
Let Users Tell the Story: Evaluating User Experience with Experience Reports

Hannu Korhonen

Nokia Research
P.O. Box 1000.
00045 Nokia Group, Finland
Hannu.j.korhonen@nokia.com

Juha Arrasvuori

Nokia Research
P.O. Box 1000.
00045 Nokia Group, Finland
Juha.arrasvuori@nokia.com

Kaisa Väänänen-Vainio-Mattila

Tampere University of Technology
P.O. Box 589, 33101 Tampere,
Finland
Kaisa.vaananen-vainio-
mattila@tut.fi
and
Nokia Research
P.O. Box 1000.
00045 Nokia Group, Finland
kaisa.vaananen-vainio-
mattila@nokia.com

Abstract

User experience (UX) has been under extensive research in recent years. One of the core questions has been how to evaluate user experience. Several methods such as diaries, experience sampling and questionnaires have been used for collecting data on user experience with a product. Although these methods provide valuable data, they may lack obtaining rich descriptions of UX in users' everyday lives. We have approached the question of UX evaluation by *experience reports* which are open-ended experience stories written by the users after using their products in real contexts of use. In this paper, we describe a field study in which 21 participants wrote 116 experience reports about UX with their personal products such as smart phones and MP3 players. The reports were analyzed with predefined context and experience categorizations to identify core experiences. We discuss our initial findings on the applicability of the method to evaluate UX.

Keywords

User Experience, Experience Report, Evaluation Method

ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI):
User Interfaces, *User-centered Design*.

Introduction

User experience (UX) is a concept formed during the past decade to describe the holistic nature of users' interaction with products. A user's interaction with a product in certain contexts defines the user experience [3], [9], [11]. Even though a multitude of UX models and frameworks have been defined, very little researched knowledge exists of the details of the actual experiences products elicit on users. There is a lack of understanding of user experience as described by the users themselves.

We conducted a ten-day field study to explore the nature of UX and to test a method for evaluating UX. 21 participants wrote experience reports of their meaningful experiences with their personal products – such as smart phones, portable music players, and heart rate monitors – that are integrated in the users' daily lives. Each report was analyzed by first determining the “triggering context” i.e. the most significant aspect of the context of use affecting the user's experience described in the report. This was done with eight context categories [6]. Then, a predefined categorization of playful experiences [10] was used to identify the “core experience” in the report. The core experience can be defined as a central, meaningful experience that results from the interaction with a product with relation to the triggering context element.

In this paper, we present a UX evaluation method based on experience reports and their analysis with the predefined context and experience categories. We discuss our initial findings on the applicability of the method to gain understanding of meaningful aspects of UX. This approach provides rich descriptions of UX and enables a deeper understanding of UX than provided by previous UX evaluation methods such as experience sampling or diaries.

Related Work

There are two main approaches capturing user experience of a product. UX can be evaluated either during the product usage or after the usage of a product has finished. In situ evaluation has often been a preferred method because it enables collection of detailed contextual data. For example, Experience Sampling Method [1] can be used to systematically study subjective experiences of people in naturalistic settings. The method is well suited for capturing momentary experiences from a continuous activity.

Retrospective evaluation of UX is another way of collecting data. Different kinds of questionnaires have been developed for this purpose. Desmet et al. have introduced Emocards which allow the users to express their emotional responses to a product [2]. AttrakDiff is a questionnaire consisting of 28 term pairs and it evaluates four aspects of user experience: pragmatic quality, attractiveness, identity, and stimulation [7]. Questionnaires provide primarily quantitative data, but do not go into depths of the qualitative nature of UX.

Diary studies are a common data collection method and they can incorporate both in situ evaluation and retrospective evaluation of user experience. The user answers to predefined questions, typically at the end of the day and describes how they experience interaction with a product.

Researchers have noted that user experience should be studied in more holistic way and that the users' subjective views to the experiences should be taken into account and analyzed in addition to short-term experiences [4], [12], [13]. Although the existing methods provide valuable data about user experience, they may fragment user experience into small elements. Our purpose, instead, is to evaluate experiences that users find themselves meaningful for them.

