these disadvantages. Therefore, we will also describe improvement suggestions in the current version of the REX approach. First of all, some users did not think that the guiding process was as useful as suggested. For instance, although REX provides means for describing emotions and aspects that were not dealt with during the evaluation at the end, some users felt that this forced them to wait in order to report their experience (see quotes from subjects S14 and S27). To mitigate this problem, some users suggested allowing explaining the selected emotions with other features (e.g., writing, speaking, others), even if these are not present at the moment of the evaluation (see a quote from subject S10). This would avoid forcing users to remember a specific aspect of their experience they were not able to report until the end of the evaluation. Furthermore, other subjects indicated that the explanation provided by REX was too focused on evaluated aspects and did not explain how the evaluation was

going to be performed (see a quote from subject S40). To this end, some users suggested explaining or providing a brief introduction on how the evaluation would be performed and how the user could report his/her experience (see quotes from subjects S05 and S40), indicating that if (s)he thought an item was not applicable, (s)he could leave it unmarked. Finally, some users indicated that the emotions sometimes did not match with what they were feeling (see a quote from subject S37) or that some items were too broad that they could not be useful for evaluating a specific application (see a quote from subject S10). Although REX has some components that are focused on the specific functionalities of the application, its items are broad so they could fit the evaluation of other applications. In this sense, it could be interesting to allow including specific evaluation items created by the development team. Moreover, regarding not finding a specific emotion for an evaluated component, we still need to evaluate to what degree is providing all possible emotions could increase the effort for providing user feedback.

> *"I could not express by writing when I wanted it..."* Experience Report Negative Aspects - Subject S14.

> *"I felt a bit restricted to what REX wanted me to evaluate..."* Experience Report Negative Aspects - Subject S27.

> *"It would be interesting to allow inputting specific observed features."* Experience Report Suggestion - Subject S10.

> *"There are some ramifications that do not make sense. I didn't know how to evaluate my experience."* Guidance Negative Aspect - Subject S40.

> *"... I would also explain that it is not mandatory to select an emotion or an item, you can just leave it unmarked."* Guidance Suggestion - Subject S05.

> *"I would explain that each emotion has a set of items and that each item could be used to report my experience."* Guidance Suggestion - Subject S40.

> *"It might be that none of the emotions were related to that component..."* Experience Report Negative Aspects - Subject S37.

"Some of the items are not specific for the evaluated application, making the evaluation evasive" Experience Report Negative Aspects - Subject S10.

5.2.3 Conclusions from Study 1

This study showed that REX was effective for identifying UX problems and improvement suggestions for mobile applications under development. The quantitative results suggest that the techniques (REX or 3E) had an effect over the number of identified UX problems while using the same amount of time. Also, REX specific items for evaluating the components from mobile applications provided further details on UX problems. Furthermore, the results from the TAM questionnaire were a positive indicator that REX could be as useful and easy to use as existing UX evaluation technologies such as 3E. Additionally, the qualitative results suggest that REX managed to meet its requirements, being perceived as a friendly and guiding method for reporting experiences. Nevertheless, REX still needs to improve in terms of guidance on how to select its items during an evaluation and allowing more freedom to users who want to report more features than the one it provides.

5.3 Study 2 - Evaluation from the Point of View of Software Engineers

In Study 2, we evaluated the feasibility of REX in the redesign of a mobile application from the point of view of software engineers. Therefore, we assessed the usefulness and ease of use of applying REX to correct UX problems presented in the ZikaZero app based on the suggestions proposed within the complete report created with the REX approach (see part of the report in the application example of REX in Table 4-1).

5.3.1 Method

At all, 31 students from a local university in the city of Manaus (Brazil) were recruited for the study by convenience and due to their knowledge. 16 out of 31 students were enrolled in the last semesters of the university in computer science. The rest of the students (15 subjects) were practitioners enrolled in a specialization course on software engineering, where new technologies for developing software were being introduced. Following the procedures from Study 1, all subjects who agreed to participate filled out a consent form and a characterization form. Around 50% of the subjects had none or low experience in software development, while the rest had at least medium experience

(median = 4 years of experience). By having novice and experienced software engineers, we wanted to collect feedback from both types of practitioners, aiming at supporting the majority of software engineers.

The redesign process was given as a graded task within the classes of novice software engineers and practitioners. In this sense, the subjects were divided into groups of three or four members. Each group received the report containing the description of problems and improvement suggestions as shown in Table 4-1. Additionally, they received a description of the step-by-step process for using the report in the redesign of an application, while having lectures explaining each part of the report. Finally, each team had to prepare a presentation where they would present the redesigned version of the ZikaZero application and which problems they have addressed, how and why.

To gather data for evaluating the acceptance of REX when employed for redesigning an application from the point of view of software engineers, again, we applied a questionnaire based on the indicators from the Technology Acceptance Model 3 (TAM 3). We applied the questionnaire from Study 1 (see Table 5-1) to evaluate perceived usefulness, perceived ease of use and intention to use regarding REX, however making the following changes:

1) Replaced the investigated “technology” term with “REX”.
2) Replaced the process investigated in the questionnaire with “redesign and application for improving its UX” with a focus on mobile applications.
3) Employed the same seven-point scale asking for the degree of agreement with the statements from the point of view of software engineers as suggested by Venkatesh and Bala (2008).
4) After the statements within the questionnaire, we included open questions, to better understand the features that made the REX report useful (or useless), easy (or difficult) to use and suitable (unsuitable) for redesigning an application.

After each subject had delivered his/her filled questionnaire, we counted the degrees of agreement per each of the items from the TAM questionnaire in order to provide an overview of the acceptance of the REX and 3E approaches. Additionally, we also analyzed the data obtained from the open questions investigating the aspects that

affected its usage and acceptance by applying the same qualitative analysis procedures as in Study 1.

5.3.2 Results

The software engineers were able to apply the suggestions from the REX approach in redesigning the ZikaZero application. An example of a redesigned version was shown in Figure 4-5, when showing how to use the redesign process of the REX approach. To verify if we could analyze the results from the TAM questionnaire regarding the perception of the REX approach from both students and software engineers as a single sample, we carried out a Mann-Whitney test (as again, the Shapiro-Wilk showed $p < 0,05$). The results for each factor (combining the results from each item) (DiStefano et al., 2009) showed $p > 0,05$, indicating that there was no significant difference in the students and software engineers' samples ($p = 0,769$ for Perceived Usefulness, $p = 0,399$ for Perceived Ease of Use and $p = 0,984$ for Intention to Use). Therefore, we have analyzed the combined results for the answers to the TAM questionnaire from software engineers and students in their last year of university. Table 5-7 shows the descriptive statistics for the Perceived Usefulness statements (PU1 to PU4), Perceived Ease of Use statements (PEoU1 to PEoU4) and the Behavioral Intention (BI1 to BI3). The median for all statements regarding the usefulness and intention to use was equal or above 5 (agreement with the statements), which suggests that software engineers thought that the report provided by REX was useful and that they would use it if given a chance. Nevertheless, the median of the degree of agreement in the perceived ease of use statements was lower (5). This suggests that although software engineers thought of the report provided by REX as easy to use, improvements should be implemented.

To better understand the reasons that made the software engineers answer positively or negatively towards the use of REX for redesigning an application, we have analyzed the answers to the open questions, looking for the justification of their answers. Below, we present our results, referring to the users as subjects using the code SXX, where XX is the number of the user from 1 to 31.

Table 5-7 Median and standard deviation for perceived usefulness, perceived ease of use and behavioral intention to use indicators for the REX approach in a UX redesign

	REX	
Item	**Median**	**Std. Dev.**
Perceived Usefulness		
PU1 – Performance	6	0,96
PU2 – Productivity	6	0,91
PU3 – Effectiveness	6	0,93
PU4 – Useful	6	0,82
Perceived Ease of Use		
PEoU1 – Clear	5	1,28
PEoU2 – Mental Effort	5	1,46
PEoU3 – Easy to use	5	1,29
PEoU4 – Controllable	5	1,21
Behavioral Intention to Use		
BI1 - Intention to use	6	1,11
BI2 - Prediction to use	6	1,19
BI3 – Usage	6	1,17

Regarding positive aspects of applying REX for redesigning a mobile application, we identified several features. For instance, software engineers indicated that the report was very organized, which facilitated understanding where to correct a specific problem (see a quote from Subject S04). Another positive aspect was that the REX report paired each description of a UX problem with a possible solution (see a quote from Subject S10). Software engineers indicated that the report made the redesign easier as it allowed finding the problems faster as it highlighted the problems (see a quote from Subject S02). Finally, some subjects indicated that the REX report allowed users to provide an overall description of features, which made it easier to understand the problems and prioritize them (see a quote from Subject S15).

> *"It was easy to understand the report as it had topics and subtopics."* Generic Positive Aspects - Subject S04.

> *"Besides showing the problems, it also presents solutions, which come in handy once we need to know what to modify in the user interface."* Generic Positive Aspects - Subject S10.

“I managed to better understand the report as it highlighted the problems with red...” Generic Positive Aspects - Subject S02.

“The descriptions at the end that were written by the users themselves helped us understand the problems and prioritize them” Generic Positive Aspects - Subject S15.

Regarding negative aspects and improvement opportunities for applying REX in the redesign of a mobile application, we found out that the report was perceived as complex (see a quote from Subject S18). Although we highlighted each component that was evaluated and paired it with the identified emotions and problems, we still included the information regarding all components on the same page, while still indicating features that were not marked by the users. Therefore, some software engineers suggested dividing the report into components or aspects (see a quote from Subject S12) and removing the aspects that were not indicated by the users during the evaluation (see a quote from Subject S14). Some novice software engineers had trouble applying the design suggestions that were suggested. The main problem with the redesign suggestions is that they offer examples, but the software development team must adapt them from scratch into their projects, designing screens for the specific suggestion (see a quote from Subject S04). Therefore, we intend to provide further support to redesign the application, by providing blank layouts that could be modified by the development team to meet their needs. Finally, some software engineers indicated that the report became difficult to employ due to the lack of support for finding specific items as there were many aspects being evaluated for a component, and that it was difficult to decide which problems to correct first (see a quote from Subject S03). Some software engineers who experienced this problem suggested providing a way to filter the items (see a quote from Subject S02), or ordering the items by the highest percentages of occurrences and per component where they occur (see a quote from Subjects S14 and S12).

“I think it is a bit complex because, in order for us to understand, it takes a little time.” Generic Negative Aspects - Subject S18.

“It would be better if it was organized in tables or pages, one per component, such as “forms” or something…” Organization Suggestion - Subject S12.

"The report is very complete, but it could omit the irrelevant items and emotions, those in which 0% of the users marked them." Organization Suggestion - Subject S14.

"The redesign process was a bit hard in the corrections, as we did not have support for developing the mockups." Redesign Process Negative Aspects - Subject S04.

"It has a lot of sub-items and it makes it difficult to prioritize…" Organization Suggestion - Subject S03.

"There could be a personalized filter according to the software engineers' needs, filtering by color, and providing charts to see the highest problems…" Organization Suggestion - Subject S02.

"... there could be a new tab, where the report organizes the items according to the highest occurrence." Organization Suggestion - Subject S14.

"It would be better if it was organized in tables or pages, one per component, such as "forms" or something…" Organization Suggestion - Subject S12.

Finally, we also found out how the software engineers prioritized the correction of the identified problems. In this sense, as REX provided the percentage of users who indicated an item or a negative emotion, some software engineers used this information to indicate which problems to correct first (see a quote from Subject S08). Among other criteria that were employed for selecting which problems to address, some software engineers indicated that they also considered correcting an issue first, if it had more chance of being experienced by users (when carrying out main tasks in the application) (see a quote from Subject S09). Another prioritization strategy was based on the criticality and impact of the item (see a quote from Subject S16). Finally, an interesting prioritization strategy was based on the experience of the software engineers themselves. Some of the subjects indicated that they considered problems with which they agreed most, as their experience was the same as the reported by the users (see a quote from Subject S10). This strategy is interesting, as the software engineers may be creating empathy and relating to the problems that the users experience.

"We focused on problems where most users felt confused. When an item had a high percentage, we knew that users had trouble understanding of using the application." Prioritization Strategy - Subject S08.

"We prioritized based on the tasks/parts that were most experienced by the users, such as navigating or accessing the main options." Prioritization Strategy - Subject S09.

"We also considered the ones we thought had the higher relevance and chance to impact in the quality of the application." Prioritization Strategy - Subject S16.

"I particularly chose the problems that I myself experienced." Prioritization Strategy - Subject S10.

5.3.3 Conclusions from Study 2

This study showed that the report provided by REX was useful for the software development team when redesigning a mobile application. The results from the TAM questionnaire were a positive indicator that REX could be useful and easy to use. However, software engineers indicated that they had difficulty in reading the report and understanding it as it had too much information. The results show that there is a need for better organizing the report results from the REX approach and allowing the filtering of UX data to facilitate the redesign process.

5.4 Discussion and Improvement Opportunities

As Study 1 showed, REX is effective in identifying more UX problems (considering the UX problems identified from the REX specific items) and suggestions while it takes the same amount of time as the 3E method. Since REX already has predefined questions and answers according to the emotions the user selects, the evaluators do not need to spend resources on adapting the evaluation session. Thus, we intend to provide a starting point for evaluators to gather UX data in future evaluations.

Regarding validity issues in Study 1, users' experience can be a threat if they were not part of the target audience of the application. However, only users who used mobile phones and were aware of the zika virus were selected to participate in the study. Also, we assigned the students to the REX approach and 3E technique guaranteeing that

none of the groups had more experienced users than the other in terms of software development and usability.

Another issue regarding the validity of our results in Study 1 is that users did not actually experience the final version of the application to indicate the location where the mosquito could breed or find information on the Zika virus and how to prevent it. However, they did simulate their interaction with the ZikaZero application through a navigable PDF created with the Balsamiq Mockups tool. There is a risk of using PDF mockups as they may not allow simulating all possible interactions with the application under development. Nonetheless, all measures were taken to verify that the available screens were sufficient to provide an overall understanding of the interaction flow. Additionally, all tasks that needed to be performed were considered within the available set of mockups. Therefore, if a user was not able to carry out a task, it meant that the application had issues, as it was reported by users. Furthermore, the mockups were tested in a pilot study (with two participants outside the study) prior to the execution of Study 1, in order to verify if users would be able to use it without the assistance of an evaluator. We highlight that the reason for choosing mockups instead of evaluating the real application with users was made for three reasons: (a) mockups allow resembling the experience of using a real application without having to develop the real application for testing (Rivero et al., 2010); (b) carrying evaluations earlier in development is needed in order to reduce the costs of correcting identified UX problems (Almahmoud et al., 2016); and (c) there is a need for further investigation of the performance of UX evaluation technologies in earlier stages of the development process (Bargas-Avila and Hornbæk, 2011). Furthermore, if software development teams intend to evaluate future prototypes of a mobile application, we suggest that the prototype is understandable and navigable on its own in order to avoid the presence of a moderator during the testing of the application. That way, users can go back to the prototype during their evaluation with REX without requiring assistance.

Although ZikaZero is a real mobile application under development and is representative of informative mobile applications, we cannot generalize our current results to all types of applications, as there are many other categories of mobile applications (Games, Health, News, others). Nevertheless, the quantitative and qualitative data can provide useful information for the applicability of the REX approach from the point of view of users, while also suggesting improvement

opportunities, which could be later implemented and tested, increasing the external validity of the study (Wohlin et al., 2012).

There can also be a threat considering the process of categorizing the qualitative data when analyzing the results from the identified UX problems, suggestions and answers from the TAM questionnaire. However, a team of UX experts carried out the classification, and they reached a consensus whenever differences were found. Additionally, in the analysis of the users' comments, we used qualitative analysis procedures (Glaser and Strauss, 2009) in order to mitigate this threat, given that it requires the entire analysis to be grounded in the collected data. Furthermore, the analysis process was performed along with another researcher, to encourage a better validation of the interpretations through the mutual agreement of the researchers.

A final limitation in Study 1 could be the instrument and measures applied for assessing technology acceptance. However, we believe that applying questionnaires was more suitable than applying interviews due to time constraints with carrying out interviews with 82 subjects for Study 1 and 31 for Study 2. Furthermore, by evaluating the time spent, a number of identified problems, false positives, suggestions, perceived usefulness and perceived ease of use, we intended to have an idea of the performance and users' acceptance of the REX approach and identify issues that should be corrected to meet their needs.

Study 2 showed the interest of software engineers in applying REX due to its organization and the pairing of UX problems with design suggestions in a single report. Although several UX evaluation methods have been proposed, few of them guide software engineers in the correction of the identified UX problems (Vermeeren et al., 2010). In product development, it is important to provide means to software development teams in order to improve the quality of the developed applications. By providing design suggestions, we aim to support software engineers in improving the UX of the developed applications, thus increasing their acceptance in the market.

One of the limitations of Study 2 was applying the REX report with students in their last semester of the university which had no experience in the industry. However, only half of the subjects were included in this category (the rest were practitioners participating in the course), and according to Carver et al. (2003), students who do not

have experience in industry may have similar skills as novice software engineers in the industry. Thus, the feedback provided by the subjects was useful for identifying aspects that made it difficult to apply the REX report from the point of view of novice software engineers, while also considering the opinion of experienced practitioners. Moreover, to mitigate the threats regarding the redesign application and its representativeness, the subjectivity of the qualitative analysis, and the applied instruments and measures, we have proceeded as described in Study 1.

Based on the qualitative data from both studies, we identified improvement opportunities in the REX approach such as:

1) Allowing explaining the selected emotions with other means when an option is not available.
2) Including an introduction to the REX approach explaining what will be evaluated in each stage and how to report an experience.
3) Dividing the UX report into components to facilitate finding UX problems in a specific component of the application.
4) Providing further explanations on how to apply the design suggestions and sketched layouts for software engineers to customize when redesigning an application.
5) Allowing software engineers to filter or order the information they need according to the type of problem or percentage of users who indicated them.

5.5 Improvements in the REX Approach

The qualitative data from both studies allowed us to identify improvement opportunities in the REX approach. These improvements have been implemented to facilitate the evaluation process from the point of view of users, while also facilitating the understanding of the evaluation report from the point of view of software engineers.

5.5.1 Improvements in the REX Technique and Evaluation Tool

The results from the first study suggested that users were confused at the beginning of their evaluation and that they did not know how to proceed in order to report their experience. To help improve this feature, we have developed a tutorial for the REX

evaluation tool. In this tutorial, the user is presented with the necessary information about what will be asked along the evaluation and how the evaluation will occur.

Figure 5-5 shows the main screens of the tutorial. In these screens, the user is told that (s)he will be guided through the evaluation process, evaluating specific parts/functionalities of the application (see part A). The user is also shown that once (s)he indicates that (s)he used a part/functionality of the application, (s)he will see a set of emotions and items to report his/her experience (see part B). Finally, the user is also told that if (s)he does not find a specific item or emotion for reporting his/her experience, (s)he can always report further details through the use of a special option, and that help would be available at all times (see part C).

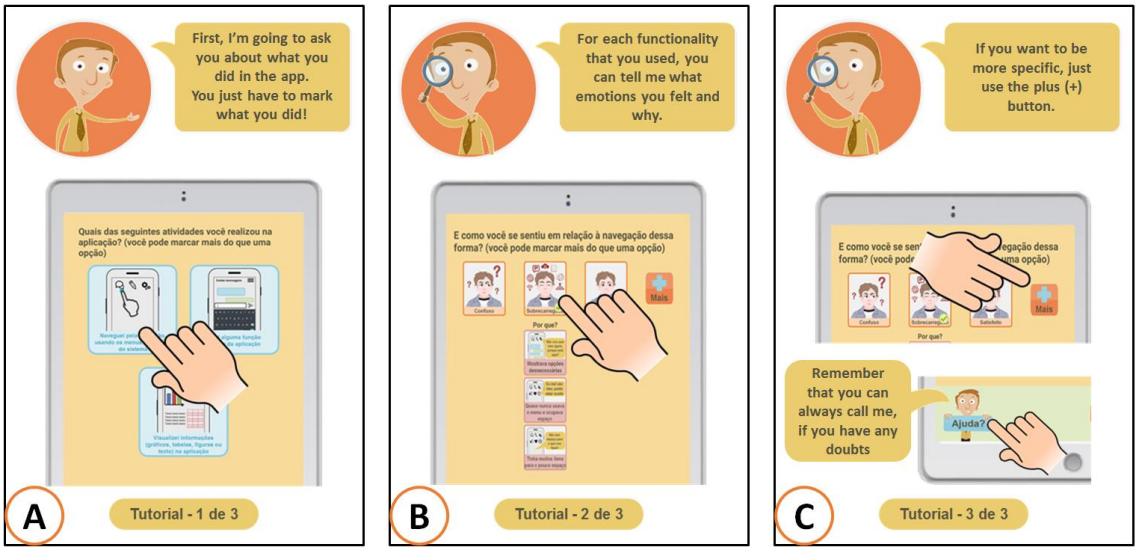


Figure 5-5 The tutorial indicating how the evaluation will occur.

One of the comments from users was the lack of the REX evaluation tool to provide means to report further comments freely. Some users did not know that they would have available space at the end of the evaluation session to provide such comments. Additionally, other users reported that it would be more useful if they were able to indicate if the problem was related to one of the parts/functionalities that were being evaluated. To facilitate the feedback giving process, we considered those comments and added a "more" button which would lead users to a new screen in which they would be able to report further details. Figure 5-6 shows how this new functionality would work. As explained in the tutorial (Figure 5-5), if users feel that they have something else to report, they can access the more comments screen within the REX approach. The extra comments will appear in the report and related to the

evaluated aspect. For instance, if the user was evaluating the navigation of the application and indicated a new problem, that new problem would be related to the navigation of the application, facilitating the understanding of the problem by the development team.

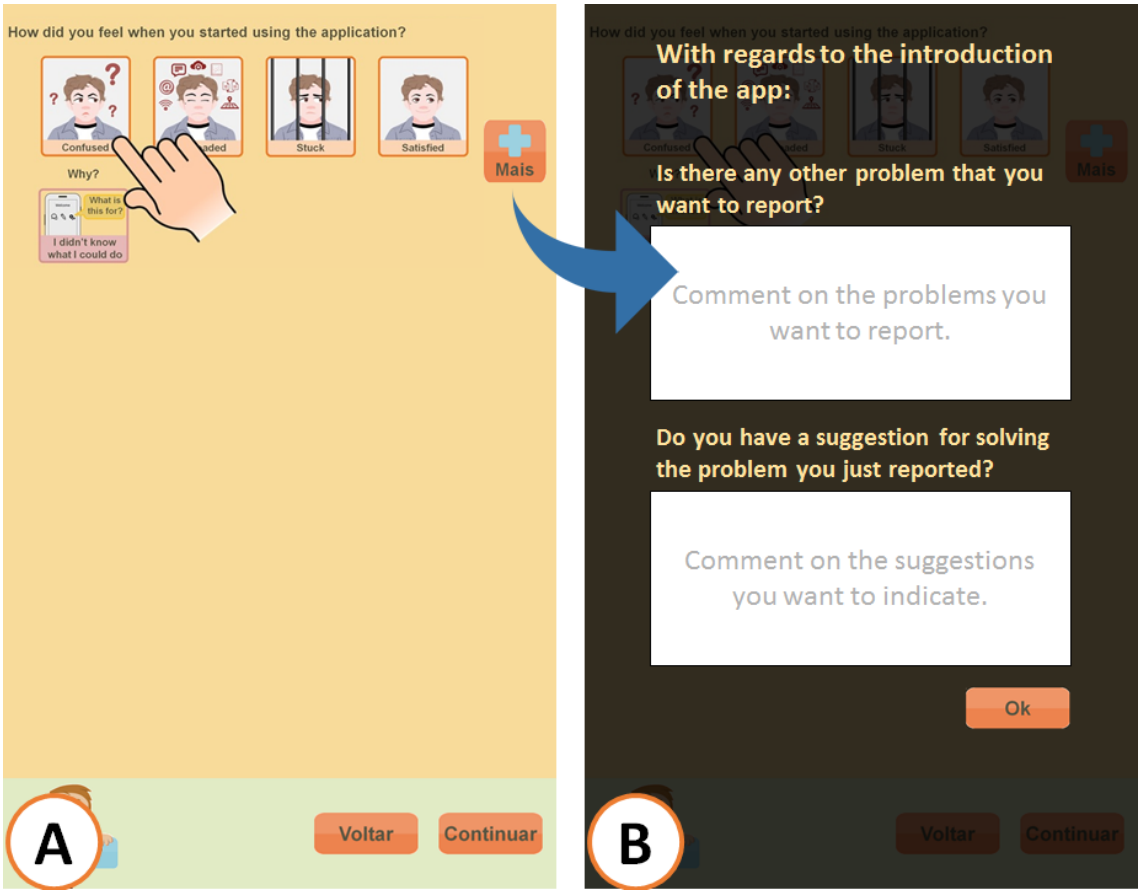


Figure 5-6 The new option for allowing users to report further problems and suggestions.

5.5.2 The REX Report Generator Tool

In the initial version of the REX approach, the report was provided through a spreadsheet, in which the data was aggregated and paired according to the evaluated aspects and emotions. However, software engineers reported difficulties in understanding the report and its organization through this initial template. To improve the usefulness of the report generated through REX, we developed a new tool called REX Report Generator, which creates complete UX evaluation reports of a mobile application, aggregating the individual reports of users.

The tool was developed as a desktop application to facilitate its use. Figure 5-7 shows screens of the initial steps of generating a report using the REX Report Generator. To create a report, the tool presents a brief explanation. Then, the evaluation moderator (i.e., the person responsible for handling the UX evaluation data) can choose between exporting a report to a PDF file or viewing the evaluation results in the tool itself. To indicate which UX data will appear in the report, the moderator has to select the individual files that each user generated after his/her UX assessment using REX.

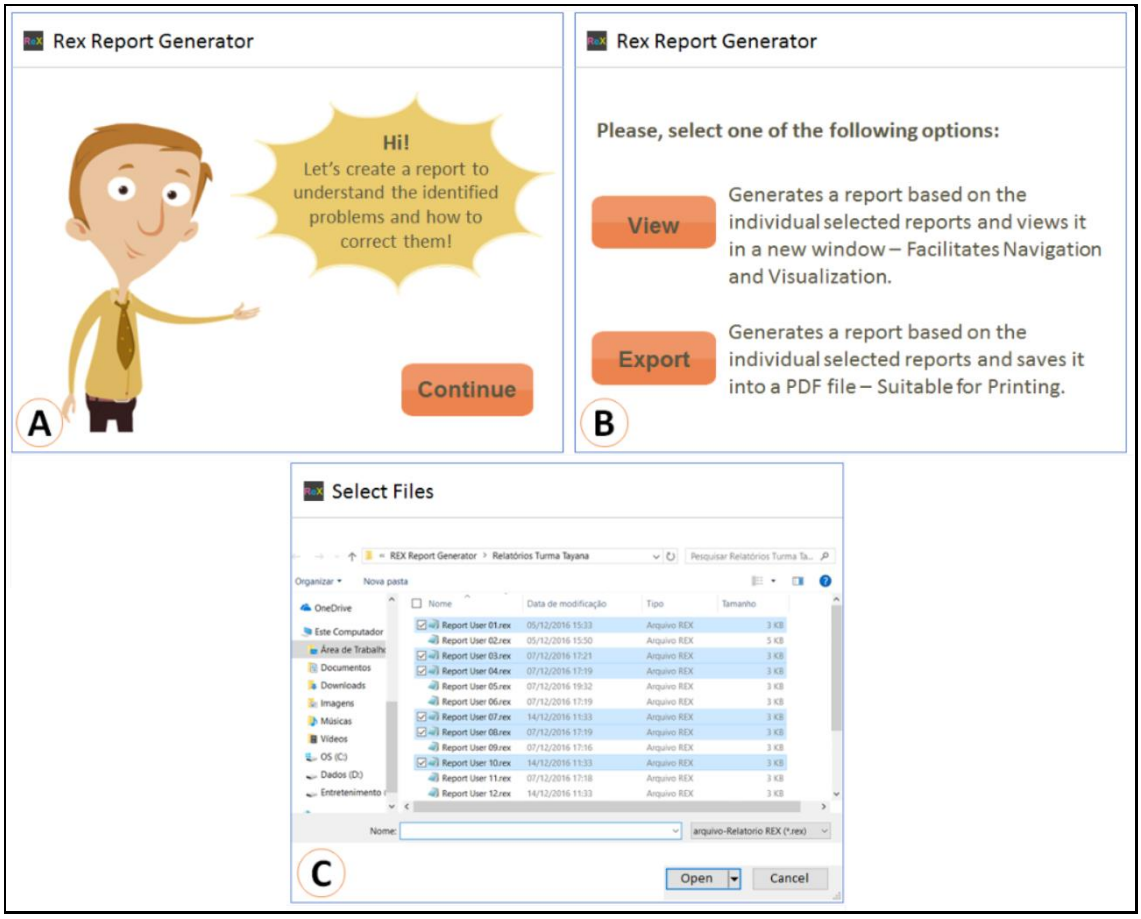


Figure 5-7 Process for generating a complete UX evaluation report using the REX Report Generator

5.5.3 The Updated Architecture and Application Process

Figure 5-8 shows the architecture from the REX approach in its second version, considering the technique, its evaluation tool, and its report generator tool. In this figure, the REX technique is composed of questions, emotional responses, and items. These are incorporated into the tool, which automates them through the moderator avatar and guides the user through the evaluation. Based on the individual evaluations, a

technical report to support the redesign process can be created using the REX report generator. The procedure for carrying out an evaluation remains the same. However, the redesign support using the REX approach will be explained below.

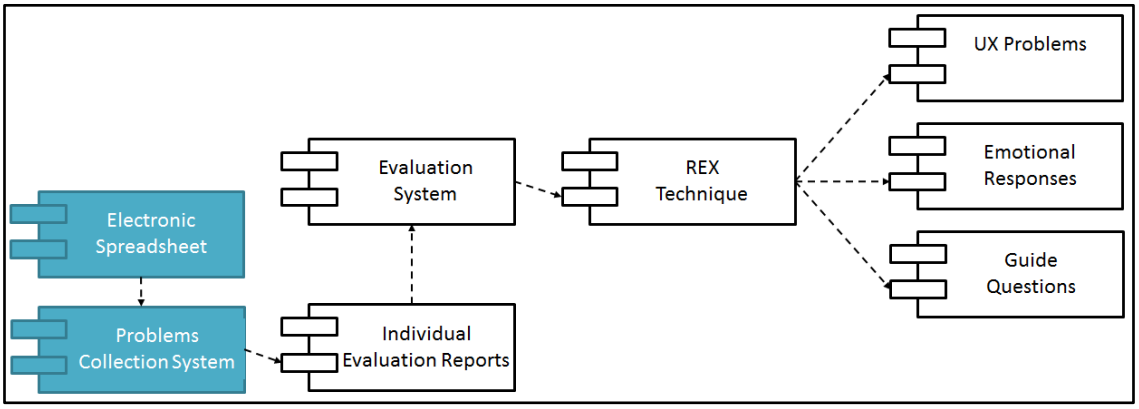


Figure 5-8 The architecture from the REX approach v2

To show an example of how the individual reports are condensed into a single complete redesigning report using the REX Report Generator, we have evaluated a mobile application called SIGNAL, which stands for System for Intelligent enGagement aNALysis in digital education contexts. SIGNAL is a teaching-learning application meant for helping the teacher to compose, run and evaluate a class, beyond verifying the students' performance with learning assessments using questionnaires with help on demand. The application tracks the students' interaction with the learning material in class and gathers data regarding if they are paying attention, if they are able to understand the didactic content, or if they are carrying out other activities with their mobile device. Figure 5-9 and Figure 5-10 show one of the screens of the SIGNAL application and part of the UX report (regarding the evaluation of that part of the application) being viewed in the REX Report Generator tool, respectively. In the report, we show: (1) the list of emotions that the users felt regarding this component, (2) the list of UX problems faced by the users when using this component, and (3) the suggestions provided based on the design suggestions or by the users themselves.

Regarding answering a questionnaire about a class topic using SIGNAL (see Figure 5-9), the report suggests that some users felt confused and trapped, while others felt satisfied. The explanations for the negative emotions are provided in the reported UX problems. These problems are organized based on the number of users who pointed them (from higher to lower). In our example, most users indicated that they did not

know if they had successfully performed the task of answering the questionnaire. This happened because when they clicked the "Answer" button (see Figure 5-9) the interface remained the same and they had to click on the other questions to continue answering the questionnaire. However, whenever the student clicked in a previously answered question, the interface showed the question without showing the marked answers. This confused students, making them answer the same questions over and over.

To support software engineers in correcting each of the identified UX problems, the development team can view suggestions within the REX report by clicking the "Suggestion" button next to the definition of a problem. For instance, regarding the UX problem of not knowing whether they had answered the questionnaire, a possible solution could be to provide a message that the answer has been saved, or update the interface, showing the next question, and keeping track of the user's answers whenever (s)he goes back to a previously answered question. We highlight that the report can contain further problems and improvement suggestions proposed by the users themselves during the overall emotional evaluation, besides the specific items from the REX approach. This is to ensure that no feedback is lost during the evaluation process.

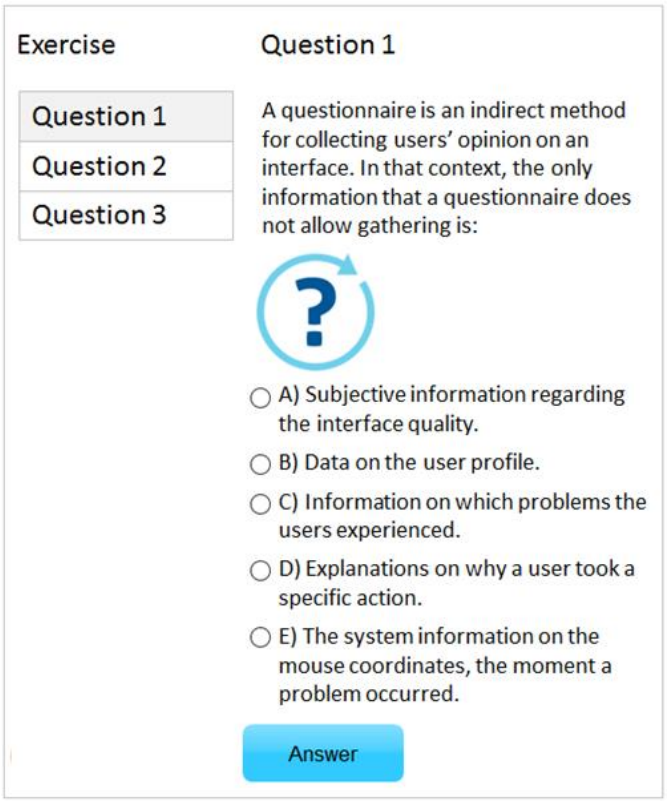


Figure 5-9 Screen from an evaluated application

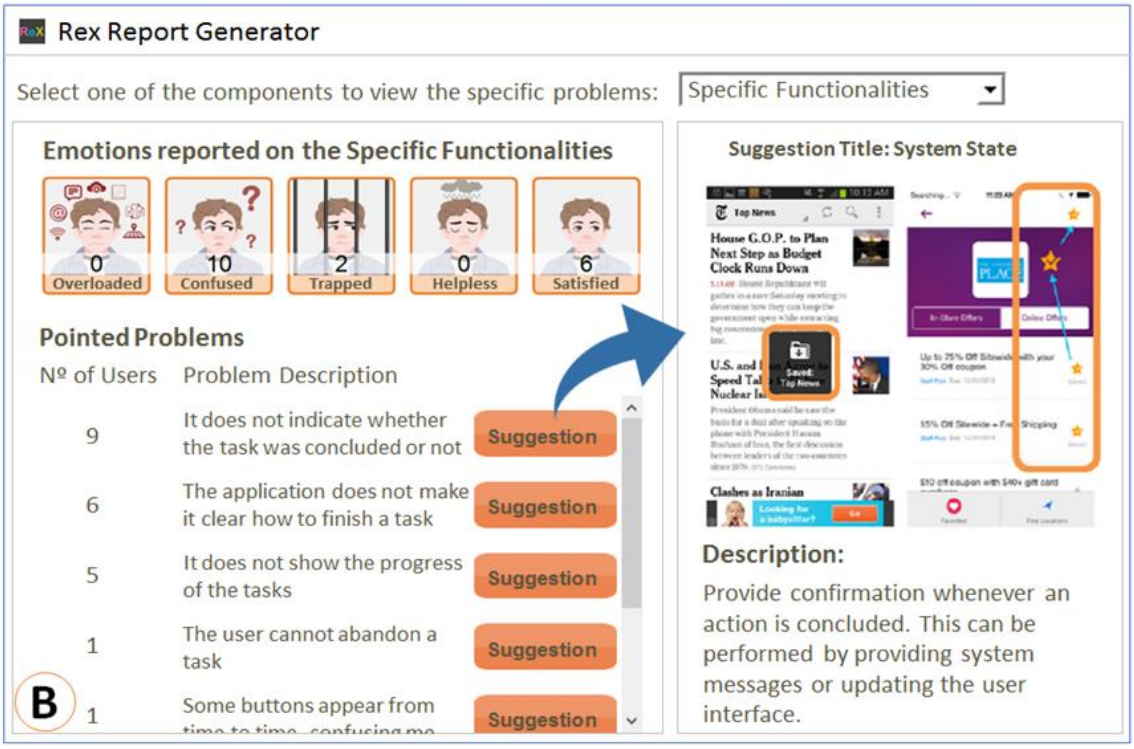


Figure 5-10 The new version of the REX UX report containing problems and improvement suggestions

Considering the evaluation and redesigned process described above, Figure 5-11 shows an activity diagram of the evaluation and redesign process of the REX approach in its second version, considering: (a) the moderator, which can be a member team deciding what to correct in the application; (b) the evaluator, which can be a user testing the application; (c) the REX evaluation tool, which automates the guiding process of the REX technique; and (d) the REX report generator tool, which automates the process of merging the individual evaluation reports and analyzing which problems should be corrected and how. The differences between this process and the original process are highlighted in blue in Figure 5-11.

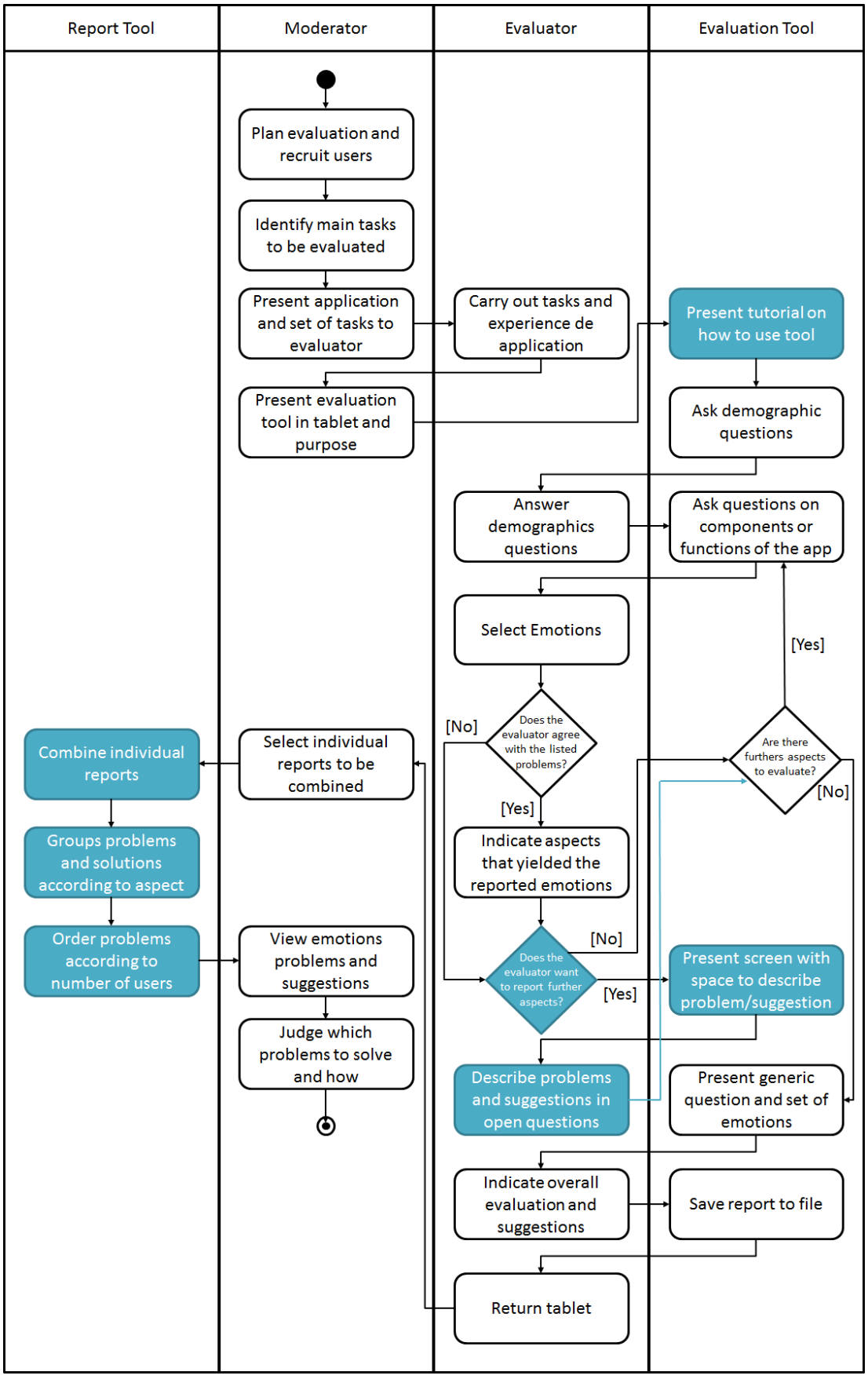


Figure 5-11 The evaluation and redesign process from the REX approach v2

5.6 Summary

In this chapter, the two studies demonstrated the use of the REX approach from the point of view of users and software engineers. The approach was employed for reporting experiences and the negative aspects that impacted them when using a mobile application called ZikaZero (Study 1) and redesigning that application based on the collected feedback (Study 2). As these studies were independent, the subjects who participated in the evaluation were not the same as the ones participating in the redesign. In that context, the results from applying a TAM questionnaire suggested that REX was perceived as useful and easy to use (from the point of view of users in Study 1 and from the point of view of software engineers in Study 2).

Given that each qualitative study provides evidence and hypotheses that can be later tested using quantitative methods, we have evaluated how including the improvement opportunities identified in these studies facilitated the application process of the REX approach from the point of view of users and software engineers. Thus, this chapter reports the first experiences using the REX approach to identify UX problems and support the redesign of an application based on the identified UX data. REX guides users through the evaluation process, and it produces better results than other forms in terms of UX aspects and improvement suggestions. Although the required time to be applied is the same as techniques such as 3E, REX is perceived as easier to use, while users indicate that it is fun, engaging, and interactive, which is difficult to achieve with static questionnaires. Additionally, the design suggestions embedded into the technique for each UX problem can guide software engineers through the redesign of the application. REX therefore, could support the UX evaluation and redesign of mobile applications, as it can provide quantitative and qualitative data in a structured way, providing suggestions to the development team to handle such data.

Also, we showed how we made improvements in the REX approach and created a new tool for software engineers to view the report. The proposed changes were implemented following the suggestions from the previously executed feasibility studies. By making the above changes, we intend to improve the usefulness of the REX approach from the point of view of users and software engineers.

In the following chapter, we present the planning and execution of an observational study that aimed at investigating the use of the REX approach in a real software development project. In such study, the new version of the REX approach and the REX Report Generator tool have been employed to gather UX data and provide improvement suggestions.

6 An Observational Study of REX

This chapter presents how we carried out an observational study to verify to which extent the REX approach could be applied in a real software development project. REX was employed by users to evaluate a mobile educational application, and a discussion meeting was held with the software development team to discuss the improvement suggestions provided by REX to support the redesign process. This chapter presents the findings from the observational study, where users indicated their satisfaction with the REX approach, while the members of the development team agreed with the usefulness of the REX report and its improvement suggestions.

6.1 Introduction

The development of the SIGNAL application (see Subsection 5.5.2) was requested as part of a research project that was financially supported by a multinational company. To improve the quality of the application, its development team requested a UX evaluation for identifying UX problems and improvements. Thus, we saw an opportunity to evaluate the REX approach in a real development project. As a result, we carried out an observational study to evaluate the acceptance and performance of the REX approach for: (a) identifying UX problems from the point of view of users, and (b) reporting UX problems and suggest solutions from the point of view of software engineers.

The remainder of this chapter is organized into three more sections. In Section 6.2, we present the planning and execution of the study. Then, in Section 6.3, we describe the quantitative and qualitative analysis of the results. Finally, Section 6.4 discusses our findings and concludes this chapter.

6.2 Planning and Execution

6.2.1 Gathering Data from the Point of View of Users

In order to gather information on the users' perception of the REX approach for reporting their experience, we carried out a UX evaluation of the SIGNAL application in class. Consequently, both students and teachers were part of the target audience of the SIGNAL application. Since only the student environment of the SIGNAL application was ready for testing, we carried out the evaluation with students. At all, 17 students from a local university in the city of Manaus (Brazil) were recruited for the study. The students were enrolled in a class on advanced topics in software engineering, which included topics on software quality and emerging software engineering methodologies. Before the study, all subjects who agreed to participate filled out a consent form and a characterization form. The consent form requested the participants' agreement in participating in the study and explained the main activities of the study and the confidentiality of the results. Additionally, the characterization form aimed at determining the users' profile.

To carry out the first part of the study, we followed the procedures shown in Figure 6-1. Before the study, we prepared the necessary materials to lecture a class on software quality using the SIGNAL mobile application. Then, in the participation day, the students entered a classroom where a teacher greeted them and explained the purpose of the SIGNAL mobile application. Each subject received a tablet with the running SIGNAL application containing all class topics and activities. The main topic of the class was usability evaluation, focusing on specific concepts, examples of evaluation methods and in which contexts they were suitable. To avoid distracting users during their experiencing of the SIGNAL application, we did not use a projector or printed material. Thus, the teacher would also use a tablet to guide users throughout the class. Additionally, all materials and presentations were downloaded into the tablet and viewed with the SIGNAL application. At all, the students performed four tasks in the one-hour duration class: (i) view a power point presentation, (ii) view a video, (iii) view and carry out an assignment, and (iv) answer a questionnaire about the lectured topics. After interacting with the application, the subjects closed the SIGNAL app and opened the REX evaluation tool to start the UX evaluation. During the UX evaluation, the

moderator and two assistants were responsible for counting the time spent per user to finish reporting his/her experience.

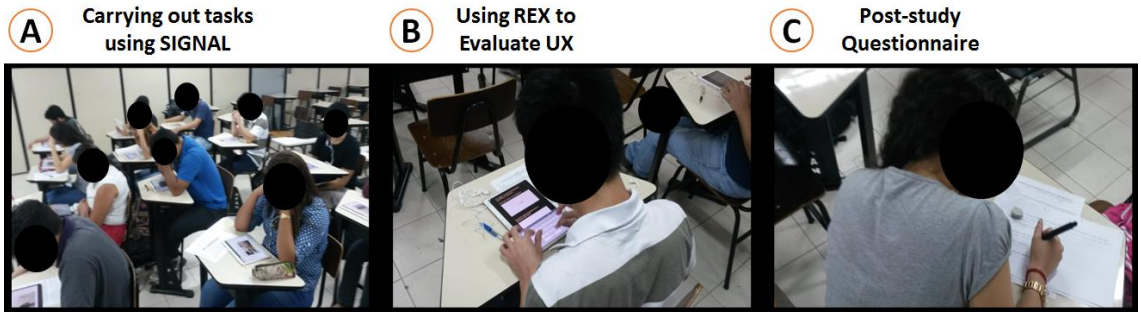


Figure 6-1 Steps for gathering data on the perception of the REX approach for carrying out a UX evaluation

After delivering their evaluation reports, the students were asked to fill in a post-study questionnaire to gather data regarding their acceptance of the REX approach. The questionnaire was the same applied in the feasibility study of the REX approach and was based on the indicators from TAM 3 (Venkatesh and Bala, 2008).

6.2.2 Gathering Data from the Point of View of the Development Team

After we identified the UX problems and gathered data on the users' perception of the REX approach, we scheduled a meeting with the SIGNAL development team to show the obtained results. Also, at this meeting, we gathered information on the applicability of the REX report to support the redesign of an application in a real software development project. The SIGNAL development team had 8 team members, including developers, analysts, and educators who acted as consultants.

Figure 6-2 shows the procedures we followed in the second part of the study, where we gathered data from the point of view of software engineers. At all, 5 out of the 8 team members participated in the meeting. The other team members did not attend the meeting due to schedule problems. All subjects who agreed to participate filled out a consent form and a characterization form. Almost all subjects had more than 3 years of experience in software development or coordinating activities in a software development project. Only one of the education consultants had less than a year of experience since this was one of the first projects in which (s)he participated.