Environment
Personal
Task
Social
Spatio-Temporal
Terminal
Service
Access Network

Table 1. Context categories used in the analysis [6]

Experience Reports as a Data Collection Method

Experience reports are a form of storytelling, where users report meaningful or memorable events with their products. Storytelling has been regarded as an effective way to get a holistic view of the user experience [4], [3]. Users remember and organize experiences as stories and those stories enable humans to communicate experiences in different situations to the particular people involved [3]. Wright et al. also state that recounting experiences is part of making sense of the experience [13]. Storytelling has been previously used e.g. as part of narrative interview [4] or to complement data that has been collected by other methods [8].

In this study, we wanted to exploit the benefits of storytelling, but without requesting participants to write complete stories with all typical story elements [5]. Instead, participants should describe significant events of their daily use of a product as they would tell about it to somebody else. We asked participants to write about the actual usage but also about their personal feelings, values and interests when using a product. In addition, we were interested in the situation where the story happened or other contextual information related to the event.

The resulting experience reports were typically less than half of an A4 sheet of typed text. They may not necessarily be the complete descriptions of the user experience, but they relay salient parts of an experience that participants consider important and meaningful, which makes experience reports a novel method for collecting user experience data about the product use.

The Study

For the field study, we recruited 21 persons (10 female, 11 male) through a common distribution list to participate in the

study. The average age of the participants was 34 years, ranging from 25 to 46 years. The participants were asked to propose their own favorite products for the study. The repertoire of the devices included smart phones, internet tablets, MP3 players, heart rate monitors, and digital cameras.

We expected the participants to be familiar with the product features and the usage of the product should be an integral part of their daily routines. Most of the participants were indeed experienced users of the products. On average, the participants had used their personal products for 17 months, and 25% of the participants had used the product actively for more than two years.

In the beginning of the study we arranged a one-hour introduction session to describe the purpose of the study and to instruct on how to report experiences. We presented an illustrative experience report to give an idea of the style of the report as well as how their feelings could be described in a report. The participants were advised to write reports during the same day or in the following day in order to capture the details of their meaningful user experiences.

In the 10 days of the study, the participants wrote a total of 116 experience reports. Both female and male participants were equally active in reporting their experiences. 53% of reports were written by male participants. The participants wrote 5.52 reports on average (median value was 6).

The Analysis Process

After the reporting period, the experience reports were analyzed independently by two researchers. The objective of the data analysis was to identify core experiences that emerged from the reports. The first step in the analysis was

1. Captivation
2. Challenge
3. Competition
4. Completion
5. Control
6. Cruelty
7. Discovery
8. Eroticism
9. Exploration
10. Expression
11. Fantasy
12. Fellowship
13. Humor
14. Nurture
15. Relaxation
16. Sensation
17. Simulation
18. Submission
19. Subversion
20. Suffering
21. Sympathy
22. Thrill

Table 2. Playful Experience (PLEX) categories [10]

to identify the “triggering context” which was expected to significantly affect the experience. Although there can be multiple context elements in a report, there is usually a single element which can be seen as triggering the most meaningful user experience, i.e. the core experience. We used context elements defined by Guarneri et al. [6] to categorize the triggering contexts (Table 1). By first identifying the context element helped us to focus on a core experience in the report.

In the next step the reports were analyzed sentence by sentence to identify issues that would describe the core experience. The identified keywords and expressions were mapped against Playful Experience (PLEX) categorization (Table 2) and their descriptions [10]. These categories have been previously used to identify playful user experiences with digital games, and we wanted to explore whether they could also be used to describe experiences from other kinds of products. Reports that did not map to any of the PLEX categories were noted as potential new categories.

After the individual evaluators’ analysis was completed, both results were combined and conflicting cases were discussed to find agreement on the described core experiences.

Although the evaluators shared an understanding of the triggering context and PLEX categories, it was discovered that the interpretations varied and the Kappa value (.34) showed only moderate agreement. This was mainly caused due to different emphases in the stories. In addition, some experience reports had multiple possible interpretations.