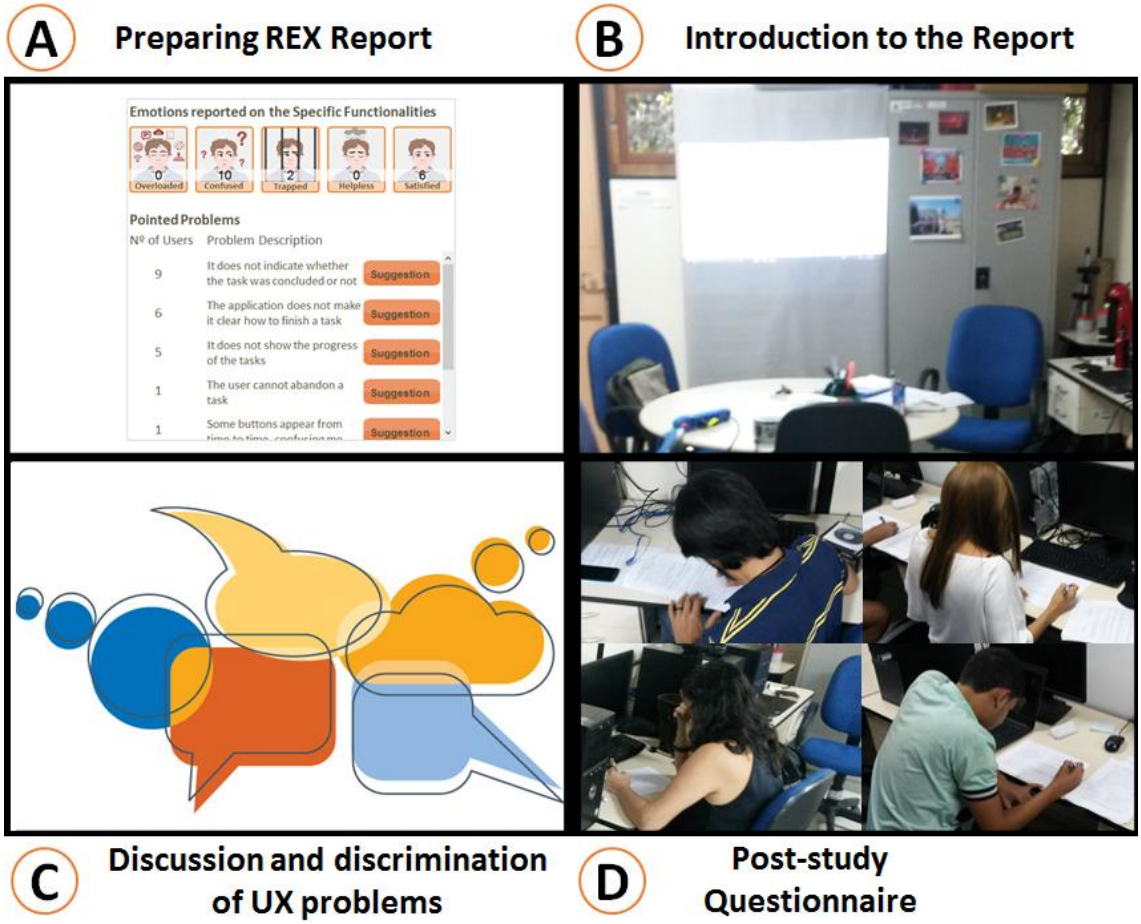


Figure 6-2 Steps for gathering data on the perception of the REX approach for providing a UX report with suggestions

We used the REX Report Generator (see Subsection 5.5.2) to create a complete report of the evaluations held during the software quality class. After that, we presented each member of the development team with the printed version of the report and provided a projector with the dynamic version of the report to be viewed in the desktop tool. After introducing the report and explaining how it worked, we invited the team leader to lead the meeting and review the identified emotions and UX problems with the rest of the development team. During the meeting, the development team had to discriminate the identified problems, indicating whether they were: (a) not a problem, (b) a problem with low correction priority, and (c) a problem with high correction priority. The purpose of carrying out that categorization was to encourage the development to discuss the report and think of the results obtained through the use of the REX approach. The meeting lasted for one hour and thirty minutes. We highlight, that the development team did not redesign the SIGNAL application, but analyzed the reported UX problems and suggestions to discuss what should be changed.

After the discussion, to gather data for evaluating the acceptance of REX from the point of view of software engineers, again, we applied a questionnaire based on the indicators from the Technology Acceptance Model 3 (TAM 3). In the following subsection, we present the results from this study evaluating REX from the point of view of users and software engineers.

6.3 Results

6.3.1 Quantitative Analysis

We used the results from the discrimination meeting to analyze the performance of the REX approach. Thus, we counted the number of real UX problems and false positives that the users reported. Also, we analyzed the time spent by users to carry out a UX evaluation using REX. At all, with the participation of 17 users of the SIGNAL application, we managed to identify a total of 27 different UX problems (see examples in Subsection 5.5.2), each of them paired with a suggestion (either automatically generated by the REX approach or suggested by the users). The results shown in Table 6-1 suggest that REX does not take long time to be applied (in most cases, less than 30 minutes), but it assists users in reporting several UX problems and few or no false positives.

Table 6-1 Individual and overall quantitative results on time spent, total identified problems and false positives from users applying the REX approach

ID	U01	U02	U03	U04	U05	U06	U07	U08	U09	U10	U11	U12	U13	U14	U15	U16	U17	**Mean**	**Std. Dev.**
Time Spent	8	5	19	10	24	18	8	5	13	36	13	12	11	13	10	21	30	**15,1**	**8,6**
Total Problems	2	7	5	4	4	2	1	0	3	14	3	5	1	4	2	11	12	**4,7**	**4,1**
Total False Positives	0	0	0	0	1	4	0	0	0	3	0	0	2	0	0	1	1	**0,7**	**1,2**

After the subjects (users and members from the SIGNAL development team) had delivered their TAM-based questionnaires, we counted the degrees of agreement with each of the TAM items in order to provide an overview of the acceptance of the REX approach. Table 6-2 and Table 6-3 show the TAM results, which suggest that both users and software engineers considered that the REX approach is useful, easy to use and they would use it for finding UX problems and suggestions in mobile applications.

6.3.2 Qualitative Analysis

We analyzed the data obtained from the open questions in the post-study questionnaire to investigate the aspects that affected the usage and acceptance of the REX approach. To perform this analysis, we applied the same qualitative analysis procedures in Sections 5.2 and 5.3 (Glaser and Strauss, 2009), creating codes related to the citations within the questionnaires. After that, we created the relationship codes. We managed to identify codes related to the difficulties and facilitators of the REX approach; the reasons that would make users or software engineers use it again for performing a UX evaluation or identify problems and improvement suggestions; and the aspects that could be modified to improve its acceptance.

Table 6-2 Individual results, median and standard deviation for perceived usefulness, perceived ease of use and behavioral intention to use indicators for the REX approach from the point of view of users

	U01	U02	U03	U04	U05	U06	U07	U08	U09	U10	U11	U12	U13	U14	U15	U16	U17	**Median**	**Std. Dev.**
PU1	5	4	7	5	4	6	6	5	6	7	6	7	5	6	7	6	7	**6**	**1,0**
PU2	5	2	7	6	7	6	5	5	7	6	6	6	4	6	7	5	6	**6**	**1,3**
PU3	6	1	3	5	7	6	3	3	7	3	5	5	5	3	7	7	5	**5**	**1,8**
PU4	6	5	5	5	7	7	6	4	7	6	7	6	6	7	7	5	7	**6**	**1,0**
PEoU1	7	1	7	7	7	5	7	7	6	7	7	7	7	6	7	5	7	**7**	**1,5**
PEoU2	7	5	7	7	7	5	7	7	6	7	6	7	7	6	7	6	6	**7**	**0,7**
PEoU3	7	5	7	7	7	6	7	7	6	7	7	7	7	6	7	6	6	**7**	**0,6**
PEoU4	7	4	5	6	7	5	7	7	6	7	7	7	7	6	7	7	6	**7**	**0,9**
BI1	5	1	7	6	7	5	7	5	7	6	6	5	4	4	7	6	6	**6**	**1,5**
BI2	5	1	7	6	7	3	7	5	7	6	6	5	4	5	7	6	6	**6**	**1,6**
BI3	5	1	6	4	7	2	4	5	7	5	5	6	4	3	7	4	4	**5**	**1,7**

Table 6-3 Individual results, median and standard deviation for perceived usefulness, perceived ease of use and behavioral intention to use indicators for the REX approach from the point of view of the development team

	TM01	TM02	TM03	TM04	TM05	**Median**	**Std. Dev.**
PU1	6	6	7	7	6	**6**	**0,5**
PU2	6	6	7	6	6	**6**	**0,4**
PU3	7	6	7	7	7	**7**	**0,4**
PU4	7	7	7	7	7	**7**	**0,0**
PEoU1	7	5	7	5	5	**5**	**1,1**
PEoU2	7	6	6	6	5	**6**	**0,7**
PEoU3	7	5	7	5	6	**6**	**1,0**
PEoU4	7	7	7	7	7	**7**	**0,0**
BI1	7	7	6	7	6	**7**	**0,5**
BI2	7	7	7	7	6	**7**	**0,4**
BI3	7	6	7	7	6	**7**	**0,5**

Below, we present our results from the point of view of users and software engineers. We refer to (a) the users through the code UXX, where XX is the number of the user who participated in the UX evaluation from 1 to 17; and (b) the software engineers using the code TMX, where X is the number of the team member who participated in the discrimination meeting from 1 to 5.

Overall, the qualitative results from the point of view of users showed that they were satisfied using REX and that the goals of providing an easy to use, guiding and appealing evaluation method were met (see quote from Subjects U08, U09 and U11). Despite the positive feedback, the users faced some difficulties and suggested improvement opportunities. For instance, some users felt that some items for reporting their experience were missing (see quote from Subject U03). Although REX provides means to report aspects that were not considered in the initial set of items, these aspects could be refined in each application of the REX approach in order to complement the original set. Additionally, some users reported that they wanted to have more powerful tools to report complex experiences. For instance, one subject indicated that (s)he felt both positive and negative emotions towards the application and that (s)he could not indicate that in REX (see quote from Subject U09). Although REX provides means to report multiple emotions, if the user wants to report multiple features that impacted his/her experience, (s)he must do it in the same space. This could be modified to

provide further details on which aspects were positive and which were negative. Another aspect that confused the users was the need for a tutorial in the app. Since the original question was not straight forward (i.e., "How was your first contact with the app"), some users felt confused whether they were evaluating the tutorial or the capability of understanding the application without it. As a result, since the SIGNAL application did not have a tutorial, some users reported having described the same problems in different parts of the evaluation (see quote from Subject U16). Finally, although subject U02 disagreed with most TAM statements, he did not provide information on the aspects that affected his answers.

> *"Besides having a nice interface, the method was easy to understand, which didn't demand much time."* - User 08 on Positive Aspects

> *"It explains what needs to be done in every stage."* - User 09 on Positive Aspects

> *"The friendly interface, the color and the way in which it is presented, through a person, made me feel at ease."* - User 11 on Positive Aspects

> *"I felt that there were some missing options in order for me to report 'why' I had chosen an emotion."* – User 03 on Problems

> *"Trying to explain contradictory emotions, such as liking the system at a whole, but not liking a specific functionality."* – User 09 on Problems

> *"I just felt a bit confused on what was being evaluated in the beginning"* – User 16 on Problems

The users also had some suggestions in order to improve future evaluations using REX. For instance, the users wanted to report further emotions, whenever they described an aspect that affected their experience that was not considered in the original set of items (see a quote from Subject U04). Also, they suggested providing parts of the application inside REX in order to mark the exact location of a problem (see a quote from Subject U01). Finally, some users indicated that it could be useful to know how much of the evaluation had been performed (see a quote from Subject U05). These suggestions can be taken into consideration in future versions of the REX approach.

"Allowing including further options and emotions at the end…" – User 04 on Suggestions

"I would allow the evaluators to provide part of the evaluated application. This would help, as we would be able to report further problems." – User 01 on Suggestions

"Maybe provide a status of the progress of using the method." – User 05 on Suggestions

Regarding positive aspects of the REX report for supporting the understanding of UX problems and suggesting improvements in an application, we identified several features. For instance, software engineers indicated that the report was very organized, facilitating its understanding (see a quote from Subject TM5 below). Another positive aspect was that the REX report provided suggestions for each identified problem and that it could be useful for improving the quality of the application (see quotes from Subjects TM4 and TM3). Finally, one interesting aspect of the report is that some of the suggestions would encourage the development team to discuss different aspects of the application that were not considered before (see quote from Subject TM2). We noticed such flow of ideas during our observation of the discrimination meeting. For instance, when discussing problems regarding the login of the app, one user indicated in the REX report that (s)he did not know if (s)he had performed well, as (s)he clicked the "login" button by mistake without entering his/her name. The development team indicated that this would not be a problem, as there were only 17 students in the class and they would know who the student was. However, as the discussion continued, the development team started thinking of the necessity of different registration alternatives, especially if the number of students increased, as well as the number of classes. This discussion made the development team focus on other login aspects and how these aspects would affect the use of the application.

"The report is divided into topics, which makes it easier to understand the flow of problems." - Team Member 5

"The suggestions, even those proposed by the users, were clear and objective." - Team Member 4

"It is useful for turning the application more appealing and motivating." - Team Member 3

"Some suggestions brought opportunities to explore features that we have not even thought about. Even things that the evaluator had not even thought about." - Team Member 2

Regarding negative aspects and improvement opportunities for applying the REX report, we found out that the main problems were regarding its repetitive items or items that were not clear enough to identify the exact part of the interface that should be modified (see a quote from Subjects TM2 and TM4). Although the REX Report Generator removes duplicated answers for the items that were marked by more than one user, it does not combine or analyses the answers provided by the users with their own words. That caused the report to list problems or suggestions that were provided by more than one user, and even allowed that the users provided vague opinions. In order to improve the quality of the report, the software engineers suggested combining the open answers of the users or carrying out an analysis prior to the meeting to avoid wasting time in verifying if a problem or suggestion was already reported (see quotes from Subjects TM2 and TM4). Regarding the details of the reported problems, the development team suggested asking the user to provide information on where to find the problem (see a quote from Subject TM1). This could be implemented in the REX approach by tracking whenever the user refers to an interface element (e.g., buttons, labels, input fields, others) and asking the user to provide more details on such element. Finally, the team members suggested providing further graphs to understand the data better (see a quote from Subject TM1). Although the REX approach provides quantitative data on the number of subjects who indicated a UX problem and the number of felt emotions, we could provide a comparison of such quantitative data according to the mobile application components that were evaluated. The problems and suggestions described above are being analyzed to improve the support provided by REX when understanding UX problems and improvement suggestions.

"Some problems were not properly detailed." - Team Member 2 on Difficulties

"Some of the items were repetitive." - Team Member 4 on Difficulties

"As the moderator knows who provided the suggestions, (s)he could rewrite them or at least make sure that they are clear for the development team. It could also be done if the moderator and the development team discussed the problems and suggestions and verify if they are clear and well written." - Team Member 2 on Suggestions

"It would be better if the report could combine errors that were pointed by more than one user." - Team Member 4 on Suggestion

"The report could be more precise. For instance, when the user starts talking about a 'button,' it could ask the user: 'which button', so the development team could know exactly where to correct the problem, instead of trying to guess the location." - Team Member 1 on Suggestions

"It could show statistics, with graphics, not only counting the number of users." - Team Member 1 on Suggestions

Finally, we also gathered information regarding the rationale behind the prioritization of the correction of the identified UX problems. The development team had to categorize the reported UX problems into: (a) no problem, (b) low priority, and (c) a high priority. In this sense, when reaching a consensus considering the different opinions among the team, some of the criteria for classifying a problem were: (i) the priority of fixing that problem to allow the use of the application (see quote from Subject TM1); (ii) the degree of impact in the UX (see quote from Subject TM2); and (iii) the relevance of fixing the problem in order to achieve the goals when using the application (see quote from Subject TM5). These strategies suggest, that although achieving tasks with success with the application, the development team is also interested in increasing its ease of use and improve the emotional response of using SIGNAL. During the meeting, we observed that the development team wanted the application to provide a positive UX in order to encourage and motivate users during classes in order to improve their learning of the class topics.

"We classified the problems according to their degree of priority, in which the most important were the ones that were indispensable for the use of the application" - Team Member 1

"We classified the problems according to their negative impact for the application, aiming at providing a better experience for the user" - Team Member 2

"We related the problems to the relevance of the results of the application" - Team Member 5

6.4 Summary

This chapter reported the use of the REX approach in the context of a real mobile application under development. The observational study in the evaluation of the SIGNAL application provides examples of how to apply the REX approach for both identifying UX problems and using its automated generated report to verify improvement opportunities.

After applying REX for reporting experiences and analyzing its generated report, the results from applying the TAM-based questionnaire suggested that REX was perceived as useful and easy to use from the point of view of users and software engineers. REX is effective for allowing the report of UX problems, emotional responses and suggestions, when compared to other approaches in similar contexts (Rivero and Conte, 2015). Also, it only requires short evaluation sessions (with an average time of 15 minutes). Additionally, the automatically generated REX report allows software engineers to easily identify the emotions felt by users, the problems they faced when using the mobile application, and what could be done to improve the experience of using it. Finally, the discussions held when analyzing the reported problems could be a basis for generating new ideas in order to include further features in the developed application that could improve its quality.

Regarding validity issues in our study, the reported users' experience can be a threat if the subjects were not part of the target audience of the application. In our study, only students from class were selected to participate in the UX evaluation of the SIGNAL application. Another issue could be the context in which the application was used. However, the users performed several activities in the context of a real class on software quality, and they only used tablets to avoid their distraction with other class materials (e.g., projector, board, others). Also, the data collection on the applicability of the REX report in a real software development project was performed with a software

development team, whose members had more than 3 years of experience, indicating the needs from practitioners working in the software development industry.

The applied instruments and measures for assessing technology acceptance could be a threat to the validity of our results. However, we believe that applying questionnaires was more suitable than applying interviews due to time constraints with the students and the development team. Furthermore, by evaluating the time spent, a number of identified problems, false positives, suggestions, perceived usefulness and perceived ease of use, we intended to have an idea of the performance and acceptance of the REX approach, while identifying issues that should be corrected to meet future needs. Furthermore, regarding the subjectivity of the data classification as a threat to the validity of our results in the qualitative analysis, we used qualitative analysis procedures (Glaser and Strauss, 2009) in order to mitigate this threat, given that it requires the entire analysis to be grounded in the data collected.

Finally, one last limitation could be the representativeness of the SIGNAL application with respect to other mobile applications. Although SIGNAL is a real educational mobile application under development, as there are many other categories of mobile applications (Games, Health, News, others), we cannot state that our results apply to all types. Nevertheless, the quantitative and qualitative data can provide useful information for the applicability of the REX approach from the point of view of users, while also suggesting improvement opportunities to software engineers.

Based on the qualitative data from the observational study, we identified improvement opportunities in the REX approach such as:

- Providing means for users to view, select or indicate in which part of the application the problem occurred (specifically, which interface elements)
- Allowing users to report further emotions and aspects that impacted them. We intend to increase the set of original items from the REX approach based on the results from its application in different contexts, thus making the approach more complete.
- Provide further feedback to the user on what is being evaluated and how much of the evaluation is left.

- Allowing for reporting contradicting emotional responses and better visualizing such information in the report, to allow software engineers to better understand the results.
- Avoid listing duplicated reported UX problems within the report, by combining descriptions dealing with the same interface element and/or related to the same aspect.
- Providing further statistics and graphs within the report, to facilitate its analysis.

The REX approach achieved a positive acceptance rate from both users and software engineers. Also, during the discrimination meeting, the development team showed interest in continuing applying REX for evaluating the SIGNAL application after correcting the currently identified problems. As future work, we intend to evaluate how including the improvement opportunities identified in this study can facilitate the application process of the REX approach from the point of view of users and software engineers. In the following chapter, we discuss the results of this doctoral dissertation and future work.

7 Conclusions and Future Work

This chapter presents our final remarks on our proposal, the REX approach, a UX evaluation method and a set o tools capable of providing improvement suggestions on mobile applications. The future perspectives for this work are provided, guiding the next stages of this research in the field of software quality in terms of UX.

7.1 Epilogue

This doctoral dissertation presented research on mobile applications' quality in terms of user experience. In this research, we have adopted components from mobile applications, design suggestions, and emotional responses to guide the UX evaluation process and the suggestion of improvements, yielding the development of the Redesigning for EXperience method and evaluation tool, and the REX Report Generator tool support.

For the initial proposal, evaluation, and evolution of the REX approach, this research adopted a research methodology based on empirical studies. The REX approach was proposed based on the results from a systematic mapping extension and the results from two initial exploratory studies. Through this studies, we managed to identify four main features that should be present in future technologies for the UX evaluation of mobile applications: (FI) to facilitate feedback giving, (FII) to reduce the influence of a moderator, (FIII) to make the evaluation pleasant, and (FIV) to provide suggestions to solve the identified problems.

By integrating the above features into a single approach, the Redesigning for EXperience proposal was developed. REX is an approach that aims to support users and software engineers in the UX evaluation and redesign of mobile applications. After trying to motivate and guide users during the reporting of their experiences, REX provides suggestions for software engineers to deal with the identified UX problems.

Up to the moment of writing this dissertation, the REX approach and its tool support have been evaluated through two major studies analyzing the opinion of both users and software engineers. The first study evaluated the acceptance and performance of the REX approach when compared to another UX evaluation technique (3E). REX managed to identify more specific and generic problems than the other approach while demanding the same amount of time. Additionally, the results from the TAM questionnaire suggested that REX was perceived as useful and easy to use to report an experience. Software engineers also suggested that the results from a UX evaluation using REX could be useful for guiding the redesign of the application based on the suggestions provided within the REX report. Finally, the results from applying REX in the context of a real software development project corroborated the results from the first empirical study, suggesting the acceptance of the REX approach and its usefulness for guiding users and software engineers in the evaluation and improvement of the application, respectively.

7.2 Contributions

The main contributions from this doctoral dissertation are:

- Secondary studies in the field of User eXperience evaluation, aiming at identifying the current state of the art with respect to UX evaluation methodologies. We carried out an analysis of the results from the literature review by Vermeeren et al. (2010) and a Systematic Mapping Study to characterize UX evaluation methods, which implied in:
    - The establishment of a body of knowledge on UX evaluation methods, which was used in this research.
    - The proposal of a set of features to be included in future proposals of UX evaluation methods to meet the needs of the software industry when evaluating mobile applications.
- The development of technologies for the UX evaluation of mobile applications and the suggestion of improvements based on the identified problems, which implied in:
    - The creation of a UX evaluation technique called REX that is specific for mobile applications, which is embedded into a mobile application (UX evaluation tool) generating individual UX evaluation reports.

- The development of a Desktop tool called REX Report Generator that allows generating complete UX evaluation reports and allows the visualization of the UX data and suggestions for improvement.

- Definition of empirical studies to evaluate the proposed technologies, which allowed:
  - The evaluation and evolution of the REX approach from the point of view of users and software engineers.
- Dissemination of the results of this research. During the development of this research, several research papers related to the results of this research or research collaborations have been published or been submitted to publication venues. Here we provide a list of the publications from this research or collaborations:

Research Papers from this Research

- Rivero, L. and Conte, T. (2017). Redesigning for Experience: An Approach for Guiding Users and Software Engineers in the Evaluation and Redesign of Mobile Apps. In Interacting with Computers Journal. (Submitted for Evaluation)
- Rivero, L. and Conte, T. (2017). How to Evaluate User eXperience in Software Development: A Systematic Mapping Study. In International Journal of Human-Computer Studies. (Submitted for Evaluation)
- Rivero, L. and Conte, T. (2017). Evaluating the REX Approach for the Identification of UX Problems and Solutions in Mobile Apps: A Case Study. In XXXI Brazilian Symposium on Software Engineering. (Submitted for Evaluation)
- Rivero, L. and Conte, T. (2016). How Novice Software Engineers Apply User Interface Design Suggestions: An Empirical Study. In 28th International Conference on Software Engineering & Knowledge Engineering (SEKE 2016), 600-604.
- Rivero, L. and Conte, T. (2015). Using a Study to Assess User eXperience Evaluation Methods from the Point of View of Users. In: International Conference on Enterprise Information Systems (ICEIS), 88-95.

Research Papers from Research Collaborations

- Ferreira, B., Rivero, L., Valentim, N., Zilse, R., Koster, A., and Conte, T. (2016). Evaluation of UX Methods: Lessons Learned When Evaluating a Multi-user Mobile Application. In International Conference on Human-Computer Interaction, 279-290.
- Nascimento, I., Silva, W., Lopes, A., Rivero, L., Gadelha, B., Oliveira, E. and Conte, T. (2016). An Empirical Study to Evaluate the Feasibility of a UX and Usability Inspection Technique for Mobile Applications. In 28th International Conference on Software Engineering & Knowledge Engineering (SEKE 2016), 595-599.
- Cavalcante, E., Rivero, L. and Conte, T. (2015). MAX: A Method for Evaluating the Post-use User eXperience through Cards and a Board. In Proceedings of the 27th International Conference on Software Engineering and Knowledge Engineering, 495-500.
- Cavalcante, E., Rivero, L. and Conte, T. (2015). Evaluating the Feasibility of MAX: A Method Using Cards and a Board for Assessing the Post-Use UX. International Journal of Software Engineering and Knowledge Engineering, 25(10), 1759-1764.
- Estacio, B., Valentim, N., Rivero, L., Conte, T. and Prikladnicki, R. (2015). Evaluating the Use of Pair Programming and Coding Dojo in Teaching Mockups Development: An Empirical Study. In: 48th Hawaii International Conference on System Sciences (HICSS 2015), 5084-5093.
- Rivero, L., Vincenzi, A., Maldonado, J., Conte, T. (2015). Evaluating Software Engineers' Acceptance of a Technique and Tool for Web Usability Inspection. In: 27th International Conference on Software Engineering and Knowledge Engineering (SEKE 2015), 140-145.
- Ferreira, B., Rivero, L., Lopes, A., Marques, A. and Conte, T. (2014). Apoiando o Ensino de Qualidade de Software: Um Serious Game para o Ensino de Usabilidade. In: Congresso Brasileiro de Software: Teoria e Prática - VII Fórum de Educação em Engenharia de Software, 12-21.

- Ferreira, B., Rivero, L., Lopes, A., Marques, A., and Conte, T. (2014). UsabiliCity: Um Jogo de Apoio ao Ensino de Propriedades de Usabilidade de Software Através de Analogias. In: Simpósio Brasileiro de Informática na Educação (SBIE 2014), 1273-1282.
- Kawakami, G., Rivero, L. and Conte, T. (2014). Identificando Oportunidades de Melhoria em Tecnologias de Inspeção de Usabilidade para o Contexto Web Móvel através de Estudos Experimentais. In: 11th Workshop en Ingeniería del Software Experimental (ESELAW 2014), 336-349.
- Rivero, L. and Conte, T. (2014). Web DUE: Uma Técnica de Inspeção de Usabilidade de Mockups de Aplicações Web Guiada Por Zonas Próprias de Páginas Web. In: XIII Simpósio Brasileiro de Qualidade de Software, Concurso de Dissertações e Teses, 325-339.
- Rivero, L.; Kalinowski, M. and Conte, T. (2014). Practical Findings from Applying Innovative Design Usability Evaluation Technologies for Mockups of Web Applications. In: 47th Annual Hawaii International Conference on System Sciences, 3054-3063.
- Rivero, L.; Kawakami, G. and Conte, T. (2014). Using a Controlled Experiment to Evaluate Usability Inspection Technologies for Improving the Quality of Mobile Web Applications Earlier in their Design. In: 28th Brazilian Symposium on Software Engineering, 161-170.
- Rivero, L.; Valle, R. and Conte, T. (2014). Aplicando Design e Avaliação de Usabilidade para Melhorar a Qualidade de um Aplicativo Web Móvel. In: XIII Simpósio Brasileiro de Qualidade de Software (SBQS 2014), 247-254.

7.3 Future Perspectives

The development of this research allowed us to develop a set of UX evaluation technologies for mobile applications, including the REX approach and the REX Report Generator. The results from this research provide new perspectives that can be explored in future work. The main future work is detailed as follows:

7.3.1 Evolution of the Proposed Technologies

As shown in Chapter 6, the results from the observational study where we evaluated the REX approach motivated the development of further functionalities. For instance, it may be necessary to include alternatives for allowing the users to report the exact location (interface element) of the UX problem, while also allowing the REX Report Generator to count duplicated UX problems based on their descriptions. Furthermore, providing means to aggregate the data to provide further information could be done in the REX Report Generator. In this sense, future work could be to allow viewing which users pointed contradicting emotions, allowing software engineers to understand the results better.

Another possible evolution could be developed based on the analysis of the results from each UX evaluation. New problems and improvement suggestions could be incorporated into the REX approach through a collaborative network. Thus, it could be interesting to develop an administrator mode for the REX approach, in order to allow researchers and software engineers to update the list of identified UX problems and suggestions that are proposed within the items from the REX approach. By doing so, software engineers could have access to an updated catalog of items and suggestions. These future work still needs to be discussed, in order to address what are the positive and negative aspects of allowing software development teams to suggest items into the REX approach, as this may impact the difficulty of the approach depending on the type of and a number of items that are suggested.

7.3.2 Further Evaluations

Despite having carried two major empirical studies from the point of view of both users and software engineers to evaluate the REX approach, several new evaluations may be performed. First of all, the performed studies (both exploratory and the studies evaluating the REX approach) could be replicated, while increasing the number of subjects, mainly in real software development scenarios, in order to increase the statistical significance of the results.

Additionally, it may be necessary to continue the analysis of the qualitative data following the procedures proposed in Grounded Theory research (Glaser and Strauss, 2009). By doing so, the theoretical saturation could be reached, increasing our

understanding of the aspects that affect the applicability of the REX approach and its improvement opportunities.

The execution of the empirical studies of the REX approach suggested further paths of investigation in the field of UX evaluation. For instance, it is necessary to verify to what extent the discussion meeting of the report could be useful for generating further ideas in the improvement of the evaluated application, and if a specific process could be proposed for enhancing the generation of ideas.

Another aspect to be investigated could be the effect of the REX approach in the development team regarding the empathy that they create with users. The empirical studies suggested that software engineers understood the UX problems and were interested in finding solutions that could engage users in the use of their products. It is necessary to verify to what extent the REX approach can support such empathy creation process, its effect in the correction of the identified problems, and how software engineers handle future projects, after participating in such meetings.

Finally, future studies comparing the REX approach with other UX evaluation technologies (or the combination of two or more technologies) could be performed. The reporting of empirical studies indicating the specific contexts in which our proposal or other technologies perform better could be useful for practitioners in choosing the most appropriate technology while suggesting future work for researchers in the field of software quality in terms of UX.

We argue that our findings will be useful for the promotion and improvement of the current practice and research of UX. Also, we expect that by suggesting the REX approach, software engineers can assess and improve the quality of mobile applications, meeting users' expectations.

References

Almahmoud, J., Almalki, A., Alrashed, T. and Alwabil, A. (2016). Prototyping Complex Systems: A Diary Study Approach to Understand the Design Process. In International Conference of Design, User Experience, and Usability, 187-196.

Balters, S. and Steinert, M. (2015). Capturing emotion reactivity through physiology measurement as a foundation for affective engineering in engineering design science and engineering practices. Journal of Intelligent Manufacturing, 1-23.

Bargas-Avila, J. and Hornbæk, K. (2011). Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2689-2698.

Bargas-Avila, J. and Hornbæk, K. (2012). Foci and blind spots in user experience research. interactions, 19(6), 24-27.

Bevan, N. (2008). Classifying and selecting UX and usability measures. In International Workshop on Meaningful Measures: Valid Useful User Experience Measurement, 13-18.

Birnbaum, Z. (1956). On a use of the Mann-Whitney statistic. In Proceedings of the third Berkeley symposium on mathematical statistics and probability, 13-17.

Borchers, J. (2001). A pattern approach to interaction design. Ai & Society, 15(4), 359-376.

Budgen, D. (2013). Design Patterns: Magic or Myth?. IEEE software, (2), 87-90.

Campos, G., Bandeira, A. and Sardi, S. (2015). Zika virus outbreak, Bahia, Brazil. Emerging infectious diseases, 21(10), 1885.

Carver, J., Jaccheri, L., Morasca, S. and Shull, F. (2003). Issues in using students in empirical studies in software engineering education. In Software Metrics Symposium, 2003. Proceedings. Ninth International, 239-249.

Chen, T., Pan, Z. and Zheng, J. (2008). EasyMall - An Interactive Virtual Shopping System. In Fuzzy Systems and Knowledge Discovery, 2008. FSKD'08, 669-673.

Chung, E., Hong, J., Lin, J., Prabaker, M., Landay, J. and Liu, A. (2004). Development and evaluation of emerging design patterns for ubiquitous computing. In

Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques, 233-242.

Cockburn, A. (2008) Using Both Incremental and Iterative Development, STSC CrossTalk (USAF Software Technology Support Center), 21(5), 27–30.

Corder, G. and Foreman, D. (2014). Nonparametric statistics: A step-by-step approach. John Wiley & Sons.

de la Vara, J., Wnuk, K., Berntsson-Svensson, R., Sánchez, J. and Regnell, B. (2011). An Empirical Study on the Importance of Quality Requirements in Industry. In SEKE, 438-443.

Desmet, P. (2005). Measuring emotion: Development and application of an instrument to measure emotional responses to products. In Funology, Springer Netherlands, 111-123.

Desmet, P., Overbeeke, K. and Tax, S. (2001). Designing products with added emotional value: Development and appllcation of an approach for research through design. The design journal, 4(1), 32-47.

DiStefano, C., Zhu, M. and Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. Practical Assessment, Research & Evaluation, 14(20), 1-11.

Dix, A., Finlay, J., Abowd, G. and Beale, R. (1998). Human-Computer Interaction, Prentice - Hall, London, 1327-1331.

Duyne, D., Landay, J., and Hong, J. (2002). The design of sites: patterns, principles, and processes for crafting a customer-centered Web experience. Addison-Wesley Longman Publishing.

Ervasti, M., Dashti, S., Reilly, J., Bray, J., Bayen, A. and Glaser, S. (2011). iShake: mobile phones as seismic sensors--user study findings. In Proceedings of the 10th international Conference on Mobile and Ubiquitous Multimedia, 43-52.

Fernandez, A., Insfran, E. and Abrahão, S. (2011). Usability evaluation methods for the web: A systematic mapping study. Information and Software Technology, 53(8), 789-817.

Frey, J., Mühl, C., Lotte, F. and Hachet, M. (2013). Review of the use of electroen-cephalography as an evaluation method for human-computer interaction. arXiv preprint arXiv:1311.2222.

Gegner, L. and Runonen, M. (2012). For What it is Worth: Anticipated eXperience Evaluation. 8th International Conference on Design and Emotion.

Glaser, B. and Strauss, A. (2009). The discovery of grounded theory: Strategies for qualitative research. Transaction Publishers.

Google. (2016) Android Design Patterns. https://www.google.com/design/spec/patterns

GWI (2014) GWI Device Summary. Available at: http://quantifiedself.com/docs/GWI_Device_Summary_Q3_2014.pdf

Hassenzahl, M. (2005). The thing and I: understanding the relationship between user and product. In Funology, Springer Netherlands, 31-42.

Hassenzahl, M. and Tractinsky, N. (2006). User experience-a research agenda. Behaviour & information technology, 25(2), 91-97.

Hassenzahl, M., Diefenbach, S. and Göritz, A. (2010). Needs, affect, and interactive products–Facets of user experience. Interacting with computers, 22(5), 353-362.

Hearst, M., Laskowski, P. and Silva, L. (2016). Evaluating information visualization via the interplay of heuristic evaluation and question-based scoring. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 5028-5033.

Hornbæk, K., and Frøkjær, E. (2005). Comparing usability problems and redesign proposals as input to practical systems development. In Proceedings of the SIGCHI conference on Human factors in computing systems, 391-400.

Hu, N., Pavlou, P. and Zhang, J. (2006). Can online reviews reveal a product's true quality?: empirical findings and analytical modeling of Online word-of-mouth communication. In Proceedings of the 7th ACM conference on Electronic commerce, 324-330.

Hvannberg, E., Law, E. and Lérusdóttir, M. (2007). Heuristic evaluation: Comparing ways of finding and reporting usability problems. Interacting with computers, 19(2), 225-240.

International Organization for Standardization, ISO DIS 9241-210 (2010). Ergonomics of human system interaction - Part 210: Human-centered design for interactive systems (formerly known as 13407). International Standardization Organization (ISO).

International Organization for Standardization, ISO/IEC 25010 (2011). Systems and software engineering -- SQuaRE - Software product Quality Requirements and Evaluation -- System and software quality models.

Isbister, K., Höök, K., Sharp, M. and Laaksolahti, J. (2006). The sensual evaluation instrument: developing an affective evaluation tool. In Proceedings of the SIGCHI conference on Human Factors in computing systems, 1163-1172.

Isomursu, M., Kuutti, K. and Väinämö, S. (2004). Experience clip: method for user participation and evaluation of mobile concepts. In Proceedings of the eighth conference on Participatory design: Artful integration: interweaving media, materials and practices, 83-92.

Isomursu, M., Tähti, M., Väinämö, S. and Kuutti, K. (2007). Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. International Journal of Human-Computer Studies, 65(4), 404-418.

Jordan, P. (2003). Designing pleasurable products: An introduction to the new human factors, CRC Press.

Juristo, N., Moreno, A., Sanchez-Segura, M. and Baranauskas, M. (2007). A glass box design: making the impact of usability on software development visible. In Human-Computer Interaction–INTERACT 2007, 541-554.

Khalid, H., Shihab, E., Nagappan, M. and Hassan, A. (2014). What do mobile app users complain about? A study on free iOS apps. In: IEEE Software, 1(1), 2014, 70-77.

Kircher, M. and Völter, M. (2007). Guest Editors' Introduction: Software Patterns. IEEE software, (4), 28-30.

Koukouletsos, K., Khazaei, B., Dearden, A. and Tseles, D. (2006). Comparing patterns and guidelines in web design. In Proceedings of the 1st International Scientific Conference eRA.

Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E. and Sinnelä, A. (2011). UX Curve: A method for evaluating long-term user experience. Interacting with Computers, 23(5), 473-483.

Lanzilotti, R., Ardito, C., Costabile, M. and De Angeli, A. (2011). Do patterns help novice evaluators? A comparative study. International journal of human-computer studies, 69(1), 52-69.

Law, E. (2011). The measurability and predictability of user experi-ence. In Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems, 1-10.

Lin, J., Abi-Rached, H. and Lahav, M. (2004). Virtual guiding avatar: An effective procedure to reduce simulator sickness in virtual environments. In Proceedings of the SIGCHI conference on Human factors in computing systems, 719-726.

Mafra, S., Barcelos, R. and Travassos, G. (2006). Aplicando uma metodologia baseada em evidência na definiçao de novas tecnologias de software. In Proceedings of the 20th Brazilian Symposium on Software Engineering (SBES 2006), 239-254. (In Portuguese)

McDonnell, T., Ray, B. and Kim, M. (2013). An empirical study of api stability and adoption in the android ecosystem. In 29th IEEE International Conference on Software Maintenance (ICSM), 70-79.

Mendes, E., Mosley, N. and Counsell, S. (2006). The need for web engineering: an introduction. In Web Engineering, Springer Berlin Heidelberg, 1-27.

Mu, H., and Jiang, S. (2011). Design patterns in software development. In 2011 IEEE 2nd International Conference on Software Engineering and Service Science.

Neil, T. (2014). Mobile design pattern gallery: UI patterns for smartphone apps. " O'Reilly Media, Inc.".

Nielsen, J. (1994). Heuristic evaluation. Usability inspection methods, 17(1), 25-62.).

Rajeshkumar, S., Omar, R. and Mahmud, M. (2013). Taxonomies of user experience (UX) evaluation methods. In 2013 International Conference on Research and Innovation in Information Systems (ICRIIS), 533-538.

Read, J. and MacFarlane, S. (2006). Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In Proceedings of the 2006 conference on Interaction design and children, 81-88.

Ricca, F., Scanniello, G., Torchiano, M., Reggio, G. and Astesiano, E. (2010). On the effort of augmenting use cases with screen mockups: results from a preliminary empirical study. In Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, 40.

Rivero, J., Rossi, G., Grigera, J., Burella, J., Luna, E. and Gordillo, S. (2010). From mockups to user interface models: an extensible model driven approach. In International Conference on Web Engineering, 13-24.

Rivero, L. and Conte, T. (2015). Using a Study to Assess User eXperience Evaluation Methods from the Point of View of Users. In Proceedings of the 17th International Conference on Enterprise Information Systems, 88-95.

Rivero, L. and Conte, T. (2016). TR-USES-2016-002 Systematic Mapping on UX Technologies: Extraction of User eXperience Evaluation Methods. USES Research Group Technical Report. Available at: https://uses.induspam.com/relatorios-tecnicos/

Ross, P., and Keyson, D. (2007). The case of sculpting atmospheres: towards design principles for expressive tangible interaction in control of ambient systems. Personal and Ubiquitous Computing, 11(2), 69-79.

Sarwar, M., and Soomro, T. (2013). Impact of Smartphone's on Society. European Journal of Scientific Research, 98(2), 216-226.

Scherer, K. (2005). What are emotions? And how can they be measured?. Social science information, 44(4), 695-729.

Seffah, A. (2010). The evolution of design patterns in HCI: from pattern languages to pattern-oriented design. In Proceedings of the 1st International Workshop on Pattern-Driven Engineering of Interactive Computing Systems, 4-9.

Shapiro, S. and Wilk, M. (1965). An analysis of variance test for normality (complete samples). Biometrika, 591-611.

Shull, F., Carver, J. and Travassos, G. (2001). An empirical methodology for introducing software processes. ACM SIGSOFT Software Engineering Notes, 26(5), 288-296.

Squires, D. and Preece, J. (1999). Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. Interacting with computers, 11(5), 467-483.

Tähti, M. and Niemelä, M. (2006). 3e–expressing emotions and experiences. In Proc. WP9 Workshop on Innovative Approaches for Evaluating Affective Systems. HUMAINE (Human-Machine Interaction Network on Emotion).

Tidwell, J. (1998). Interaction Design Patterns. PLOP'98 Conference on Pattern Languages of Programming. Available at: http://www.mit.edu/~jtidwell/interaction_patterns.html

Travassos, G., Shull, F., Fredericks, M. and Basili, V. (1999). Detecting defects in object-oriented designs: using reading techniques to increase software quality. In ACM Sigplan Notices, 34(10), 47-56.

Väätäjä, H., Koponen, T. and Roto, V. (2009). Developing practical tools for user experience evaluation: a case from mobile news journalism. In European Conference on Cognitive Ergonomics: Designing beyond the Product---Understanding Activity and User Experience in Ubiquitous Environments, 23.

Vermeeren, A., Kort, J., Cremers, A. and Fokker, J. (2008). Comparing UX measurements, a case study. In Proceedings of the International Workshop on Meaningful Measures: Valid Useful Experience Measurement, 72-78.

Vermeeren, A., Law, E., Roto, V., Obrist, M., Hoonhout, J. and Väänänen-Vainio-Mattila, K. (2010). User experience evaluation methods: current state and development needs. In Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, 521-530.

Welie, M. (2011). A Pattern Library for Interaction Design. Available at: http://www.welie.com/patterns/index.html.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M., Regnell, B. and Wesslén, A. (2012). Experimentation in software engineering. Springer Science & Business Media.

Appendix A: References from the Review on UX Evaluation Technologies – Initial Review

[S1] Schubert, E. (1999). Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space. Australian Journal of Psychology, 51(3), 154-165.

[S2] Tähti, M. and Niemelä, M. (2006). 3e–expressing emotions and experiences. In Proc. WP9 Workshop on Innovative Approaches for Evaluating Affective Systems. HUMAINE (Human-Machine Interaction Network on Emotion).

[S3] Lavie, T. and Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. International journal of human-computer studies, 60(3), 269-298.

[S4] Russell, J., Weiss, A. and Mendelsohn, G. (1989). Affect grid: a single-item scale of pleasure and arousal. Journal of personality and social psychology, 57(3), 493.

[S5] Ståhl, A., Höök, K., Svensson, M., Taylor, A. and Combetto, M. (2009). Experiencing the affective diary. Personal and Ubiquitous Computing, 13(5), 365-378.

[S6] Hassenzahl, M., Burmester, M. and Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In Mensch & Computer 2003, Vieweg+ Teubner Verlag, 187-196.

[S7] Väätäjä, H., Koponen, T. and Roto, V. (2009). Developing practical tools for user experience evaluation: a case from mobile news journalism. In European Conference on Cognitive Ergonomics: Designing beyond the Product---Understanding Activity and User Experience in Ubiquitous Environments, VTT Technical Research Centre of Finland, 23.

[S8] Gegner, L. and Runonen, M. (2012). For What it is Worth: Anticipated eXperience Evaluation. 8th International Conference on Design and Emotion.

[S9] Jordan, P. (2002). Designing pleasurable products: An introduction to the new human factors. CRC press.

[S10] Froehlich, J., Chen, M., Consolvo, S., Harrison, B. and Landay, J. (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, 57-70.

[S11] Abeele, V. and Zaman, B. (2009). Laddering the User Experience! In User Experience Evaluation Methods in Product Development (UXEM'09)-Workshop.

[S12] Kahneman, D., Krueger, A., Schkade, D., Schwarz, N. and Stone, A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. Science, 306(5702), 1776-1780.

[S13] Boyle, G. (1984). Reliability and validity of Izard's differential emotions scale. Personality and Individual Differences, 5(6), 747-750.

[S14] Laurans, G. and Desmet, P. (2006). Using self-confrontation to study user experience: A new approach to the dynamic measurement of emotions while interacting with products. In Design & Emotion.

[S15] Desmet, P., Overbeeke, K. and Tax, S. (2001). Designing products with added emotional value: Development and application of an approach for research through design. The design journal, 4(1), 32-47.

[S16] Hole, L. and Williams, O. (2007). The emotion sampling device (ESD). In Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it, 177-178.

[S17] Isomursu, M., Kuutti, K. and Väinämö, S. (2004). Experience clip: method for user participation and evaluation of mobile concepts. In Proceedings of the eighth conference on Participatory design: Artful integration: interweaving media, materials and practices, 83-92.

[S18] Larson, R. and Csikszentmihalyi, M. (1983). The experience sampling method. In H. T. Reis (Ed.), Naturalistic Approaches to Studying Social Interaction. New Directions for Methodology of Social and Behavioral Science, 41-56.

[S19] Den Uyl, M. and Van Kuilenburg, H. (2005). The FaceReader: Online facial expression recognition. In Proceedings of Measuring Behavior, 589-590.

[S20] Cowie, R., Douglas-Cowie, E., Savvidou, S., McMahon, E., Sawey, M. and Schröder, M. (2000). 'FEELTRACE': An instrument for recording perceived emotion in real time. In ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion.

[S21] Read, J. and MacFarlane, S. (2006). Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In Proceedings of the 2006 conference on Interaction design and children, pp. 81-88.

[S22] IJsselsteijn, W., van den Hoogen, W., Klimmt, C., de Kort, Y., Lindley, C., Mathiak, K. and Vorderer, P. (2008). Measuring the experience of digital game enjoyment. In Proceedings of Measuring Behavior, 88-89.

[S23] Geneva Emotion Research Group. (2002). Geneva Appraisal Questionnaire (GAQ): Format, Development, and Utilization.

[S24] Scherer, K. (2005). What are emotions? And how can they be measured?. Social science information, 44(4), 695-729.

[S25] Obrist, M., Moser, C., Alliez, D., Holocher, T. and Tscheligi, M. (2009). Connecting TV & PC: an in-situ field evaluation of an unified electronic program guide concept. In Proceedings of the seventh european conference on European interactive television conference, 91-100.

[S26] Van der Heijden, H. and Sørensen, L. (2002). Measuring attitudes towards mobile information services: an empirical validation of the HED/UT scale. Technical University of Denmark, Center for Tele-Information, 765-777.

[S27] Madsen, M. and Gregor, S. (2000). Measuring human-computer trust. In Proceedings of Eleventh Australasian Conference on Information Systems, 6-8.

[S28] Ryan, R., Connell, J. and Plant, R. (1990). Emotions in nondirected text learning. Learning and Individual Differences, 2(1), 1-17.

[S29] Karapanos, E., Zimmerman, J., Forlizzi, J. and Martens, J. (2010). Measuring the dynamics of remembered experience over time. Interacting with Computers, 22(5), 328-335.

[S30] Schütte, S. (2002). Designing feelings into products: Integrating kansei engineering methodology in product development.

[S31] Abowd, G., Atkeson, C., Bobick, A., Essa, I., MacIntyre, B., Mynatt, E. and Starner, T. (2000). Living laboratories: the future computing environments group at the Georgia Institute of Technology. In CHI'00 extended abstracts on Human factors in computing systems, 215-216.

[S32] Bolger, N., Davis, A. and Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. Annual review of psychology, 54(1), 579-616.

[S33] De Waard, D. (1996). The measurement of drivers' mental workload. Groningen University, Traffic Research Center.

[S34] Bakker, S., Markopoulos, P. and de Kort, Y. (2008). OPOS: an observation scheme for evaluating head-up play. In Proceedings of the 5th Nordic Conference on Human-Computer interaction: Building Bridges, 33-42.

[S35] Mehrabian, A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. Current Psychology, 14(4), 261-292.

[S36] Herman, M. and Koczkodaj, W. (1996). A Monte Carlo study of pairwise comparison. Information Processing Letters, 57(1), 25-29.

[S37] Ward, R. and Marsden, P. (2003). Physiological responses to different web page designs. International Journal of Human-Computer Studies, 59(1), 199-212.

[S38] Desurvire, H., Caplan, M. and Toth, J. (2004). Using heuristics to evaluate the playability of games. In CHI'04 extended abstracts on Human factors in computing systems, 1509-1512.

[S39] Watson, D., Clark, L. and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. Journal of personality and social psychology, 54(6), 1063-1070.

[S40] Desmet, P. (2005). Measuring emotion: Development and application of an instrument to measure emotional responses to products. In Funology, Springer Netherlands, 111-123.

[S41] Witmer, B. and Singer, M. (1998). Measuring presence in virtual environments: A presence questionnaire. Presence: Teleoperators and virtual environments, 7(3), 225-240.

[S42] Mugge, R., Schifferstein, H. and Schoormans, J. (2006). A longitudinal study on product attachment and its determinants. European advances in consumer research, 7, 641-647.

[S43] Mandryk, R., Inkpen, K. and Calvert, T. (2006) Using psychophysiological techniques to measure user experience with entertainment technologies. Behaviour and Information Technology, 25(2), 141-158.

[S44] Fallman, D. and Waterworth, J. (2005). Dealing with user experience and affective evaluation in hci design: A repertory grid approach. In Workshop Paper, CHI, 2-7.

[S45] Bradley, M. and Lang, P. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry, 25(1), 49-59.

[S46] Isbister, K., Höök, K., Sharp, M. and Laaksolahti, J. (2006). The sensual evaluation instrument: developing an affective evaluation tool. In Proceedings of the SIGCHI conference on Human Factors in computing systems, 1163-1172.

[S47] Kujala, S., Walsh, T., Nurkka, P. and Crisan, M. (2014). Sentence Completion for Understanding Users and Evaluating User Experience. Interacting with Computers, 26(3), 238-255.

[S48] Kirakowski, J., and Corbett, M. (1993). SUMI - The Software Usability Measurement Inventory. British Journal of Educational Technology, 24 (3), 210-212

[S49] Kim, J., Gunn, D., Schuh, E., Phillips, B., Pagulayan, R. and Wixon, D. (2008). Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, 443-452.

[S50] Vermeeren, A., Kort, J., Cremers, A. and Fokker, J. (2008). Comparing UX measurements, a case study. In Proceedings of the International Workshop on Meaningful Measures: Valid Useful Experience Measurement, 72-78.

[S51] Venkatesh, V., Morris, M., Davis, G. and Davis, F. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.

[S52] Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E. and Sinnelä, A. (2011). UX Curve: A method for evaluating long-term user experience. Interacting with Computers, 23(5), 473-483.

[S53] Väänänen-Vainio-Mattila, K., and Wäljas, M. (2009). Developing an expert evaluation method for user eXperience of cross-platform web services. In Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era, 162-169.

[S54] Burmester, M., Mast, M., Jäger, K. and Homans, H. (2010). Valence method for formative evaluation of user experience. In Proceedings of the 8th ACM conference on Designing Interactive Systems, 364-367.