Examples of Experience Report Analysis

Out of the 116 experience reports, we present three examples and describe how they were interpreted. The experience reports are presented as they were written by the

participants. The underlining highlights key aspects that were considered important when defining the core experience.

Report #1 [female, 43 years]

“The other summer, my in-laws with their 16-year-old daughter rented a beautiful cottage by the lake. But was their daughter happy? No way, being buried there in the middle of nowhere, while all her friends were in the city. And there was wailing and gnashing of teeth. But then I, the cool Auntie, arrived with my E90+N810. Now the daughter could browse here and there and email with all her friends. There she was, all curled up there in the corner of the sofa, quietly and smiling, tapping and clicking away. She was happy, her parents were happy, we were happy.”

In the report, the user is taking care of a teenager’s social well-being. The user is describing more the social situation and the experience is a result of being able to make someone else happy. The personal device is used as a tool to achieve this. Both evaluators agreed that the triggering context was *Social* and the core experience was *Nurture*.

Report #2 [male, 33 years]

“Okay, random play is officially my new favourite feature of the iPod. I can’t believe how many gems I have hidden in this little device. I was getting really bored with my music selection, or so I thought. Turns out I was getting bored with the sequence I was listening to the songs, but not it’s like the entire library has been given new live and it’s exciting for me to listen to my music again. It’s quite invigorating to go from Irish folk music to Norwegian black metal to Dutch prog-opera metal ; -)”

In this report, the evaluators initially disagreed on whether the triggering context was *Task* or *Terminal*. On one hand, it

could be argued that the report is about the task of accessing new music. On the other hand, the reported experience was about “the favorite feature” of the device which helped discovering new music. *Terminal* was agreed to be the more appropriate interpretation for the triggering context and the core experience was defined as *Discovery*.

Report #3 [male, 28 years, translated]

“On Tuesday morning, riding once again to work with the bicycle. However, I went through a somewhat longer route. The normal routines when departing included strapping on the meter. You have to store the execution of the weekly exercise program to the meter when you have acquired one. While bicycling, I thought that I should again begin to bring the exercises to the web exercising diary. That could be a good idea if starting to participate to a local running competition. The only inconvenient thing is the data transfer. Transmitting through Polar’s voice link is a bit troublesome and time-consuming, but on the other hand I could begin using exercise programs offered by the service instead of just those prepared by the meter. Perhaps. I arrived there at the office and the round took a bit beyond half an hour. The meter confirmed my feelings during the bicycling in that I had recurred from Sunday’s run. Tomorrow I could then run a somewhat more intense route.”

In this report about a Polar heart rate monitor and belt, the evaluators originally focused on different parts of the story, which resulted in completely different analyses of the report. The first evaluator focused on data transfer part and interpreted *Access Network* as a triggering context and *Suffering* as a core experience because of the poor data transfer speed. The second evaluator focused on the user’s personal feelings about his physical condition and therefore, concluded that the triggering context would be *Personal* and

core experience is *Control*. Both analyses were seen valid and this story was given two core experiences.

Summary and Discussion

Our primary goal for the study was to test experience reports as a UX evaluation method. We asked the users to describe meaningful user experiences that emerged from the interaction with their personal products. During the data analysis, it became apparent that the participants are not only describing experiences that are a result of *interaction* with the device, but there are also other types of experience reports. Some experience reports were emphasizing the situation or context in which the participants were using the product, and how they experienced that situation, rather than product interaction. In some other reports, the participants were describing experiences that are related to the role of the product in their lives without referring to any specific interaction or situation. Based on this initial analysis, we grouped experience reports into three types and labeled them *Interaction Experience* (e.g. Report #2, 43% of the reports), *Context Experience* (e.g. Report #1, 33%), and *Product Experience* (20%) reports. The remaining 4% of the reports were of indistinguishable type, e.g. the story was so short and non-descriptive that no experience could be identified.

The most common experiences found according to PLEX categorization were *Control*, *Simulation*, and *Fellowship*. The PLEX categories covered well the experiences in the interaction experience reports. 91% of those reports could be evaluated with the categories. In context experience reports, PLEX categories managed to identify elicited user experiences quite comprehensively (76%). Product experiences were not covered very well (only 20%) and this is one of the areas that need to be developed further in this method.