[S55] Kirakowski, J., Claridge, N. and Whitehand, R. (1998). Human centered measures of success in web site design. In Proceedings of the Fourth Conference on Human Factors & the Web.

[S56] Lucero, A. (2009). Co-designing interactive spaces for and with designers: supporting mood-board making. Eindhoven, the Netherlands: Eindhoven University of Technology.

[S57] Sproll, S., Peissner, M. and Sturm, C. (2010). From product concept to user experience: exploring UX potentials at early product stages. In Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, 473-482.

[S58] Wimmer, B., Wöckl, B., Leitner, M. and Tscheligi, M. (2010). Measuring the dynamics of user experience in short interaction sequences. In Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, 825-828.

[S59] Sainz, F., Madrid, R., and Madrid, J. (2010). Introducing co-design for digital technologies in rural areas. In Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, 769-772.

[S60] Vermeulen, I., Roth, C., Vorderer, P. and Klimmt, C. (2010). Measuring user responses to interactive stories: Towards a standardized assessment tool. In Interactive Storytelling, 38-43.

[S61] Duarte, L., de Sá, M. and Carriço, L. (2010). Physiological data gathering in mobile environments. In Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing-Adjunct, 405-406.

[S62] Pu, P., Chen, L. and Hu, R. (2010). A user-centric evaluation framework for recommender systems. In Proceedings of the fifth ACM conference on Recommender systems, 157-164.

[S63] Knijnenburg, B., Meesters, L., Marrow, P. and Bouwhuis, D. (2009). User-centric evaluation framework for multimedia recommender systems. In User Centric Media, 366-369.

[S64] Liu, W., Huang, D. and Zhang, Y. (2010). Research on fuzzy comprehensive evaluation of user experience. In 2010 IEEE Youth Conference on Information Computing and Telecommunications (YC-ICT), 122-125.

[S65] Mashapa, J. and van Greunen, D. (2010). User experience evaluation metrics for usable accounting tools. In Proceedings of the 2010 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists, 170-181.

[S66] Arrasvuori, J., Korhonen, H. and Väänänen-Vainio-Mattila, K. (2010). Exploring playfulness in user experience of personal mobile products. In Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction, 88-95.

[S67] Vanattenhoven, J. and De Roeck, D. (2009). Investigating the Use of the Experience Clip Method. In User Centric Media, 79-86.

[S68] Lew, P., Olsina, L. and Zhang, L. (2010). Integrating quality, quality in use, actual usability and user experience. In 6th Central and Eastern European Software Engineering Conference (CEE-SECR), 117-123.

[S69] Al-Wabil, A., Alabdulqader, E., Al-Abdulkarim, L. and Al-Twairesh, N. (2010). Measuring the user experience of digital books with children: An eyetracking study of interaction with digital libraries. In International Conference for Internet Technology and Secured Transactions (ICITST), 1-7.

[S70] Burmester, M., Mast, M., Jäger, K. and Homans, H. (2010). Valence method for formative evaluation of user experience. In Proceedings of the 8th ACM conference on Designing Interactive Systems, 364-367.

[S71] Wu, W. and Tang, M. (2010). Assessing the affective impact of social navigation tools in facilitating exploratory search. Proceedings of the American Society for Information Science and Technology, 47(1), 1-2.

[S72] Lew, P., Zhang, L., and Olsina, L. (2010). Usability and user experience as key drivers for evaluating GIS application quality. In 18th International Conference on Geoinformatics, 1-6.

[S73] Drachen, A., Nacke, L., Yannakakis, G. and Pedersen, A. (2010). Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games, 49-54.

[S74] Finstad, K. (2010). The usability metric for user experience. Interacting with Computers, 22(5), 323-327.

[S75] Karapanos, E., Zimmerman, J., Forlizzi, J. and Martens, J. (2010). Measuring the dynamics of remembered experience over time. Interacting with Computers, 22(5), 328-335.

[S76] Funk, M., Rozinat, A., Karapanos, E., de Medeiros, A. and Koca, A. (2010). In situ evaluation of recommender systems: Framework and instrumentation. International journal of human-computer studies, 68(8), 525-547.

[S77] Zaman, B. and Abeele, V. (2010). Laddering with young children in User eXperience evaluations: theoretical groundings and a practical case. In Proceedings of the 9th International Conference on Interaction Design and Children, 156-165.

[S78] Rodden, K., Hutchinson, H., and Fu, X. (2010). 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, 2395-2398.

[S79] Väätäjä, H. and Roto, V. (2010). Mobile questionnaires for user experience evaluation. In CHI'10 Extended Abstracts on Human Factors in Computing Systems, 3361-3366.

[S80] Korhonen, H., Arrasvuori, J. and Väänänen-Vainio-Mattila, K. (2010). Let users tell the story: evaluating user experience with experience reports. In CHI'10 Extended Abstracts on Human Factors in Computing Systems, 4051-4056.

[S81] Barnum, C. and Palmer, L. (2010). More than a feeling: understanding the desirability factor in user experience. In CHI'10 Extended Abstracts on Human Factors in Computing Systems, 4703-4716.

[S82] Karapanos, E., Martens, J. and Hassenzahl, M. (2010). On the retrospective assessment of users' experiences over time: memory or actuality?. In CHI'10 Extended Abstracts on Human Factors in Computing Systems, 4075-4080.

[S83] Petrie, H. and Precious, J. (2010). Measuring User Experience of websites: Think aloud protocols and an emotion word prompt list. In CHI'10 Extended Abstracts on Human Factors in Computing Systems, 3673-3678.

[S84] De Moor, K., Ketyko, I., Joseph, W., Deryckere, T., De Marez, L., Martens, L. and Verleye, G. (2010). Proposed framework for evaluating quality of experience in a mobile, testbed-oriented living lab setting. Mobile Networks and Applications, 15(3), 378-391.

[S85] Brooks, P. and Hestnes, B. (2010). User measures of quality of experience: why being objective and quantitative is important. Network, IEEE, 24(2), 8-13.

[S86] Nordahl, R. and Korsgaard, D. (2010). Distraction as a measure of presence: using visual and tactile adjustable distraction as a measure to determine immersive presence of content in mediated environments. Virtual reality, 14(1), 27-42.

[S87] O'Brien, H. and Toms, E. (2010). The development and evaluation of a survey to measure user engagement. Journal of the American Society for Information Science and Technology, 61(1), 50-69.

[S88] Knijnenburg, B., Willemsen, M. and Hirtbach, S. (2010). Receiving recommendations and providing feedback: The user-experience of a recommender system. In E-Commerce and Web Technologies, 207-216.

[S89] Bach, C., Gauducheau, N., and Salembier, P. (2011). Combining interviews and scales in the multidimensional evaluation of user experience: a case study in 3D games. In Proceedings of the 29th Annual European Conference on Cognitive Ergonomics, 157-160.

[S90] Pu, P., Chen, L. and Hu, R. (2011). A user-centric evaluation framework for recommender systems. In Proceedings of the fifth ACM conference on Recommender systems, 157-164.

[S91] Knijnenburg, B., Willemsen, M. and Kobsa, A. (2011). A pragmatic procedure to support the user-centric evaluation of recommender systems. In Proceedings of the fifth ACM conference on Recommender systems, 321-324.

[S92] Chen, Z. and Zhu, S. (2011). The research of mobile application user experience and assessment model. In International Conference on Computer Science and Network Technology (ICCSNT), 2832-2835.

[S93] Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K. and Sinnelä, A. (2011). Identifying hedonic factors in long-term user experience. In Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces, 17.

[S94] Mirza-Babaei, P., Long, S., Foley, E. and McAllister, G. (2011). Understanding the contribution of biometrics to games user research. In Proc. DiGRA, 329-347.

[S95] Wan, A., Sadiq, S. and Li, X. (2011). Exploratory Study on Learners' Experience in an eLearning System. In AMCIS.

[S96] Ervasti, M., Dashti, S., Reilly, J., Bray, J., Bayen, A. and Glaser, S. (2011). iShake: mobile phones as seismic sensors--user study findings. In Proceedings of the 10th international Conference on Mobile and Ubiquitous Multimedia, 43-52.

[S97] Gürkök, H., Hakvoort, G., Poel, M. and Nijholt, A. (2011). User expectations and experiences of a speech and thought controlled computer game. In Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology, 53.

[S98] Roth, C., Klimmt, C., Vermeulen, I. and Vorderer, P. (2011). The experience of interactive storytelling: comparing “Fahrenheit” with “Façade”. In Entertainment Computing–ICEC 2011, 13-21.

[S99] Mulwa, C., Lawless, S., Sharp, M. and Wade, V. (2011). A Web-Based Framework for User-Centred Evaluation of End-User Experience in Adaptive and Personalized E-Learning Systems. In Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, 351-356.

[S100] Pasch, M. and Landoni, M. (2011). Recognizing bodily expression of affect in user tests. In Affective Computing and Intelligent Interaction, 264-271.

[S101] Castro, L., Favela, J. and García-Peña, C. (2011). Naturalistic enactment to stimulate user experience for the evaluation of a mobile elderly care application. In Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, 371-380.

[S102] Nacke, L., Stellmach, S. and Lindley, C. (2010). Electroencephalographic assessment of player experience: A pilot study in affective ludology. Simulation & Gaming, 42(5), 632-655

[S103] Othman, M., Petrie, H. and Power, C. (2011). Engaging visitors in museums with technology: scales for the measurement of visitor and multimedia guide experience. In Human-Computer Interaction–INTERACT 2011, 92-99.

[S104] Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E. and Sinnelä, A. (2011). UX Curve: A method for evaluating long-term user experience. Interacting with Computers, 23(5), 473-483.

[S105] Némery, A., Brangier, E., and Kopp, S. (2011). First validation of persuasive criteria for designing and evaluating the social influence of user interfaces: justification of a guideline. In Design, user experience, and usability. theory, methods, tools and practice, 616-624.

[S106] Kolski, C., Uster, G., Robert, J. M., Oliveira, K. and David, B. (2011). Interaction in mobility: the evaluation of interactive systems used by travellers in transportation contexts. In Human-Computer Interaction. Towards Mobile and Intelligent Interaction Environments, 301-310.

[S107] Wu, Z., Lin, T., and Tang, N. (2011). How do visual angle and physical screen size affect presence and emotional responses of game players?. In International Conference on Consumer Electronics, Communications and Networks (CECNet), 4128-4131.

[S108] Freeman, B. (2011). Triggered think-aloud protocol: using eye tracking to improve usability test moderation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1171-1174.

[S109] Szostek, A., and Karapanos, E. (2011). Introducing VERO: visual experiential requirements organizer. In CHI'11 Extended Abstracts on Human Factors in Computing Systems, 2101-2106.

[S110] Isomursu, M., Ervasti, M., Kinnula, M. and Isomursu, P. (2011). Understanding human values in adopting new technology—a case study and methodological discussion. International journal of human-computer studies, 69(4), 183-200.

[S111] Schulze, K. and Krömker, H. (2010). A framework to measure user experience of interactive online products. In Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research, 14.

[S112] Felismino, M., Luiz, R., Castela, M. and Marques, P. (2010). Optimization of user satisfaction on an augmented reality based guided tour. In Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research, 43.

[S113] Wu, O., Chen, Y., Li, B. and Hu, W. (2011). Evaluating the visual quality of web pages using a computational aesthetic approach. In Proceedings of the fourth ACM international conference on Web search and data mining, 337-346.
[S114] Oksanen, K. and Hämäläinen, R. (2011). Assessing Game Experiences Caused by Educational Collaborative Game. In Proceedings of the 7th European Conference on Management Leadership and Governance: ECGBL 2011, 431.
[S115] Van Schaik, P. and Ling, J. (2011). An integrated model of interaction experience for information retrieval in a Web-based encyclopaedia. Interacting with Computers, 23(1), 18-32.
[S116] Göbel, S., Gutjahr, M. and Steinmetz, R. (2011). What Makes a Good Serious Game–Conceptual Approach Towards a Metadata Format for the Description and Evaluation of Serious Games. In Proceedings of the European Conference on Games Based Learning. Academic Conferences, 202-210.
[S117] Mirza-Babaei, P. (2011). Biometrics to improve methodologies on understanding player's gameplay experience. In Proceedings of the 25th BCS Conference on Human-Computer Interaction, 546-549.
[S118] Gross, A. and Bongartz, S. (2012). Why do I like it?: investigating the product-specificity of user experience. In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design, 322-330.
[S119] Varsaluoma, J. and Kentta, V. (2012). DrawUX: web-based research tool for long-term user experience evaluation. In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design, 769-770.
[S120] Pitkänen, J., Pitkäranta, M. and Nieminen, M. (2012). Usability testing in real context of use: the user-triggered usability testing. In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design, 797-798.
[S121] Moczarny, I., De Villiers, M. and Van Biljon, J. (2012). How can usability contribute to user experience?: a study in the domain of e-commerce. In Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference, 216-225.
[S122] Read, J. (2012). Evaluating artefacts with children: age and technology effects in the reporting of expected and experienced fun. In Proceedings of the 14th ACM international conference on Multimodal interaction, 241-248.

[S123] Belk, M., Germanakos, P., Zaharias, P. and Samaras, G. (2012). Adaptivity Considerations for Enhancing User-Centric Web Experience. In ACHI 2012, The Fifth International Conference on Advances in Computer-Human Interactions, 348-353.

[S124] Robert, J. and Larouche, A. (2012). The dimensions of user experience with interactive systems. In Proceedings of IADIS International Conference-Interfaces and Human Computer Interaction, 89-96.

[S125] Walsh, T. and Nurkka, P. (2012). Approaches to cross-cultural design: two case studies with UX web-surveys. In Proceedings of the 24th Australian Computer-Human Interaction Conference, 633-642.

[S126] Lew, P., Abbasi, M., Rafique, I., Wang, X. and Olsina, L. (2012). Using Web Quality Models and Questionnaires for Web Applications Understanding and Evaluation. In Eighth International Conference on the Quality of Information and Communications Technology (QUATIC), 20-29.

[S127] Ciuciu, I. and Debruyne, C. (2012). Assessing the user satisfaction with an ontology engineering tool based on social processes. In On the Move to Meaningful Internet Systems: OTM 2012 Workshops, 242-251.

[S128] Witchel, H., Westling, C., Healy, A., Chockalingam, N. and Needham, R. (2012). Comparing four technologies for measuring postural micromovements during monitor engagement. In Proceedings of the 30th European Conference on Cognitive Ergonomics, 189-192.

[S129] Sikorski, M. (2012). A Cross-Disciplinary UX Evaluation of a CRM System. In I-UxSED, 31-36.

[S130] Weiss, B. (2012) Analyzing User Behavior for Quality Assessment. In International Conference on Interfaces and Human Computer Interaction, 223-226.

[S131] Matthias, H. and Beckhaus, S. (2012). Adaptive generation of emotional impact using enhanced virtual environments. Presence: Teleoperators and Virtual Environments, 21(1), 96-116.

[S132] Sunnari, M., Arhippainen, L., Pakanen, M. and Hickey, S. (2012). Studying user experiences of autostereoscopic 3D menu on touch screen mobile device. In Proceedings of the 24th Australian Computer-Human Interaction Conference, 558-561.

[S133] Song, B. and Yun, H. (2012). Flow and presence analysis of interactive digital contents. In 7th International Conference on Computing and Convergence Technology (ICCCT), 1272-1276.

[S134] Abbasi, M., Weng, J., Wang, Y., Rafique, I., Wang, X. and Lew, P. (2012). Modeling and Evaluating User Interface Aesthetics Employing ISO 25010 Quality Standard. In Eighth International Conference on the Quality of Information and Communications Technology (QUATIC), 303-306.

[S135] Johnson, F. (2012). Using Semantic Differentials for an Evaluative View of the Search Engine as an Interactive System. In EuroHCIR 7-10.

[S136] Islam, A. and Mäntymäki, M. (2012). Continuance of professional social networking sites: A decomposed expectation-confirmation approach. In Thirty Third International Conference on Information Systems, 1-18.

[S137] Djamasbi, S., Mehta, D. and Samani, A. (2012). Eye Movements, Perceptions, and Performance. In Proceedings of the eighteenth Americas Conference on Information Systems (AMCIS), 1.

[S138] Mulwa, C., Lawless, S., O'Keeffe, I., Sharp, M. and Wade, V. (2012). A recommender framework for the evaluation of end user experience in adaptive technology enhanced learning. International Journal of Technology Enhanced Learning, 4(1-2), 67-84.

[S139] Rauschenberger, M., Olschner, S., Cota, M., Schrepp, M. and Thomaschewski, J. (2012). Measurement of user experience: A Spanish language version of the user experience questionnaire (UEQ). In 7th Iberian Conference on Information Systems and Technologies (CISTI), 1-6.

[S140] Loscertales, A., Cerezo, E., Creatura, M., González, J., Minguez, J., Gil, R. and López, Y. (2012). NeuroGame project: multimodal videogame neuro-evaluation. In Proceedings of the 13th International Conference on Interacción Persona-Ordenador, 18.

[S141] González, J., García, R., Brunetti, J., Gil, R. and Gimeno, J. (2012). SWET-QUM: a quality in use extension model for semantic web exploration tools. In Proceedings of the 13th International Conference on Interacción Persona-Ordenador, 15.

[S142] Masip, L., Oliva, M. and Granollers, T. (2012). The open repository of heuristics. In Proceedings of the 13th International Conference on Interacción Persona-Ordenador, 4.

[S143] Elbedweihy, K., Wrigley, S., Ciravegna, F., Reinhard, D. and Bernstein, A. (2012). Evaluating semantic search systems to identify future directions of research. In The Semantic Web: ESWC 2012 Satellite Events, 148-162.

[S144] Van Beurden, M., Ijsselsteijn, W. and de Kort, Y. (2011). User experience of gesture based interfaces: a comparison with traditional interaction methods on pragmatic and hedonic qualities. In Gesture and Sign Language in Human-Computer Interaction and Embodied Communication, 36-47.

[S145] Karapanos, E., Martens, J. and Hassenzahl, M. (2012). Reconstructing experiences with iScale. International Journal of Human-Computer Studies, 70(11), 849-865.

[S146] Tang, M., Ting, P. and Sie, Y. (2012). Exploring evaluation criteria of social navigational tools on social media: a case study of aNobii. In Proceedings of the 4th Information Interaction in Context Symposium, 21-26.

[S147] Fleury, A. (2012). Drawing and acting as user experience research tools. In Proceedings of the 10th asia pacific conference on Computer human interaction, 269-278.

[S148] Kunze, K. and Strohmeier, D. (2012). Examining subjective evaluation methods used in multimedia Quality of Experience research. In Fourth International Workshop on Quality of Multimedia Experience (QoMEX), 51-56.

[S149] Knijnenburg, B., Willemsen, M., Gantner, Z., Soncu, H. and Newell, C. (2012). Explaining the user experience of recommender systems. User Modeling and User-Adapted Interaction, 22(4-5), 441-504.

[S150] Buskermolen, D. and Terken, J. (2012). Co-constructing stories: a participatory design technique to elicit in-depth user feedback and suggestions about design concepts. In Proceedings of the 12th Participatory Design Conference: Exploratory Papers, Workshop Descriptions, Industry Cases-Volume 2, 33-36.

[S151] Tan, C., Rosser, D., Bakkes, S. and Pisan, Y. (2012). A feasibility study in using facial expressions analysis to evaluate player experiences. In Proceedings of The 8th Australasian Conference on Interactive Entertainment: Playing the System, 5.

[S152] Staiano, J., Menéndez, M., Battocchi, A., De Angeli, A. and Sebe, N. (2012). UX_Mate: from facial expressions to UX evaluation. In Proceedings of the Designing Interactive Systems Conference, 741-750.

[S153] Wiethoff, A. and Gehring, S. (2012). Designing interaction with media façades: a case study. In Proceedings of the Designing Interactive Systems Conference, 308-317.

[S154] Hellmers, J., Thomaschewski, J., Holt, E. and Wriedt, T. (2012). Usability Evaluation Methods for a Scientific Internet Information Portal. J. UCS, 18(10), 1308-1322.

[S155] Sim, G. and Horton, M. (2012). Investigating children's opinions of games: Fun Toolkit vs. This or That. In Proceedings of the 11th International Conference on Interaction Design and Children, 70-77.

[S156] Efstratiou, C., Leontiadis, I., Picone, M., Rachuri, K., Mascolo, C. and Crowcroft, J. (2012). Sense and sensibility in a pervasive world. In Pervasive Computing, 406-424.

[S157] Gualeni, S., Janssen, D. and Calvi, L. (2012). How psychophysiology can aid the design process of casual games: A tale of stress, facial muscles, and paper beasts. In Proceedings of the international conference on the foundations of digital games, 149-155.

[S158] Partala, T. and Salminen, M. (2012). User experience of photorealistic urban pedestrian navigation. In Proceedings of the International Working Conference on Advanced Visual Interfaces, 204-207.

[S159] Law, E. and Sun, X. (2012). Evaluating user experience of adaptive digital educational games with Activity Theory. International Journal of Human-Computer Studies, 70(7), 478-497.

[S160] Navalpakkam, V. and Churchill, E. (2012). Mouse tracking: measuring and predicting users' experience of web-based content. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2963-2972.

[S161] Abeele, V., Zaman, B. and De Grooff, D. (2012). User eXperience Laddering with preschoolers: unveiling attributes and benefits of cuddly toy interfaces. Personal and Ubiquitous Computing, 16(4), 451-465.

[S162] Moriarty, B., Lennon, E., DiCola, F., Buzby, K., Manzella, M. and Hromada, E. (2012). Utilizing depth based sensors and customizable software frameworks for experiential application. Procedia Computer Science, 12, 200-205.

[S163] Miles, R., Greensmith, J., Schnädelbach, H., Garibaldi, J. (2013). Towards a method of identifying the causes of poor user experience on websites. In 2013 13th UK Workshop on Computational Intelligence (UKCI), 258-265.

[S164] Zacharatos, H., Gatzoulis, C., Chrysanthou, Y., Aristidou, A. (2013). Emotion recognition for exergames using laban movement analysis. In Proceedings of Motion on Games, 61-66.

[S165] Keskinen, T., Hakulinen, J., Heimonen, T., Turunen, M., Sharma, S., Miettinen, T. and Luhtala, M. (2013). Evaluating the experiential user experience of public display

applications in the wild. In Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia, 7.

[S166] Marco, J., Cerezo, E. and Baldassarri, S. (2013). Bringing tabletop technology to all: evaluating a tangible farm game with kindergarten and special needs children. Personal and ubiquitous computing, 17(8), 1577-1591.

[S167] Pilloni, P., Mulas, F., Piredda, L. and Carta, S. (2013). Assessing the user experience design as a persuasive methodology in a real world sport application. In Proceedings of International Conference on Advances in Mobile Computing and Multimedia, 387.

[S168] Jraidi, I., Chaouachi, M. and Frasson, C. (2013). A dynamic multimodal approach for assessing learners' interaction experience. In Proceedings of the 15th ACM on International conference on multimodal interaction, 271-278.

[S169] Rantala, J., Müller, S., Raisamo, R., Suhonen, K., Väänänen-Vainio-Mattila, K. and Lantz, V. (2013). User experiences of mobile audio conferencing with spatial audio, haptics and gestures. In Proceedings of the 15th ACM on International conference on multimodal interaction, 59-66.

[S170] Arhippainen, L. (2013). A Tutorial of Ten User Experience Heuristics. In Proceedings of International Conference on Making Sense of Converging Media, 336.

[S171] van de Laar, B., Bos, D. P. O., Reuderink, B., Poel, M. and Nijholt, A. (2013). How much control is enough? Influence of unreliable input on user experience. IEEE transactions on cybernetics, 43(6), 1584-1592.

[S172] Keskinen, T., Melto, A., Hakulinen, J., Turunen, M., Saarinen, S., Pallos, T. and Salanterä, S. (2013, December). Mobile dictation for healthcare professionals. In Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia, 41.

[S173] Marco, J., Baldassarri, S. and Cerezo, E. (2013). NIKVision: Developing a Tangible Application for and with Children. J. UCS, 19(15), 2266-2291.

[S174] Plechawska-Wójcik, M., Luján-Mora, S. and Wójcik, L. (2013). Assessment of User Experience with Responsive Web Applications using Expert Method and Cognitive Walkthrough-A Case Study. In ICEIS (3), 111-118.

[S175] Lugrin, J., Cavazza, M., Charles, F., Le Renard, M., Freeman, J. and Lessiter, J. (2013). Immersive fps games: User experience and performance. In Proceedings of the 2013 ACM international workshop on Immersive media experiences, 7-12.

[S176] Janneck, M. and Dogan, M. (2013). The Influence of Stressors on Usability Tests-An Experimental Study. In WEBIST, 581-590.
[S177] Alexandros, L. and Michalis, X. (2013). The physiological measurements as a critical indicator in users' experience evaluation. In Proceedings of the 17th Panhellenic Conference on Informatics, 258-263.
[S178] Möller, S., Schmidt, S. and Beyer, J. (2013). Gaming taxonomy: An overview of concepts and evaluation methods for computer gaming qoe. In Fifth International Workshop on Quality of Multimedia Experience (QoMEX), 236-241.
[S179] Thiele, S., Meyer, L., Geiger, C., Drochtert, D. and Wöldecke, B. (2013). Virtual archery with tangible interaction. In IEEE Symposium on 3D User Interfaces (3DUI), 67-70.
[S180] Mõttus, M., Lamas, D., Pajusalu, M. and Torres, R. (2013). The evaluation of interface aesthetics. In Proceedings of the International Conference on Multimedia, Interaction, Design and Innovation, 3.
[S181] Lindley, S. and Monk, A. (2013). Measuring social behaviour as an indicator of experience. Behaviour and Information Technology, 32(10), 968-985.
[S182] Thebault, P., Decotter, D., Boussard, M. and Lu, M. (2013). Embodying services into physical places: toward the design of a mobile environment browser. ACM Transactions on Interactive Intelligent Systems (TiiS), 3(2), 8.
[S183] Bernhaupt, R. and Pirker, M. (2013). Evaluating user experience for interactive television: Towards the development of a domain-specific user experience questionnaire. In IFIP Conference on Human-Computer Interaction, 642-659.
[S184] Sonnleitner, A., Pawlowski, M., Kässer, T. and Peissner, M. (2013). Experimentally Manipulating Positive User Experience Based on the Fulfilment of User Needs. In IFIP Conference on Human-Computer Interaction, 555-562.
[S185] Zwinderman, M., Leenheer, R., Shirzad, A., Chupriyanov, N., Veugen, G., Zhang, B. and Markopoulos, P. (2013). Using Video Prototypes for Evaluating Design Concepts with Users: a comparison to usability testing. In IFIP Conference on Human-Computer Interaction, 774-781.
[S186] O'Brien, H. and Toms, E. (2013). Examining the generalizability of the User Engagement Scale (UES) in exploratory search. Information Processing and Management, 49(5), 1092-1107.
[S187] Dhir, A. and Al-kahtani, M. (2013). A Case Study on User Experience (UX) Evaluation of Mobile Augmented Reality Prototypes. J. UCS, 19(8), 1175-1196.

[S188] Sánchez, J., García, R., Brunetti, J., Gil, R. and Gimeno, J. (2013). Using SWET-QUM to Compare the Quality in Use of Semantic Web Exploration Tools. J. UCS, 19(8), 1025-1045.

[S189] O'Brien, H. and Lebow, M. (2013). Mixed-methods approach to measuring user experience in online news interactions. Journal of the American Society for Information Science and Technology, 64(8), 1543-1556.

[S190] Provost, G., Robert, J. (2013). The dimensions of positive and negative user experiences with interactive products. In International Conference of Design, User Experience, and Usability, 399-408.

[S191] Maguire, M. (2013). Using human factors standards to support user experience and agile design. In International Conference on Universal Access in Human-Computer Interaction, 185-194.

[S192] Gockel, B., Graf, H., Pagano, A., Pescarin, S. and Eriksson, J. (2013). An Approach to User Experience Evaluation for Virtual Museums.

[S193] Zen, M. (2013). Metric-based evaluation of graphical user interfaces: model, method, and software support. In Proceedings of the 5th ACM SIGCHI symposium on Engineering interactive computing systems, 183-186.

[S194] Vissers, J., De Bot, L. and Zaman, B. (2013). MemoLine: evaluating long-term UX with children. In Proceedings of the 12th International Conference on Interaction Design and Children, 285-288.

[S195] Sim, G., Cassidy, B. and Read, J. C. (2013). Understanding the fidelity effect when evaluating games with children. In Proceedings of the 12th International Conference on Interaction Design and Children, 193-200.

[S196] Norman, K. (2013). Geq (game engagement/experience questionnaire): a review of two papers. Interacting with Computers, 25(4), 278-283.

[S197] Silva, J. and El Saddik, A. (2013). Exertion interfaces for computer videogames using smartphones as input controllers. Multimedia systems, 19(3), 289-302.

[S198] Zacharatos, H., Gatzoulis, C. and Chrysanthou, Y. (2013). Affect Recognition during Active Game Playing based on Posture Skeleton Data. In GRAPP/IVAPP, 419-422.

[S199] Hart, J., Sutcliffe, A. and De Angeli, A. (2013). Love it or hate it!: interactivity and user types. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2059-2068.

[S200] Mirza-Babaei, P., Nacke, L., Gregory, J., Collins, N. and Fitzpatrick, G. (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, 1499-1508.

[S201] Huisman, G., van Hout, M., van Dijk, E., van der Geest, T. and Heylen, D. (2013). LEMtool: measuring emotions in visual interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 351-360.

[S202] Filko, D. and Martinovic, G. (2013). Emotion recognition system by a neural network based facial expression analysis. Automatika–Journal for Control, Measurement, Electronics, Computing and Communications, 54(2).

[S203] Tan, J., Rönkkö, K. and Gencel, C. (2013). A Framework for Software Usability and User Experience Measurement in Mobile Industry. In Joint Conference of the 23rd International Workshop on Software Measurement and the 2013 Eighth International Conference on Software Process and Product Measurement (IWSM-MENSURA), 156-164.

[S204] Liikkanen, L. and Gómez, P. (2013). Designing interactive systems for the experience of time. In Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces, 146-155.

[S205] Alotaibi, S. and Wald, M. (2013). Evaluation of the UTAUT model for acceptable user experiences in Identity Access Management Systems. In 8th International Conference for Internet Technology and Secured Transactions (ICITST), 232-237.

[S206] Miniukovich, A. (2013). Development of Attractiveness Walkthrough. In CHItaly (Doctoral Consortium), 84-91.

[S207] Cota, M., Thomaschewski, J., Schrepp, M. and Gonçalves, R. (2014). Efficient measurement of the user experience. A Portuguese version. Procedia Computer Science, 27, 491-498.

[S208] Weiss, B. and Willkomm, S. (2013). User Experience and Social Attribution for an Embodied Spoken Dialog System. In Computers as Sociual Actors Workshop Edinburgh Scotland UK.

[S209] Janneck, M. and Balin, M. (2013). Using Stressors in Usability Tests: Empirical Results and Practical Recommendations. International Journal of Sociotechnology and Knowledge Development (IJSKD), 5(3), 17-31.

[S210] Sánchez, J. and Vela, F. (2014). Assessing the player interaction experiences based on playability. Entertainment Computing, 5(4), 259-267.
[S211] Saariluoma, P. and Jokinen, J. (2014). Emotional dimensions of user experience: a user psychological analysis. International Journal of Human-Computer Interaction, 30(4), 303-320.
[S212] Sundar, S., Bellur, S., Oh, J., Xu, Q. and Jia, H. (2014). User experience of on-screen interaction techniques: An experimental investigation of clicking, sliding, zooming, hovering, dragging, and flipping. Human–Computer Interaction, 29(2), 109-152.
[S213] Lallemand, C., Koenig, V. and Gronier, G. (2014). 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, 11-20.
[S214] Crowle, S., Boniface, M., Poussard, B. and Asteriadis, S. (2014). A design and evaluation framework for a tele-immersive mixed reality platform. In International Conference on Augmented and Virtual Reality, 151-158.
[S215] Móro, R., Daráz, J. and Bieliková, M. (2014). Visualization of Gaze Tracking Data for UX Testing on the Web. In HT (Doctoral Consortium/Late-breaking Results/Workshops).
[S216] Ermes, V., Janß, A., Radermacher, K. and Röcker, C. (2014). Analyzing the benefits of integrative multi-dimensional assessments of usability features in interaction-centered user studies. In Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare, 227-230.
[S217] Regal, G., Busch, M., Hochleitner, C., Wokerstorfer, P. and Tscheligi, M. (2014). TACTUX: a tactile user experience assessment board. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational , 983-986.
[S218] Meedin, G. and Perera, I. (2014). Crowdsourcing towards User Experience evaluation: An intelligent user experience questionnaire (IUEQ). In International Conference on Advances in ICT for Emerging Regions (ICTer), 67-73.
[S219] Moriguchi M., Noda H., Kinoshita T. and Ookawa H. (2014). UX Improvement Framework for Large-Scale System DevelopmentIn NEC Technical Journal, 8(3), 33-36.
[S220] Auricht, M., Stark, R. and Blume, C. (2014). Integrating User Experience Validation into a New Engineering Development Process for Advanced Driver

Assistance Systems. In Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 1-6.

[S221] Rubio, G., Navarro, E. and Montero, F. (2014). APADYT: a multimedia application for SEN learners. Multimedia tools and applications, 71(3), 1771-1802.

[S222] Savioja, P., Liinasuo, M. and Koskinen, H. (2014). User experience: does it matter in complex systems?. Cognition, Technology and Work, 16(4), 429-449.

[S223] Yao, L., Liu, Y., Li, W., Zhou, L., Ge, Y., Chai, J. and Sun, X. (2014). Using physiological measures to evaluate user experience of mobile applications. In International Conference on Engineering Psychology and Cognitive Ergonomics, 301-310.

[S224] Brajnik, G. and Giachin, C. (2014). Using sketches and storyboards to assess impact of age difference in user experience. International Journal of Human-Computer Studies, 72(6), 552-566.

[S225] Kujala, S., Walsh, T., Nurkka, P. and Crisan, M. (2014). Sentence completion for understanding users and evaluating user experience. Interacting with Computers, 26(3), 238-255.

[S226] Andersen, R., Nasrollahi, K., Moeslund, T. and Haque, M. (2014). Interfacing assessment using facial expression recognition. In International Conference on Computer Vision Theory and Applications (VISAPP), 186-193.

[S227] Morán, A., Meza, V., Ramírez-Fernández, C., Grimaldo, A., García-Canseco, E., Orihuela-Espina, F. and Sucar, L. (2014). Revisiting the user experience of a virtual rehabilitation tool for the physical activation and cognitive stimulation of elders. In International Workshop on Ambient Assisted Living, 203-210.

[S228] Schrepp, M., Hinderks, A. and Thomaschewski, J. (2014). Applying the User Experience Questionnaire (UEQ) in different evaluation scenarios. In International Conference of Design, User Experience, and Usability, 383-392.

[S229] Colombo, L. and Landoni, M. (2014). A diary study of children's user experience with EBooks using flow theory as framework. In Proceedings of the 2014 conference on Interaction design and children, 135-144.

[S230] Miki, H. (2014). User experience evaluation framework for human-centered design. In International Conference on Human Interface and the Management of Information, 602-612.

[S231] Novick, D., Vicario, J., Santaella, B. and Gris, I. (2014). Empirical Analysis of Playability vs. Usability in a Computer Game. In International Conference of Design, User Experience, and Usability, 720-731.

[S232] Miniukovich, A. and De Angeli, A. (2014). Quantification of interface visual complexity. In Proceedings of the 2014 international working conference on advanced visual interfaces, 153-160.

[S233] Niforatos, E. and Karapanos, E. (2015). EmoSnaps: a mobile application for emotion recall from facial expressions. Personal and Ubiquitous Computing, 19(2), 425-444.

[S234] Novak, D., Nagle, A. and Riener, R. (2014). Linking recognition accuracy and user experience in an affective feedback loop. IEEE Transactions on Affective Computing, 5(2), 168-172.

[S235] Aggarwal, A., Niezen, G. and Thimbleby, H. (2014). User experience evaluation through the brain's electrical activity. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational, 491-500.

[S236] Pradana, D. and Ferdiana, R. (2014). Mobile applications rating assessments based on users experience perception. In Makassar International Conference on Electrical Engineering and Informatics (MICEEI), 175-179.

[S237] Liapis, A., Karousos, N., Katsanos, C. and Xenos, M. (2014). Evaluating user's emotional experience in HCI: the physiOBS approach. In International Conference on Human-Computer Interaction, 758-767.

[S238] Mendes, M., Furtado, E., Furtado, V. and de Castro, M. (2014). How do users express their emotions regarding the social system in use? A classification of their postings by using the emotional analysis of Norman. In International Conference on Social Computing and Social Media, 229-241.

[S239] Orero, J. and Rifqi, M. (2014). Design of a fuzzy affective agent based on typicality degrees of physiological signals. In International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, 304-313. Springer International Publishing.

[S240] Bollini, L. (2014). Orienteering and Orienteering Yourself. User Centered Design Methodologies Applied to Geo-referenced Interactive Ecosystems. In International Conference on Computational Science and Its Applications, 642-651.

[S241] Fernandes, F. and Paschoarelli, L. (2014). Online Shopping Websites: An Evaluation of User Experience and Interface Ergonomic Criteria from the Perspective of

Older Users. In International Conference of Design, User Experience, and Usability, 104-115.

[S242] Ihamäki, P. (2014). GameFlow experience model: understanding player enjoyment in pervasive adventure geocaching game. International Journal of Wireless and Mobile Computing, 7(6), 536-548.

[S243] Chai, J., Ge, Y., Liu, Y., Li, W., Zhou, L., Yao, L. and Sun, X. (2014). Application of frontal EEG asymmetry to user experience research. In International Conference on Engineering Psychology and Cognitive Ergonomics, 234-243.

[S244] Nagy, A. and Kovari, B. (2014). Incorporating sensory data collected on mobile devices into user experience analysis. In 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom), 487-491.

[S245] Ayzenberg, Y. and Picard, R. (2014). FEEL: A system for frequent event and electrodermal activity labeling. IEEE journal of biomedical and health informatics, 18(1), 266-277.

[S246] Guseva, A., Gusev, M. and Ristov, S. (2014). User satisfaction evaluation of the electronic student services system—iKnow. In 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 698-703.

[S247] Sosa-Tzec, O. and Siegel, M. (2014). Rhetorical evaluation of user interfaces. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational, 175-178.

[S248] Jalil, S., Myers, T. and Atkinson, I. (2014). Design Implications from the Preliminary Results of a Telemedicine Patient-Technology Interaction Study. In Proceedings of the 7th International Symposium on Visual Information Communication and Interaction, 192.

[S249] Schneegass, S. and Henze, N. (2014). Harvesting social media for assessing user experience. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational, 931-934.

[S250] Huang, C. and Stolterman, E. (2014). Temporal anchors in user experience research. In Proceedings of the 2014 conference on Designing interactive systems, 271-274.

[S251] Cowan, B. and Jack, M. (2013). Measuring Anxiety Towards Wiki Editing: Investigating the Dimensionality of the Wiki Anxiety Inventory-Editing. Interacting with Computers, 50(1).

[S252] Gatti, E., Bordegoni, M. and Camere, S. (2014). Experiences and Senses: An Experimental Based Methodology for the Design Optimization. Proceedings of the Colors of Care: The 9th International Conference on Design & Emotion, 340-348.
[S253] Jraidi, I., Chaouachi, M. and Frasson, C. (2014). A hierarchical probabilistic framework for recognizing learners' interaction experience trends and emotions. Advances in Human-Computer Interaction, 2014, 6.
[S254] Mõttus, M., Lamas, D. and Tokranova, D. (2014). Evaluating Aesthetics During Interaction Episodes. In Proceedings of the 2014 Mulitmedia, Interaction, Design and Innovation International Conference on Multimedia, Interaction, Design and Innovation, 1-14.
[S255] Mortazavi, B., Nyamathi, S., Lee, S., Wilkerson, T., Ghasemzadeh, H. and Sarrafzadeh, M. (2014). Near-realistic mobile exergames with wireless wearable sensors. IEEE journal of biomedical and health informatics, 18(2), 449-456.
[S256] Kamp, I. and Desmet, P. (2014). Measuring product happiness. In CHI'14 Extended Abstracts on Human Factors in Computing Systems, 2509-2514.
[S257] Centieiro, P., Romão, T. and Dias, A. (2014). From the lab to the world: studying real-time second screen interaction with live sports. In Proceedings of the 11th Conference on Advances in Computer Entertainment Technology, 14.
[S258] Zhu, M. and Fang, X. (2014). What nouns and adjectives in online game reviews can tell us about player experience?. In Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems, 1471-1476.
[S259] Tokuhisa, S., Kamiyama, Y. and Tokiwa, T. (2014). Personal, Physical, Social, and Creative Contextual Design for Art Education: How to Achieve Fun in Art Education for Children. Computers in Entertainment (CIE), 12(1), 2.
[S260] Lederman, R., Wadley, G., Gleeson, J., Bendall, S. and Álvarez-Jiménez, M. (2014). Moderated online social therapy: Designing and evaluating technology for mental health. ACM Transactions on Computer-Human Interaction (TOCHI), 21(1), 5.
[S261] Foglia, P., Zanda, M. and Trading, I. (2014). Towards relating physiological signals to usability metrics: a case study with a web avatar. WSEAS Transactions on Computers, 13, 624-634.
[S262] Endrass, B., Hall, L., Hume, C., Tazzyman, S. and André, E. (2014). A Pictorial Interaction Language for Children to Communicate with Cultural Virtual Characters. In International Conference on Human-Computer Interaction, 532-543.

[S263] Morán, A., Ramírez-Fernández, C., Meza-Kubo, V., Orihuela-Espina, F., García-Canseco, E., Grimaldo, A. and Sucar, E. (2015). On the Effect of Previous Technological Experience on the Usability of a Virtual Rehabilitation Tool for the Physical Activation and Cognitive Stimulation of Elders. Journal of medical systems, 39(9), 1-11.

[S264] Baraković, S. and Skorin-Kapov, L. (2015). Multidimensional modelling of quality of experience for mobile Web browsing. Computers in Human Behavior, 50, 314-332.

[S265] Lasa, G., Justel, D. and Retegi, A. (2015). Eyeface: a new multimethod tool to evaluate the perception of conceptual user experiences. Computers in Human Behavior, 52, 359-363.

[S266] Nogueira, P., Torres, V., Rodrigues, R. and Oliveira, E. (2015). An annotation tool for automatically triangulating individuals' psychophysiological emotional reactions to digital media stimuli. Entertainment Computing, 9, 19-27.

[S267] Aizpurua, A., Arrue, M. and Vigo, M. (2015). Prejudices, memories, expectations and confidence influence experienced accessibility on the Web. Computers in Human Behavior, 51, 152-160.

[S268] Yoon, S., Choi, Y. and Oh, H. (2015). User attributes in processing 3D VR-enabled showroom: Gender, visual cognitive styles, and the sense of presence. International Journal of Human-Computer Studies, 82, 1-10.

[S269] Rivero, L. and Conte, T. (2015). Using a Study to Assess User eXperience Evaluation Methods from the Point of View of Users. In Proc. of the 17th International Conference on Enterprise Information Systems (ICEIS), 88-95

[S270] Oh, J. and Lee, U. (2015). Exploring UX issues in Quantified Self technologies. In Eighth International Conference on Mobile Computing and Ubiquitous Networking (ICMU), 53-59.

[S271] Santoso, H., Isal, R., Basaruddin, T., Sadira, L. and Schrepp, M. (2014). Research-in-progress: User experience evaluation of Student Centered E-Learning Environment for computer science program. In 3rd International Conference on User Science and Engineering (i-USEr), 52-55.

[S272] Borsci, S., Federici, S., Bacci, S., Gnaldi, M. and Bartolucci, F. (2015). Assessing user satisfaction in the era of user experience: comparison of the SUS, UMUX, and UMUX-LITE as a function of product experience. International Journal of Human-Computer Interaction, 31(8), 484-495.

[S273] Mirza-Babaei, P. and Nacke, L. (2014). Introducing the biometric storyboards tool for games user research. In Games Media Entertainment (GEM), 1-7.

[S274] Hashemi, M. and Herbert, J. (2015). Baran: An Interaction-centred User Monitoring Framework.

[S275] Shaheen, S., El-Hajj, W., Hajj, H. and Elbassuoni, S. (2014). Emotion recognition from text based on automatically generated rules. In 2014 IEEE International Conference on Data Mining Workshop, 383-392.

[S276] Amanatiadis, A., Mitsinis, N. and Maditinos, D. (2015). A neural network-based approach for user experience assessment. Behaviour and Information Technology, 34(3), 304-315.

[S277] Ahmad, N., Zainal, A., Razak, F., Adnan, W. and Osman, S. (2015). User Experience Evaluation Of Mobile Spiritual Applications For Older People: An Interview And Observation Study. Journal Of Theoretical and Applied Information Technology, 72(1).

[S278] An, L., Yang, S. and Bhanu, B. (2015). Efficient smile detection by extreme learning machine. Neurocomputing, 149, 354-363.

[S279] Bampatzia, S., Antoniou, A., Lepouras, G., Roumelioti, E. and Bravou, V. (2015). Comparing game input modalities: A study for the evaluation of player experience by measuring emotional state and game usability. In 2015 IEEE 9th International Conference on Research Challenges in Information Science (RCIS), 530-531.

[S280] Dirin, A. and Nieminen, M. (2015). mLUX: usability and user experience development framework for m-learning. International Journal of Interactive Mobile Technologies (iJIM), 9(3), 37-51.

[S281] Nam, Y. (2015). Designing interactive narratives for mobile augmented reality. Cluster Computing, 18(1), 309-320.

[S282] Gupta, R. and Falk, T. (2015). Affective state characterization based on electroencephalography graph-theoretic features. In 2015 7th International IEEE/EMBS Conference on Neural Engineering (NER), 577-580.

[S283] Naya, V. and Ibáñez, L. (2015). Evaluating user experience in joint activities between schools and museums in virtual worlds. Universal Access in the Information Society, 14(3), 389-398.

Appendix B: Classification of the Identified UX Evaluation Technologies

		Q1		Q2			Q3			Q4			Q5						Q6			Q7				Q8		
Paper ID	**Technology Name**	**A**	**B**	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**	**CS**	**Sw**	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**A**	**B**	**C**	**D**	**A**	**B**	**C**
S01	2DES		X			X	X			X			N	N	X						X		X				X	
S02	3E (Expressing Experiences and Emotions)	X		X			X			X			Y	Y				X			X				X	X		
S03	Aesthetics scale	X		X			X			X	X	X	Y	Y		X					X				X		X	
S04	Affect Grid	X		X			X			X			N	N	X						X				X		X	
S05	Affective Diary		X	X			X					X	Y	N	X						X			X	X			X
S06	AttrakDiff	X		X			X			X	X	X	N	N	X						X				X		X	
S07	Attrak-Work questionnaire	X		X			X			X		X	Y	Y				X			X				X		X	
S08	AXE (Anticipated eXperience Evaluation)	X				X	X			X	X		Y	Y	X				X	X		X			X	X		
S09	Co-discovery	X		X			X			X			N	N	X				X	X	X		X			X		
S10	Context-aware ESM (MyExperience)		X	X			X					X	Y	Y			X				X			X				X
S11	Contextual Laddering	X		X			X			X			Y	N	X						X				X	X		
S09	Controlled observation	X		X			X			X			N	N	X				X	X	X		X					X
S12	Day Reconstruction Method	X		X			X				X		N	N	X						X				X			X
S13	Differential Emotions Scale (DES)	X				X	X				X		N	N	X						X				X		X	
S14	EMO2	X				X	X			X			N	N	X						X				X			X
S15	Emocards	X		X			X			X			N	N	X				X		X				X			X
S16	Emotion Sampling Device (ESD)	X				X	X					X	Y	N	X						X			X				X
S17	Experience clip	X		X			X					X	Y	Y			X				X		X			X		
S18	Experience Sampling Method (ESM)	X				X	X					X	N	N	X						X			X			X	
S19	Facereader	X				X	X			X			Y	Y	X						X		X				X	

		Q1		Q2			Q3			Q4			Q5						Q6			Q7				Q8		
Paper ID	**Technology Name**	**A**	**B**	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**	**CS**	**Sw**	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**A**	**B**	**C**	**D**	**A**	**B**	**C**
S20	Feeltrace		X	X			X			X			N	N	X						X		X				X	
S21	Fun Toolkit	X		X			X			X	X		Y	Y	X						X	X			X			X
S22	Game experience questionnaire (GEQ)	X				X	X				X		Y	Y				X			X		X		X		X	
S23	Geneva Appraisal Questionnaire	X		X			X				X		N	N	X						X	X			X			X
S24	Geneva Emotion Wheel	X		X			X			X			N	N	X						X				X		X	
S25	Group-based expert walkthrough	X		X			X	X	X	X			Y	Y				X	X		X		X		X			X
S26	Hedonic Utility scale (HED/UT)	X		X			X			X			Y	Y			X				X				X			X
S27	Human Computer trust	X		X			X			X			Y	Y				X			X				X		X	
S09	Immersion	X		X					X			X	N	N	X						X			X		X		
S28	Intrinsic motivation inventory (IMI)	X				X	X				X		N	N	X						X				X		X	
S29	iScale	X				X	X				X		Y	N	X						X				X			X
S30	Kansei Engineering Software		X			X	X				X		Y	Y		X					X				X		X	
S31	Living Lab Method	X				X	X				X		Y	Y				X			X		X			X		
S32	Long term diary study	X		X			X					X	N	N	X						X			X			X	
S33	Mental effort	X		X			X				X		N	N	X						X				X		X	
S09	Mental mapping	X		X			X		X	X			N	N	X						X				X	X		
S34	OPOS – Outdoor Play Observation Scheme	X		X			X					X	Y	Y				X			X		X				X	
S35	PAD	X			X		X			X			N	N	X						X				X		X	
S36	Paired comparison	X		X			X					X	N	N	X						X				X		X	
S37	Physiological arousal via electrodermal activity	X			X		X			X			Y	Y		X					X		X				X	
S38	Playability heuristics	X		X					X	X			Y	Y				X	X	X			X			X		
S39	Positive and Negative Affect Scale (PANAS)	X		X			X				X		N	N	X						X				X		X	
S40	PrEmo	X			X		X			X			N	N	X						X				X		X	
S41	Presence questionnaire	X		X			X			X			Y	Y				X			X				X		X	
S09	Private camera conversation	X		X			X			X			N	N	X				X	X	X	X			X	X		