The participants reported their experiences actively, probably because they were using their own personal products on a daily basis and therefore it was easy to write about their meaningful experiences. The participants were free to make the reports as long as they liked and the format of the report was also unspecified. We believe that some reports caught the true user experience better because of this freedom.

The analysis process was relatively simple, but it had some challenges. The first step, i.e. context element identification, was quite straightforward. As there were only eight categories to choose from and all of them are distinct, it is easy to identify the triggering context in most of the reports. Identifying the triggering context is important because it will also determine what the core experience in the report is. Identifying the exact core experiences was a more challenging task. Although both evaluators were familiar with PLEX categories used in the analysis, different core experiences were initially concluded. There are two main challenges in this process. First, the experience reports do not follow any specific structure and identifying different keywords will result in different core experiences. Second, PLEX categories describe playful aspects of the experiences which were not strongly present in the majority of the experience reports. In the future, we will define in more detail the procedure to determine key aspects in the reports and revise the PLEX category descriptions to cover the experiences more broadly.

Based on the results of this initial study, we believe that experience reports can be developed to a valuable tool to evaluate UX. These reports describing users' meaningful experiences, with the context and experience categorizations in the analysis, will increase product designers' and researchers' understanding of UX.

References

- [1] Csikszentmihalyi, M., Larson, R., Prescott, S. The Ecology of Adolescent Activity and Experience. *Journal of Youth and Adolescence* 6, 3 (1977), 281-294
- [2] Desmet, P., Overbeeke, C. J., Tax, S. J. E. T. Designing Products with Added Emotional Value; Development and Application of an Approach for Research through Design. *The Design Journal* 4, 1 (2001), 32-47
- [3] Forlizzi, J., Ford, S. The Building Blocks of Experience: An Early Framework for Interaction Designers. In *proc DIS ACM Press* (2000), 419-423
- [4] Geven, A., Schrammel, J., Tscheligi, M. Narrations and Storytelling as Methodological Key Elements for Studying User Experience. In *proc NordiCHI COST294-MAUSE* (2006), 79-83
- [5] Gruen, D., Rauch, T., Redpath, S., Ruettinger, S. The Use of Stories in User Experience Design. *International Journal of Human-Computer Interaction* 14, 3&4 (2002), 503-534
- [6] Guarneri, R., Sollund, A. M., Marston, D., Fosbak, E., Berntsen, B., Nygreen, G., Gylterud, G., Bars, R., Kerdraon, A., *Eperspace – Ist Integrated Project*, 2004, Available at <http://www.ist-eperspace.org/deliverables/D5.1.pdf>
- [7] Hassenzahl, M., Brumester, M., Koller, F. Attrakdiff: Ein Fragebogen Zur Messung Wahrgenommener Hedonischer Und Pragmatischer Qualität. in *Mensch & Computer 2003. Interaktion in Bewegung*, Ziegler, J., Szwillus, G. (Eds) B.G. Teubner (2003), 187-196
- [8] Isomursu, M., Kuutti, K., Väinämö, S. Experience Clip: Method for User Participation and Evaluation of Mobile Concepts. In *proc PDC ACM* (2004), 83-92
- [9] Kankainen, A. Ucpd: User-Centered Product Concept Design. In *proc DUX* (2003), 1-13
- [10] Korhonen, H., Montola, M., Arrasvuori, J. Understanding Playful User Experience through Digital Games. In *proc DPPI Université de Technologie de Compiègne* (2009), 274-285
- [11] Mäkelä, A., Fulton Suri, J. Supporting Users' Creativity: Design to Induce Pleasurable Experiences. In *proc International Conference on Affective Human Factors Design* (2001), 387-394
- [12] Norman, D. A., Memory Is More Important Than Actuality, *interactions* 16, 2 (2009), 24-26
- [13] Wright, P., McCarthy, J., Meekison, L. Making Sense of Experience. in *Funology: From Usability to Enjoyment*, Blythe, M., Monk, A. F., Overbeeke, K., Wright, P. (Eds) Kluwer Academic Publisher (2003), 43-53