		Q1		Q2			Q3			Q4			Q5						Q6			Q7				Q8		
Paper ID	**Technology Name**	**A**	**B**	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**	**CS**	**Sw**	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**A**	**B**	**C**	**D**	**A**	**B**	**C**
S42	Product Attachment Scale	X		X			X					X	N	N	X						X			X				X
S09	Product Personality Assignment	X		X			X			X			N	N	X						X	X			X	X		
S09	Property checklists	X		X					X	X			N	N	X				X	X	X		X			X		
S43	Psychophysiological measurements	X			X		X			X			Y	Y				X			X		X				X	
S09	Reaction checklists	X		X			X			X			N	N	X				X	X	X		X	X			X	
S44	Repertory Grid Technique (RGT)	X		X			X			X			Y	N	X						X				X			X
S45	Self Assessment Scale (SAM)	X		X			X			X			N	N	X						X				X		X	
S46	Sensual Evaluation Instrument	X				X	X			X	X		Y	Y				X			X		X		X	X		
S47	Sentence Completion	X		X			X			X	X	X	Y	N	X				X		X	X			X	X		
S48	SUMI	X			X		X			X	X	X	Y	Y	X						X				X		X	
S49	TRUE Tracking Realtime User Experience	X				X	X	X			X		Y	Y				X			X		X		X		X	
S50	TUMCAT		X			X	X					X	Y	Y			X	X			X		X	X				X
S51	UTAUT	X		X			X				X		Y	Y	X						X				X		X	
S52	UX Curve	X		X			X				X		Y	Y	X						X				X			X
S53	UX Expert evaluation	X		X					X		X		Y	Y				X			X		X					X
S54	Valence method	X		X			X			X	X		Y	Y	X						X		X		X			X
S55	WAMMI (Website Analysis and Measurement Inventory)	X			X		X				X		Y	Y		X					X				X		X	
S56	Workshops + probe interviews	X		X			X				X		Y	Y	X						X	X			X	X		

Footnote:
Q1 – Type of Technology: (a) Method, (b)Tool
Q2 – Availability: (a) Available For Free, (b) Available Under a License, (c) Not Available
Q3 – Information Source: (a) Users, (b) The Development Team, (c) UX Experts
Q4 – Location: (a) Lab or Industry, (b) Field – Specific, (c) Field – Free, (d) Others
Q5 – Type of Assessed Product: (a) Computer Science, (b) Software Applications, (a) Generic, (b) Web Application, (c) Mobile Application, (d) Others
Q6 – Product Development Phase: (a) Conceptual Ideas, (b) Design Models, (c) Functional prototypes
Q7 – Assessed Period of Experience: (a) Before Usage, (b) During Usage – Single Episode, (c) During Usage – Long Term Usage, (d) After Usage
Q8 – Collected Data: (a) Qualitative, (b) Quantitative, (c) Both

		Type of Tech.			UX Data Source			Locat.		Assessed Application				Eval. Artifact			Assessed UX Period				Type of Data			Correct.		Availability		
Paper ID	**Technology Name**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**
S057	UX Concept Testing	X	X	X	X			X	X	X				X			X	X		X			X	X		X		
S058	Emocards in Short Interaction Sequences		X		X			X		X						X				X		X			X	X		X
S059	User Testing Within Co-Design	X		X	X			X		X						X		X		X			X	X		X	X	
S060	Standardized Assessment Tool for Interactive Stories	X			X			X					X			X				X		X			X	X		
S061	MobPro			X	X				X			X				X		X	X			X			X			X
S062	ResQue (Recommender systems' Quality of user experience)	X			X			X					X			X				X		X			X	X		
S063	Generic framework for evaluating the user experience			X	X			X					X			X				X		X			X			X
S064	Fuzzy Comprehensive Evaluation Of User Experience	X		X	X		X	X					X			X				X		X			X	X		
S065	User Experience Evaluation Metrics	X	X	X	X			X					X			X		X		X			X		X	X		
S066	Identify meaningful playfulness experiences	X			X				X	X						X			X		X				X	X		
S067	Experience Clip		X	X	X				X			X				X		X		X	X				X			X
S068	Quality, Quality in use, actual Usability and User experience (2Q2U)			X	X			X		X						X		X					X	X				X
S069	Measuring the User Experience of Digital Books	X	X	X	X			X					X			X		X		X			X		X	X	X	
S070	Valence Method		X	X	X			X		X						X		X		X			X		X	X		
S071	An approach for evaluating the affective aspects of exploratory search	X			X			X					X			X				X		X			X			X
S072	Quality, Quality in use, actual Usability and User experience (2Q2U)			X	X			X		X						X		X					X	X				X
S073	Physiological measures and Questionnaire	X		X	X			X					X			X		X		X		X			X	X	X	
S074	Usability Metric for User Experience (UMUX)	X			X			X	X	X						X				X		X			X	X		
S075	i-Scale	X			X			X		X						X			X	X			X		X			X

		Type of Tech.			UX Data Source			Locat.		Assessed Application				Eval. Artifact			Assessed UX Period				Type of Data			Correct.		Availability		
Paper ID	**Technology Name**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**
S076	Framework for in situ evaluation of recommender systems	X		X	X				X				X			X		X		X			X		X			X
S077	Laddering with young children in User eXperience evaluations		X		X			X					X			X				X	X				X	X		
S078	HEART (Happiness, Engagement, Adoption, Retention, and Task success) framework	X		X	X			X	X		X					X		X		X			X		X			X
S079	Mobile Questionnaires	X			X				X			X				X		X		X		X			X			X
S080	Identify meaningful playfulness experiences	X			X				X	X						X			X		X				X	X		
S081	Microsoft's Reaction Cards	X	X		X			X		X						X		X		X			X		X			X
S082	Quality, Quality in use, actual Usability and User experience (2Q2U)			X	X			X		X						X		X					X	X				X
S083	Think aloud protocols and an emotion word prompt list		X		X			X			X					X				X		X			X	X		
S084	Framework for Evaluating Quality of Experience	X		X	X				X			X				X			X				X		X			X
S085	User Measures of Quality of Experience	X		X	X				X				X			X			X	X		X			X	X		
S086	The adjustable distraction (AD) method			X	X			X		X						X		X			X				X	X		
S087	Engagement Survey	X			X			X			X					X				X		X			X	X		
S088	User-centric approach to recommender system evaluation	X			X			X					X			X				X		X			X	X		
S089	Combining interviews and scales in the multidimensional evaluation of User Experience	X	X		X			X					X			X				X			X		X	X		
S090	ResQue (Recommender systems' Quality of user experience)	X			X			X					X			X				X		X			X	X		
S091	Procedure to evaluate recommender systems for experience products	X		X	X			X					X			X		X		X		X			X	X		
S092	Four-Dimensional Assessment System Model		X		X		X	X				X				X				X		X			X	X		
S093	UX Curve	X			X			X		X						X			X	X			X		X	X		

		Type of Tech.			UX Data Source			Locat.		Assessed Application				Eval. Artifact			Assessed UX Period				Type of Data			Correct.		Availability		
Paper ID	**Technology Name**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**
S094	Mixed-methods approach of GSR and observation-based techniques		X	X	X			X					X			X		X		X			X		X		X	
S095	Survey for measuring engagement in eLearning system	X			X			X					X			X				X		X				X			X
S096	Complementary UX data collection methods	X	X	X	X			X	X			X				X		X		X			X	X				X
S097	UX Evaluation of Brain-computer interfaces	X			X			X					X			X	X			X		X			X	X		
S098	User experience questionnaire	X			X			X					X			X				X		X			X			X
S099	EFEx - Evaluation Framework For End-User Experience In Evaluating Adaptive Systems			X	X			X					X			X		X					X		X			X
S100	Nonintrusive recognition system for bodily expression of affect			X	X			X					X			X		X				X			X			X
S101	Naturalistic enactment		X	X	X				X			X				X				X	X				X	X		
S102	Methodology for studying affective ludology using EEG measurement			X	X			X					X			X		X				X			X		X	
S103	Museum Experience Scale (MES) and Multimedia Guide Scale (MGS)	X			X				X				X			X				X		X			X	X		
S104	UX Curve	X			X			X		X						X			X	X			X		X	X		
S105	Guidelines to measure and assess the persuasive dimensions of user experiences	X					X	X		X						X		X				X		X		X		
S106	A model-based design approach in transport			X	X							X				X		X			X				X			X
S107	Combination of UX methods	X		X	X			X					X			X		X		X		X			X	X		
S108	Triggered Think-Aloud Protocol		X	X	X			X			X					X		X		X			X		X	X		
S109	VERO: Visual Experiential Requirements Organizer	X			X				X	X				X			X						X		X			X
S110	Schwartz's value model for modelling the subjective value perceived by users	X	X	X	X				X	X						X		X		X			X		X	X		

		Type of Tech.			UX Data Source			Locat.		Assessed Application				Eval. Artifact			Assessed UX Period				Type of Data			Correct.		Availability		
Paper ID	**Technology Name**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**
S111	UX scalable combined evaluation method set	X	X	X	X			X			X			X	X	X		X		X			X	X				X
S112	Metrics used to assess the quality of the user experience	X			X			X					X			X				X		X			X			X
S113	Computational aesthetics approach			X	X				X		X					X				X		X			X	X		
S114	Combination of Game Experience Questionnaire and Physiological Measures	X		X	X			X					X			X		X		X		X			X	X		
S115	Integrated experience acceptance model	X			X				X		X					X				X		X			X	X		
S116	Extensible metadata format for Serious Games (MDF-SG)	X			X				X				X			X				X		X			X			X
S117	Biometric Storyboards	X		X	X			X					X			X		X		X			X		X	X		
S118	Using questionnaires to evaluate all components of UX	X			X			X				X				X				X		X			X	X		
S119	DrawUX	X			X				X	X						X				X			X		X			X
S120	UXblackbox	X		X	X			X		X						X		X					X		X			X
S121	UX questionnaire	X			X			X			X					X				X			X		X			X
S122	The Fun Toolkit	X			X			X		X						X				X		X			X	X		
S123	AttrakDiff2	X			X			X		X						X				X		X			X			X
S124	10 UX dimensions	X			X				X	X						X				X	X				X	X		
S125	UX Web Survey	X			X				X	X				X		X	X			X	X				X	X		
S126	Quality, Quality in use, actual Usability and User experience (2Q2U)			X	X			X		X						X		X					X	X				X
S127	Post-study System Usability Questionnaire	X			X			X		X						X				X			X		X	X		
S128	Technologies for measuring postural micromovements			X	X			X		X						X		X				X			X		X	
S129	UX Evaluation framework		X	X	X			X	X	X						X			X	X	X				X	X		
S130	Speech recognition for UX evaluation		X	X	X			X					X			X		X				X			X			X
S131	Evaluation and rating method for MoEs	X			X			X					X			X				X		X		X		X		
S132	Adjective Card Selection Method		X		X			X					X			X				X	X				X	X		

		Type of Tech.			UX Data Source			Locat.		Assessed Application				Eval. Artifact			Assessed UX Period				Type of Data			Correct.		Availability		
Paper ID	**Technology Name**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**
S133	Flow and Presence Measuring Tool	X			X			X					X			X				X		X			X	X		
S134	Interface aesthetics requirements evaluation	X			X		X	X	X		X					X				X		X			X	X		
S135	Semantic Differentials for UX evaluation	X			X			X					X			X				X		X			X	X		
S136	Decomposed Expectation-Confirmation Model and Questionnaire	X			X			X					X			X				X		X			X	X		
S137	Eye tracking			X	X			X					X			X		X				X			X			X
S138	Recommender framework for the evaluation of end user experience	X			X			X					X			X				X	X				X			X
S139	User Experience Questionnaire	X			X			X		X						X				X		X			X	X		
S140	Multimodal Videogame Neuro-Evaluation			X	X			X					X	X		X	X	X				X			X	X		
S141	Semantic Web Exploration Tools Quality in Use Model (SWET-QUM)	X		X	X			X			X					X		X		X		X			X	X		
S142	Open-HEREDEUX (Open Heuristic REsource for Designening and Evaluating User eXperience)	X				X	X	X		X						X				X		X			X			X
S143	Flexible and comprehensive methodology for evaluating different semantic search approaches	X	X	X	X			X					X			X		X		X			X		X			X
S144	Self-Reporting Questionnaire	X			X			X					X			X	X					X			X			X
S145	i-Scale	X			X			X		X						X			X	X			X		X			X
S146	Self-report user experience measures	X			X			X					X			X				X		X			X	X		
S147	Participant-generated drawings and drama workshops as user experience research methods	X	X		X			X		X				X			X				X			X		X		
S148	Holistic comparison tool for multimedia quality evaluation methods	X	X			X		X					X			X		X		X			X		X			X
S149	Framework for UX evaluation of recommender systems	X			X			X					X			X				X		X			X	X		
S150	Co-constructing stories	X	X		X			X		X				X			X				X				X	X		

		Type of Tech.			UX Data Source			Locat.		Assessed Application				Eval. Artifact			Assessed UX Period				Type of Data			Correct.		Availability		
Paper ID	**Technology Name**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**
S151	Automated facial expressions analysis to infer player experiences			X	X				X				X			X		X				X			X	X		
S152	UX_Mate			X	X			X		X						X		X				X			X	X		
S153	Interview process based on the laddering technique		X		X				X				X			X				X			X		X	X		
S154	User Experience Questionnaire (UEQ)	X			X				X		X					X				X		X			X	X		
S155	The Fun Toolkit	X			X			X		X						X				X		X			X	X		
S156	Interviews and experience sampling techniques		X	X	X				X		X					X			X			X			X			X
S157	Psychophysiological measurements			X	X			X					X			X		X				X			X		X	
S158	AttrakDiff and Task Load	X			X			X				X				X				X		X			X	X		
S159	Activity Theory with Observations and Interviews		X	X	X				X				X			X		X		X	X				X	X		
S160	Mouse Tracking			X	X			X			X					X		X				X			X	X		
S161	User eXperience Laddering with preschoolers		X		X			X					X		X					X	X				X	X		
S162	User Engagement Metrics			X	X			X					X			X		X				X			X			X
S163	Electromyography (EMG) and interaction data			X	X			X			X					X		X				X			X		X	
S164	Emotion Recognition for Exergames using Laban Movement Analysis			X	X			X					X			X		X				X			X		X	
S165	Experiential user experience evaluation method	X			X				X				X			X	X			X		X			X	X		
S166	Video analysis			X	X			X					X			X		X					X		X	X		
S167	Custom motivation survey	X			X				X			X				X			X			X			X			X
S168	Dynamic Multimodal Approach for Assessing Learners' Interaction Experience	X		X	X			X					X			X		X		X		X			X			X
S169	Combination of UX methods	X	X		X			X				X				X		X		X			X		X			X
S170	Ten User Experience Heuristics	X				X	X	X		X					X	X					X				X	X		
S171	Questionnaire for Evaluating UX	X			X			X					X			X				X		X			X	X		

		Type of Tech.			UX Data Source			Locat.		Assessed Application				Eval. Artifact			Assessed UX Period				Type of Data			Correct.		Availability		
Paper ID	**Technology Name**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**
S172	Experiential user experience evaluation method	X			X				X				X			X	X			X		X			X	X		
S173	Video analysis			X	X			X					X			X		X					X		X	X		
S174	Expert Method and Cognitive Walkthrough	X					X	X			X	X			X			X				X			X	X		
S175	ITC-SOPI Presence questionnaires	X			X			X					X			X				X		X			X	X		
S176	Combination of UX methods	X	X	X	X			X		X						X		X		X			X		X	X		
S177	PhysiOBS			X	X			X		X						X		X				X			X			X
S178	Gaming QoE	X			X			X					X			X				X		X			X			X
S179	IPQ questionnaire	X			X			X					X			X				X		X			X			X
S180	VisAWI questionnaire	X			X			X			X					X				X		X			X	X		
S181	Rating Scale for Empathised Enjoyment	X			X			X					X			X				X		X			X	X		
S182	Satisfaction Questionnaire	X			X			X					X			X				X		X			X	X		
S183	iTV-UX questionnaire	X			X			X					X			X				X		X			X	X		
S184	User Needs Questionnaire	X			X			X		X						X				X		X			X			X
S185	Combination of UX Questionnaires and Observation	X		X	X			X		X				X	X		X	X					X		X	X		
S186	User Engagement Scale (UES)	X			X			X		X						X				X		X			X	X		
S187	Combination of UX Methods	X	X		X			X		X						X		X		X			X		X	X		
S188	Semantic Web Exploration Tools Quality in Use Model - SWET-QUM	X		X	X			X					X			X		X		X		X			X	X		
S189	Mixed-Methods Approach	X		X	X			X					X			X		X		X		X			X	X		
S190	10 dimensions to characterize UX			X	X			X		X				X	X	X	X	X	X	X			X		X			X
S191	Framework for user experience design			X	X			X		X				X				X				X			X			X
S192	User Experience (UX) aspects applied to virtual museums (VM) - VMUXE		X		X			X					X			X				X			X		X	X		
S193	Metric-Based Evaluation of Graphical User Interfaces			X	X				X		X					X		X				X			X	X		
S194	MemoLine	X			X			X		X						X			X	X			X		X	X		
S195	Smileyometer and Again Again method	X			X			X		X					X	X				X		X			X	X		

		Type of Tech.			UX Data Source			Locat.		Assessed Application				Eval. Artifact			Assessed UX Period				Type of Data			Correct.		Availability		
Paper ID	**Technology Name**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**
S196	GEQ (Game Engagement/Experience Questionnaire)	X			X			X					X			X				X		X			X	X		
S197	UX Questionnaire	X			X			X					X			X				X			X		X	X		
S198	Affect Recognition			X	X			X					X			X		X					X		X			X
S199	Semi-structured UX interview and questionnaire	X	X		X			X			X					X				X			X		X	X		
S200	Biometric Storyboards (BioSt)			X	X			X					X			X		X					X		X		X	
S201	LEMtool (Layered Emotion Measurement tool)	X			X			X			X					X				X			X		X			X
S202	Emotion Recognition System			X	X			X		X						X		X					X		X			X
S203	UX Evaluation Framework	X	X	X	X			X		X						X	X	X		X			X		X	X		
S204	Six design guidelines for time-considerate design			X	X			X		X				X	X	X		X					X		X	X		
S205	UTAUT	X			X			X		X						X				X			X		X	X		
S206	Attractiveness Walkthrough	X				X	X	X		X					X	X				X	X				X	X		
S207	Portuguese version of the UEQ	X			X			X		X						X				X			X		X	X		
S208	UX Evaluation Questionnaire considering AttrakDiff	X			X			X					X			X				X			X		X			X
S209	Combination of UX Methods	X	X	X	X			X		X						X		X		X			X		X	X	X	
S210	Playability Model to evaluate the player experience	X		X	X			X					X			X		X		X	X				X	X	X	
S211	Experience questionnaire	X			X			X		X						X				X			X		X	X		
S212	User Engagement Questionnaire for Interaction Techniques	X			X			X			X					X				X			X		X			X
S213	Psychological Needs-Driven Ux Expert Evaluation	X					X	X		X						X				X			X		X			X
S214	Immersive User eXperience (IUX) model		X	X	X			X					X			X		X		X			X		X			X
S215	Gaze Tracking			X	X			X			X					X		X					X		X	X		
S216	Integrative Multi-Dimensional Assessments of Usability Features	X	X	X	X			X		X						X		X		X			X		X	X		

		Type of Tech.			UX Data Source			Locat.		Assessed Application				Eval. Artifact			Assessed UX Period				Type of Data			Correct.		Availability		
Paper ID	**Technology Name**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**
S217	Tactile User Experience Assessment Board		X		X			X		X						X				X			X		X			X
S218	Intelligent User Experience Questionnaire	X			X				X	X						X				X	X				X	X		
S219	Quantitative UX Evaluation Tools			X	X			X					X			X		X				X			X		X	
S220	MINARGUS			X	X			X					X			X		X				X			X			X
S221	Quick-UX	X			X			X					X			X				X		X			X	X		
S222	UX indicators for complex work systems	X			X			X					X			X				X		X			X	X		
S223	Physiological Measures and UEQ	X		X	X			X				X				X		X		X		X			X	X	X	
S224	UX Questionnaire	X			X			X		X					X		X			X			X		X	X		
S225	Sentence Completion	X			X			X		X				X			X				X				X	X		
S226	Facial Expression Recognition			X	X			X					X			X		X		X		X			X			X
S227	TAM-based UX questionnaire	X			X			X					X			X				X		X			X			X
S228	User Experience Questionnaire	X			X			X		X						X				X		X			X	X		
S229	UX Adapted Scale	X			X				X				X			X		X				X			X			X
S230	User Experience Evaluation Framework	X	X	X	X			X		X				X	X	X	X	X		X		X			X			X
S231	Playability vs. Usability in a Computer Game			X	X			X					X			X		X			X				X			X
S232	Quantification of Interface Visual Complexity	X		X	X		X	X			X					X				X		X			X			X
S233	EmoSnaps: A Mobile Application for Emotion Recall from Facial Expressions	X			X				X			X				X			X				X		X	X		
S234	User experience in an affective feedback loop	X		X	X			X					X			X		X				X		X		X		
S235	User Experience Evaluation Through The Brain's Electrical Activity		X	X	X			X		X						X		X		X			X		X		X	
S236	User Experience Questionnaire (UEQ)	X			X				X		X					X				X		X			X	X		
S237	Physiological Observation	X		X	X			X		X						X		X		X			X		X			X
S238	Classification of User postings by using the emotional analysis of Norman	X			X				X		X					X				X	X			X				X
S239	Fuzzy Affective Agent			X	X			X					X			X		X				X			X		X	

		Type of Tech.			UX Data Source			Locat.		Assessed Application				Eval. Artifact			Assessed UX Period				Type of Data			Correct.		Availability		
Paper ID	**Technology Name**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**
S240	Open-Air Mobile User Test			X	X				X			X				X		X			X				X	X		
S241	Evaluation of User Experience and Interface Ergonomic Criteria	X			X		X	X			X					X	X			X		X			X	X		
S254	GameFlow experience model			X	X			X					X			X		X		X		X		X				X
S243	Frontal alpha asymmetry and UX Questionnaires	X		X	X			X		X						X		X		X		X			X	X	X	
S244	Incorporating sensory data into user experience analysis			X	X				X			X				X		X				X			X	X		
S245	FEEL—Frequent EDA and Event Logging	X		X	X			X				X				X		X					X		X			X
S246	Framework of UX questions	X			X			X					X			X				X			X		X	X		
S247	Rhetorical Evaluation of User Interfaces	X				X	X	X				X				X				X	X				X	X		
S248	Clinical User-Experience Evaluation (CUE)	X	X	X	X			X					X			X		X		X	X				X			X
S249	UX Evaluation with Facebook Data			X	X				X	X						X				X			X		X			X
S250	Temporal Anchors in User Experience Research	X	X		X			X		X						X				X	X				X			X
S251	Wiki anxiety inventory-editing (WAI-E)	X			X			X					X			X				X		X			X	X		
S252	Experimental Based Methodology for the Design Optimization	X		X	X			X	X	X						X				X		X		X				X
S253	Hierarchical Probabilistic Framework to Model the User's Experience	X		X	X			X					X			X	X	X		X		X			X		X	
S254	User-Reported Aesthetic Value	X	X	X	X			X			X					X		X		X			X		X	X	X	
S255	ExerSurvey	X			X			X					X			X				X			X		X	X		
S256	Measuring Product Happiness	X			X			X		X						X				X		X			X			X
S257	UX Evaluation of Sports App	X	X		X			X					X			X				X			X		X	X		
S258	The Lexical Analysis	X		X	X				X				X			X				X		X			X			X
S259	Workshop for Children	X	X		X			X					X			X				X			X		X	X		
S260	Moderated Online Social Therapy		X		X			X					X			X				X	X				X			X
S261	Relating Physiological Signals to Usability Metrics			X	X			X			X					X		X				X			X		X	

		Type of Tech.			UX Data Source			Locat.		Assessed Application				Eval. Artifact			Assessed UX Period				Type of Data			Correct.		Availability		
Paper ID	**Technology Name**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(c)**	**(d)**	**(a)**	**(b)**	**(c)**	**(a)**	**(b)**	**(a)**	**(b)**	**(c)**
S262	Facial Scales for Children	X			X			X					X			X				X		X			X	X		
S263	Direct Observation			X	X			X					X			X		X						X		X		X
S264	UX evaluation questionnaire	X			X			X			X					X				X		X			X	X		
S265	Panas-X and Eye Tracking	X		X	X			X		X						X		X		X		X			X	X		X
S266	Tool for Automatically Triangulating Individuals' Psychophysiological Emotional Reactions			X	X			X		X						X		X				X			X			X
S267	UX Interview		X		X			X			X					X				X	X				X			X
S268	Custom UX Questionnaires	X			X			X			X					X				X		X			X	X		
S269	Combination of UX methods	X	X		X			X			X				X					X			X		X	X		
S270	Examination of UX reviews	X			X				X				X			X				X	X				X	X		
S271	UEQ using Indonesian Language	X			X			X					X			X				X		X			X	X		
S272	UMUX and UMUX-Lite	X			X			X		X						X				X		X			X	X		
S273	Biometric Storyboards Tool			X	X			X					X			X		X				X			X		X	X
S274	External Context Detector - BARAN framework			X	X			X		X						X		X				X			X			X
S275	Emotion Classification in English Sentences	X			X				X	X						X				X			X		X			X
S276	Neural network-based approach for user experience assessment			X	X			X		X						X		X				X			X			X
S277	UX Interview and observation		X	X	X			X				X				X		X		X			X		X	X		
S278	Smile Detection			X	X			X		X						X		X				X			X			X
S279	Self-Assessment Manikin (SAM)	X			X			X					X			X				X		X			X			X
S280	mLUX: Usability and User Experience Development Framework for M-Learning		X		X			X					X			X				X	X				X			X
S281	Post activity questionnaire	X			X			X					X			X				X		X			X	X		
S282	Affective State Characterization based on Electroencephalography			X	X			X		X						X		X				X			X		X	X
S283	UX Survey	X			X			X					X			X				X		X			X	X		

Footnote:
SQ1 – Type of Technology: (a) Written Reporting, (b) Oral Reporting, (c) Observation/Monitoring
SQ2 – Information Source: (a) Users, (b) The Development Team, (c) UX Experts
SQ3 – Location: (a) Controlled environment, (b) Field
SQ4 – Type of Assessed Application: (a) Generic, (b) Web Application. (c) Mobile Application, (d) Others
SQ5 – Type of Assessed Artifact: (a) Conceptual Ideas, (b) Design Models, (c) Func. Prot. or Finished App.
SQ6 – Assessed Period of Experience: (a) Before Usage, During Usage – Single Ep., (c) During Usage – Long Term, (d) After Usage
SQ7 – Collected Data: (a) Qualitative, (b) Quantitative, (c) Both
SQ8 – Supports Correction of Identified Problems: (a) Yes, (b) No
SQ9 – Availability: (a) Available for Free, (b) Available Under a License, (c) Not Available

Appendix C: REX Technique

ID	Aspect	Question/Answer	Emotion	Item	Suggestion
1	Tutorial	How was your first contact with the app?			
1_1	No Tutorial Available	I started Using How did you feel about the app?			
1_1_1			Confused		Check whether the following problems must be corrected: [1_1_1_1]
1_1_1_1				I did not understand the purpose of the app	Design Suggestions: Tutorials and Invitations (Use Less Text, No Frontloading, Make It Rewarding, Tutorials should be skippable, Anything conveyed in the tutorial should be accessible at all times)
1_1_2			Satisfied		
1_2	Available Tutorial	Before using, the app showed me an introduction of how to use it. How did you feel about the introduction of the app?			
1_2_1			Confused		Check whether the following problems must be corrected: [1_2_1_1] [1_2_1_2] [1_2_1_3]
1_2_1_1				I did not know that I was in the tutorial	Apply the following heuristic - Visibility of System Status
1_2_1_2				I did not understand the tutorial of the app	Design Suggestions: Tutorials and Invitations (No Frontloading)
1_2_1_3				I did not understand the words or images	Apply the following heuristic - Match between system and the real world
1_2_2			Overloaded		Check whether the following problems must be corrected: [1_2_2_1] [1_2_2_2]

ID	Aspect	Question/Answer	Emotion	Item	Suggestion
1_2_2_1				It explained too much and not what was important	Design Suggestions: Tutorials and Invitations (No Frontloading)
1_2_2_2				It had a lot of text, it did not summarize	Design Suggestions: Tutorials and Invitations (Use Less Text)
1_2_3			Stuck		Check whether the following problems must be corrected: [1_2_3_1] [1_2_3_2] [1_2_3_3]
1_2_3_1				I was not able to test what it was showing	Design Suggestions: Tutorials and Invitations (Make It Rewarding)
1_2_3_2				I could not skip the tutorial	Design Suggestions: Tutorials and Invitations (Tutorials should be skippable)
1_2_3_3				I wasn't able to access it if I closed it	Design Suggestions: Tutorials and Invitations (Anything conveyed in the tutorial should be accessible at all times)
1_2_4			Satisfied		
2	Functionalities	Which of the following tasks did you carry out in the application?			
2_1	Navigation	I navigated through the application, using the menus and options How was the navigation of the app?			
2_1_1	Persistent Navigation	The main menu was always visible in the screen How did you feel about the navigation being that way?			
2_1_1_1			Confused		Check whether the following problems must be corrected: [2_1_1_1_1] [2_1_1_1_2] [2_1_1_1_3] [2_1_1_1_4]
2_1_1_1_1				The items in the menu did not show what happened in the app	Design Suggestions: Navigation Patterns - persistent (menu categories have status indicators)

ID	Aspect	Question/Answer	Emotion	Item	Suggestion
2_1_1_1_2				I did not understand that the items in the menu where for navigating	Apply the following heuristic - Consistency and standards
2_1_1_1_3				I did not know where I was going after using a menu item	Apply the following heuristic - Match between system and the real world
2_1_1_1_4				I did not know where I was when navigating through the application	Apply the following heuristic - Visibility of System Status
2_1_1_2			Overloaded		Check whether the following problems must be corrected: [2_1_1_2_1] [2_1_1_2_2] [2_1_1_2_3]
2_1_1_2_1				It showed items that were not necessary most of the time	Design Suggestions: Navigation Patterns (Transient navigation)
2_1_1_2_2				I almost never used the menu and it occupied unnecessary space	Design Suggestions: Navigation Patterns (Transient navigation)
2_1_1_2_3				It had a lot of items for the available space	Design Suggestions: Navigation Patterns (Transient navigation)
2_1_1_3			Satisfied		
2_1_2	Transient Navigation	The main menu was hidden, unless I activated it How did you feel about the navigation being that way?			
2_1_2_1			Confused		Check whether the following problems must be corrected: [2_1_2_1_1] [2_1_2_1_2] [2_1_2_1_3] [2_1_2_1_4]
2_1_2_1_1				It had few options and was empty	Design Suggestions: Navigation Patterns (Persistent navigation)
2_1_2_1_2				I did not understand that the menu items allow navigating through the app	Apply the following heuristic - Consistency and standards

ID	Aspect	Question/Answer	Emotion	Item	Suggestion
2_1_2_1_3				I did not know where I was going after activating a navigation item	Apply the following heuristic - Match between system and the real world
2_1_2_1_4				I did not know where I was after using a navigation item	Apply the following heuristic - Visibility of System Status
2_1_2_2			Overloaded		Check whether the following problems must be corrected: [2_1_2_2_1]
2_1_2_2_1				I was forced to access the menu all the time to navigate through the app	Design Suggestions: Navigation Patterns (Persistent navigation)
2_1_2_3			Helpless		Check whether the following problems must be corrected: [2_1_2_3_1]
2_1_2_3_1				The menu items did not show what was happening in the app	Design Suggestions: Navigation Patterns - persistent (menu categories have status indicators)
2_1_2_4			Satisfied		
2_2	Specific Functionalities of the Application	I used a function of the application Which of the following functionalities did you use in the app?			
2_2_1	Registration	I registered And how did you feel about your registration?			
2_2_1_1			Confused		Check whether the following problems must be corrected: [2_2_1_1_1]
2_2_1_1_1				It took me to a web page outside the app	Design Suggestions: Input Forms (Registration)
2_2_1_2				Overloaded	Check whether the following problems must be corrected: [2_2_1_2_1] [2_2_1_2_2]
2_2_1_2_1				I had to fill in a lot of information	Design Suggestions: Input Forms (Registration)
2_2_1_2_2				It did not help me fill in the data	Design Suggestions: Input Forms (Registration)

ID	Aspect	Question/Answer	Emotion	Item	Suggestion
2_2_1_3			Stuck		Check whether the following problems must be corrected: [2_2_1_3_1]
2_2_1_3_1				It did not allow me to customize the app	Design Suggestions: Input Forms (Registration with Personalization)
2_2_1_4			Satisfied		
2_2_2	Login	I logged in And how did you feel about the login?			
2_2_2_1			Helpless		Check whether the following problems must be corrected: [2_2_2_1_1]
2_2_2_1_1				It didn't help me if I forgot information or how to fill data	Design Suggestions: Input Forms (Sign In)
2_2_2_2			Overloaded		Check whether the following problems must be corrected: [2_2_2_2_1]
2_2_2_2_1				I had to fill in a lot of information	Design Suggestions: Input Forms (Sign In)
2_2_2_3			Stuck		Check whether the following problems must be corrected: [2_2_2_3_1] [2_2_2_3_2]
2_2_2_3_1				It did not provide an alternative to sign in	Design Suggestions: Input Forms (Sign In)
2_2_2_3_2				It did not allow login in with an alternative registration	Design Suggestions: Input Forms (Sign In)
2_2_2_4			Satisfied		
2_2_3	Searching	I carried out a search And how did you feel about carrying out a search?			
2_2_3_1			Confused		Check whether the following problems must be corrected: [2_2_3_1_1] [2_2_3_1_2]
2_2_3_1_1				It did not provide the data in an organized way	Design Suggestions: Search, Sort, and Filter (Search Results/View Results)

ID	Aspect	Question/Answer	Emotion	Item	Suggestion
2_2_3_1_2				The results from the search did not match with what I was searching	Apply the following heuristic - Consistency and standards
2_2_3_2			Stuck		Check whether the following problems must be corrected: [2_2_3_2_1] [2_2_3_2_2] [2_2_3_2_3] [2_2_3_2_4]
2_2_3_2_1				It did not allow focusing the search in categories	Design Suggestions: Search, Sort, and Filter (Scoped Search)
2_2_3_2_2				It did not allow ordering the data	Design Suggestions: Search, Sort, and Filter (Onscreen Sort, Sort Overlay, Sort Form)
2_2_3_2_3				It id not allow refining the results or filtering	Design Suggestions: Search, Sort, and Filter (Onscreen Filter, Filter Overlay, Filter Form, Filter Drawer)
2_2_3_2_4				It did not allow me to carry out an advanced search	Design Suggestions: Search, Sort, and Filter (Search Form, Sort Form)
2_2_3_3			Helpless		Check whether the following problems must be corrected: [2_2_3_3_1] [2_2_3_3_2] [2_2_3_3_3]
2_2_3_3_1				It did not autocomplete what I was entering	Design Suggestions: Search, Sort, and Filter (Search with Auto-Complete)
2_2_3_3_2				It did not help me fill in the data	Design Suggestions: Search, Sort, and Filter (Explicit Search)
2_2_3_3_3				It did not provide another way to enter da data	Design Suggestions: Search, Sort, and Filter (Explicit Search)
2_2_3_4			Satisfied		
2_2_4	Specific Functionalities of the Application	I used one of the specific functions of the app And how did you feel about the specific functionalities of the app?			

ID	Aspect	Question/Answer	Emotion	Item	Suggestion
2_2_4_1			Confused		Check whether the following problems must be corrected: [2_2_4_1_1] [2_2_4_1_2] [2_2_4_1_3] [2_2_4_1_4] [2_2_4_1_5]
2_2_4_1_1				I did not see how to finish the task	Design Suggestions: Feedback and Affordance (Affordance)
2_2_4_1_2				I did not understand how the actions worked	Design Suggestions: Feedback and Affordance (System Status)
2_2_4_1_3				I did not understand what each action would accomplish	Design Suggestions: Tools (Toolbar, Toolbox, Inline Actions, Contextual Tools, Bulk Actions)
2_2_4_1_4				The app did not show me if I had finished the task	Design Suggestions: Feedback and Affordance (Confirmation) Tools (Multi-State Button)
2_2_4_1_5				The app did not show the progress of the task	Design Suggestions: Feedback and Affordance (System Status)
2_2_4_2			Overloaded		Check whether the following problems must be corrected: [2_2_4_2_1]
2_2_4_2_1				It had a lot of options and it was disorganized	Design Suggestions: Tools (Toolbar, Toolbox)
2_2_4_3			Stuck		Check whether the following problems must be corrected: [2_2_4_3_1] [2_2_4_3_2]
2_2_4_3_1				I wasn't able to cancel or drop a task	Apply the following heuristic - User control and freedom
2_2_4_3_2				I wasn't able to use the app with the locked screen	Design Suggestions: Tools (Lock Screen Controls)
2_2_4_4			Satisfied		
2_3	Information	I viewed information in the application. What type of information did you view in the application?			
2_3_1	Graphs	I viewed graphs How did you feel about the graphs?			
2_3_1_1			Confused		Check whether the following problems must be corrected: [2_3_1_1_1] [2_3_1_1_2]

ID	Aspect	Question/Answer	Emotion	Item	Suggestion
2_3_1_1_1				I did not manage to understand the graphs	Design Suggestions: Charts (Keep it simple, Sparkline)
2_3_1_1_2				It had no labels or footnotes to understand it better	Design Suggestions: Charts (Integrated Legend)
2_3_1_2			Overloaded		Check whether the following problems must be corrected: [2_3_1_2_1]
2_3_1_2_1				They were very complex and had a lot of information	Design Suggestions: Charts (Keep it simple, Sparkline, Surprise—math quiz!, Integrated Legend, Thresholds)
2_3_1_3			Stuck		Check whether the following problems must be corrected: [2_3_1_3_1] [2_3_1_3_2] [2_3_1_3_3] [2_3_1_3_4]
2_3_1_3_1				I wasn't able to filter the information	Design Suggestions: Charts (Chart with Filters, Pivot Table)
2_3_1_3_2				I couldn't see additional data	Design Suggestions: Charts (Overview plus Data)
2_3_1_3_3				I wasn't able to view them well or increase their size	Design Suggestions: Charts (Zoom)
2_3_1_3_4				I wasn't able to Interact with it, it was static	Design Suggestions: Charts (Interactive Timeline, Data Point Details, Drill Down, Interactive Preview)
2_3_1_4			Satisfied		
2_3_2	Tables	I viewed tables How did you feel about the tables?			
2_3_2_1			Confused		Check whether the following problems must be corrected: [2_3_2_1_1] [2_3_2_1_2] [2_3_2_1_3] [2_3_2_1_4]
2_3_2_1_1				I wasn't able to see the separation of the date	Design Suggestions: Tables (Basic Table)
2_3_2_1_2				The headings didn't help me understand the data	Design Suggestions: Tables (Headerless Table)
2_3_2_1_3				It did not present visual indicators	Design Suggestions: Tables (Table with Visual Indicators)

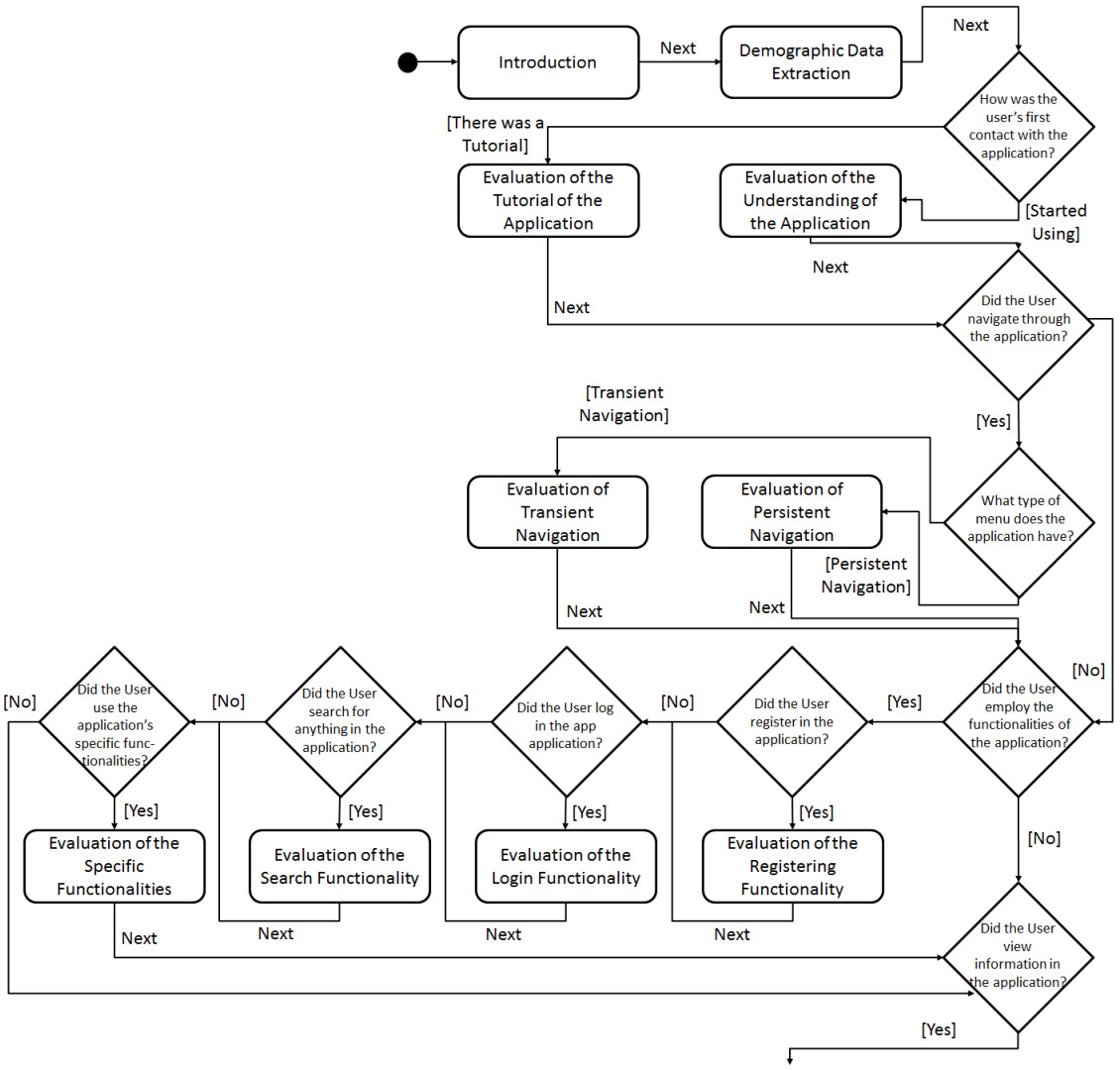
ID	Aspect	Question/Answer	Emotion	Item	Suggestion
2_3_2_1_4				It was difficult to understand	Design Suggestions: Tables (Grouped Rows)
2_3_2_2			Overloaded		Check whether the following problems must be corrected: [2_3_2_2_1]
2_3_2_2_1				There were lots of columns	Design Suggestions: Tables (Fixed Column)
2_3_2_3			Satisfied		
2_3_3	Figures and Text	I viewed figures and text How did you feel about the Figures and Text?			
2_3_3_1			Confused		Check whether the following problems must be corrected: [2_3_3_1_1] [2_3_3_1_2]
2_3_3_1_1				I wasn't able to understand it	Apply the following heuristic - Match between system and the real world
2_3_3_1_2				I wasn't able to see it easily	Apply the following heuristic - Aesthetic and minimalist design Apply the following heuristic - User control and freedom
2_3_3_2			Overloaded		Check whether the following problems must be corrected: [2_3_3_2_1]
2_3_3_2_1				There was a lot of information in the same screen	Apply the following heuristic - Aesthetic and minimalist design
2_3_3_3			Satisfied		
3	Feedback	How was the answer of the app when you made a mistake when carrying out a task?			
3_1		I did not make a mistake			
3_2		I made a mistake, but the system did not show me any error messages			Check whether the following problems must be corrected: [3_3_1_1] [3_3_1_2] [3_3_1_3] [3_3_2_1]

ID	Aspect	Question/Answer	Emotion	Item	Suggestion
3_3		I made a mistake and the system showed me a message about the error And how did you feel about the provided feedback?			
3_3_1			Confused		Check whether the following problems must be corrected: [3_3_1_1] [3_3_1_2] [3_3_1_3]
3_3_1_1				I did not understand what was the error	Design Suggestions: Feedback and Affordance (Error Messages)
3_3_1_2				I wasn't able to see the message easily	Design Suggestions: Feedback and Affordance (Error Messages)
3_3_1_3				It didn't help me to correct the problem	Design Suggestions: Feedback and Affordance (Error Messages)
3_3_2			Overloaded		Check whether the following problems must be corrected: [3_3_2_1]
3_3_2_1				It had a lot of information instead of a summary	Apply the following heuristic - Aesthetic and minimalist design
3_3_3			Satisfied		
4	Help	How was the app in terms of help?			
4_1		I did not need help			
4_2		I needed help, but the application did not have a help option			Check whether the following problems must be corrected: [4_3_1_1] [4_3_1_2] [4_3_1_3] [4_3_2_1] [4_3_3_1]
4_3		I needed help and accessed the available option in the application And how did you feel about the provided help?			
4_3_1			Confused		Check whether the following problems must be corrected: [4_3_1_1] [4_3_1_2] [4_3_1_3]
4_3_1_1				It did not describe the functionalities	Design Suggestions: Help (How-Tos, Tutorials)

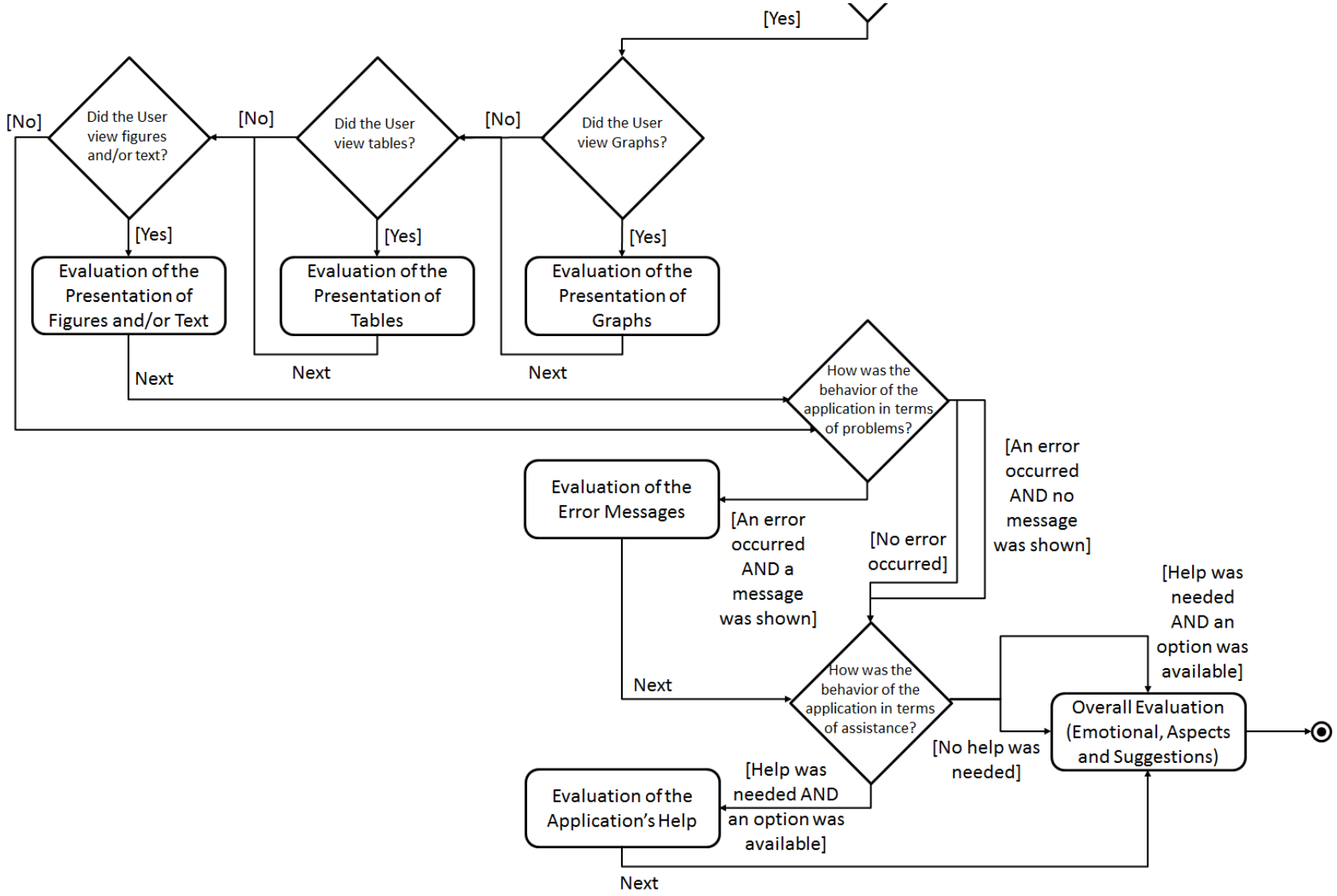
ID	Aspect	Question/Answer	Emotion	Item	Suggestion
4_3_1_2				It did not describe how the app worked	Design Suggestions: Help (User Guide/Help System, Feature Tours)
4_3_1_3				It did not have answers to frequently asked questions	Design Suggestions: Help (FAQs)
4_3_2			Overloaded		Check whether the following problems must be corrected: [4_3_2_1]
4_3_2_1				It had a lot of information instead of summarizing	Design Suggestions: Tutorials and Invitations (Use Less Text) Apply the following heuristic - Aesthetic and minimalist design
4_3_3			Helpless		Check whether the following problems must be corrected: [4_3_3_1]
4_3_3_1				It did not have a way to report problems in the app	Design Suggestions: Help (Capture Feedback)
4_3_4			Satisfied		
5	Overall	Overall, how did you feel about the application?			
			Emotions Proposed by Scherer (2005)		

Evaluated Aspect
- **Tutorial**
- **Functionalities**
- **Feedback**
- **Help**
- **Overall**

Appendix D: State Diagram for Applying REX



It continues in the next page


[Yes]
Did the User view Graphs?
[No]
Did the User view tables?
[No]
Did the User view figures and/or text?
[No]
[Yes]
Evaluation of the Presentation of Graphs
[Yes]
Evaluation of the Presentation of Tables
[Yes]
Evaluation of the Presentation of Figures and/or Text
Next
Next
Next
How was the behavior of the application in terms of problems?
[An error occurred AND no message was shown]
[No error occurred]
[An error occurred AND a message was shown]
Evaluation of the Error Messages
[Help was needed AND an option was available]
Next
How was the behavior of the application in terms of assistance?
Overall Evaluation (Emotional, Aspects and Suggestions)
[No help was needed]
[Help was needed AND an option was available]
Evaluation of the Application's Help
Next