
Comparison of UX Evaluation Methods That Measures the UX over Time

Ayako Hashizume

Tokyo Metropolitan University
Tokyo, Japan
hashiaya@tmu.ac.jp

Masaaki Kurosu

The Open University of Japan
Chiba, Japan
masaakikurosu@spa.nifty.com

Yuuki Ueno

Otsuka Business Service
Tokyo, Japan
uenoyuuki@gmail.com

Copyright is held by the author/owner(s).
CHIItaly '17, September 18-20, 2017, Cagliari, Italy.

Abstract

Evaluation of UX is deeply related to the concept of UX. UX cannot be evaluated directly from the quality in design but should be evaluated from the quality in use. As a result, the evaluation of UX should use different tools from what usability professionals have been using for evaluating the usability. There are two types of UX evaluation methods, one is the real-time method and another is the retrospective method based on memory. Because of the difficulty of conducting the real-time method for a long time, authors adopted the retrospective approach. We first focused on the UX Curve and revised it to become the UX Graph. Authors analyzed the accumulated data obtained so far by applying the UX Graph and concluded that the time dimension (horizontal axis of the graph) and the graph itself was not of much importance. Instead, they proposed the Experience Recollection Method (ERM) which accepts ambiguity in the time of use. In this paper, we compared UX evaluation methods, and clarified the advantages of the ERM.

Author Keywords

user experience (UX); evaluation; Experience Recollection Method (ERM); rating scale, satisfaction

ACM Classification Keywords

H.5.2 User Interfaces Evaluation/methodology.

Introduction

UX is the perception and reaction of users in the real environment for real products and services. UX evaluation should adopt tools and methods that are different from what have been used for evaluating the usability. Although many evaluation methods have been proposed as can be found on the "AllAboutUX" web site [1], there are many usability evaluation methods intermingled with the UX evaluation methods with little consideration of the difference between the concept of usability and UX. Based on our model of usability and UX, we defined that the usability is a part of the "quality in design" and the UX is related to the "quality in use" [7].

The purpose of UX evaluation is to grasp changes in actual experiences; hence short-term UX evaluation is not sufficient. Among methods used for evaluating UX, there are real-time methods and retrospective methods. Real-time methods include the use of the Experience Sampling Method (ESM) [4] using mobile phone, but it's difficult to conduct for a long time. On the other hands, retrospective methods including UX Curve and UX Graph fit well to the goal of UX evaluations. Because we thought that the ultimate evaluation measure is the satisfaction [6], we adopted it as the scale for evaluating the UX [8]. Thus we proposed the UX Graph method [2, 8] as a revised version of UX Curve [5], and then, the ERM (Experience Recollection Method) [8, 7] as an advanced method rooted in the UX Graph.

In this paper, we picked up typical timeline-based [9] UX evaluation methods and made a comparison among them.

A New Method for UX Evaluation the "ERM"

The UX Curve is a method to let users draw the curve from the start of use till the time of measurement in terms of the attractiveness, usability, utility and the frequency of use by assigning those values to the vertical axe and time to the horizontal axe. While it is adequate to ask real users their experiences in the real context, to provide the curve that is visual and impressive and to facilitate the grasp of the general tendency by a glance, it has some issues that should be reconsidered as follows: (1) the focus is rather on the curve itself and the content of each episode is not much focused, (2) drawing similar three curves in addition to the frequency is time-taking and boring. Thus authors developed the UX Graph with the idea of improving the UX Curve [5].

In the UX Graph, only the satisfaction that is regarded as the same with 'utility' in the context of economics is used as the sole measure. In the study on the concept of satisfaction [6], it is described that the satisfaction is an integrative concept of hedonic aspects and all the quality characteristics as well. Besides, the coordinate of each episode is thought to be important and the graph is drawn connecting those points on the sheet. In other words, the graph as the visual expression is just in a supplementary position. A WEB tool was developed to draw the UX Graph. Based on the collection of data, authors summarized the evaluation of the UX Graph as follows: (1) the visual expression is attractive but is not much substantially important, and (2) the memory of informants in terms of the time is ambiguous and the value on the horizontal axes changes the shape of the graph drastically.

Thus we decided to waste the visual expression and not to ask the exact time to informants. This is the process how we developed the Experience Recollection Method (ERM) [8, 7].

In the ERM, informants are not asked to draw the curve nor the graph. Time line is not the exact year but the roughly sectioned periods as can be seen in Figure 1. Informants are requested to write down the episodes with the satisfaction rating from +10 to -10.

Comparison of UX Evaluation Methods

Table 1 summarizes UX evaluation methods that measure the UX over time. Following methods are included in the table:

Recording Sheet for ERM: Experience Recollection Method Target Item Smartphone (iPhone 6) Male/Female Age 27

1. Write what you experienced at each phase and fill in the evaluation by +10 to -10 rating.

Phase	What you experienced	Evaluation (+10~ -10)
Expectation before purchase	I expected to get the latest model of iPhone on the day of sale.	8
Evaluation at the time of starting to use	Year 2014 I was bewildered for the larger screen compared to my previous model (iPhone5).	5
Evaluation at early days from starting to use	I got used to the large screen soon. And I felt the advantage of large screen for enjoying the game.	10
Evaluation during the use	The body was bent, but was straighten back by pushing it harder.	5
Recent evaluation	The power loss of battery is unexpectedly fast.	-5
Present evaluation	Year 2016 It's now a must to carry the backup battery.	-5
Evaluation in the near future	I will use this until the next model will appear.	-2

Figure 1: An example of ERM regarding the using smartphone

Method	a.		b.			c.						
	Nature of Method		Type of Method			Phase						
	Qualitative	Quantitative (Timeline)	Verbal	Hand-written	PC-based	Expectation	Purchase / Obtain	First Phase	Middle Phase	Last Phase	Current	Future
CORPUS [11]	○ (interview)	○	○				○	○	○	○	○	
iScale [3]		○			○		○	○	○	○	○	
UX Curve [5]		○		○			○	○	○	○	○	
UX Graph [2, 9]		○		○ (both)	○	○	○	○	○	○	○	○
ERM [9, 8]		○		○ (both)	○	○	○	○	○	○	○	○

Table 1: Comparison of six timeline-based UX evaluation methods

CORPUS [10], iScale [3], UX Curve [5], UX Graph [2, 8], and ERM [8, 7].

The column of Table 1 (a.-e.) are following aspects for clarifying the characteristics of each method.

- a. Nature of Method - Qualitative (i.e. interview) / Quantitative (i.e. numerical measures be obtained)
- b. Type of Method – Verbal / Hand-written / PC-based
- c. Period – Expectation / Purchase or Obtain / First phase / Middle phase / Last phase / Current / Future (i.e. regarding which time phase, the method can get the data)

Regarding “b. Type of Method”, PC-based method facilitates the acquisition of data from informants living in the remote area if it is conducted online by using the internet. In this respect, the iScale (if it can be used on the internet), UX Graph and ERM are better. Regarding “h. Phase”, it is necessary to include the expectation because it should be regarded as a part of UX as in the UX Graph and ERM.

Comments for ERM

Comments obtained from 53 university students included positive comments such as “the table (not the graph or curve) is suitable for reminding the past events”, “it is easy to use because episodes are not always be remembered in the chronological order”, etc.

Conclusion

Considering the characteristics of each UX evaluation method, each method has its own uniqueness. ERM, of course, has its own uniqueness and merits, thus should be used by many practitioners in the real life UX situation.

References

1. AllAboutUX <http://www.allaboutux.org/all-methods>
2. Ayako Hashizume and Masaaki Kurosu. 2016. "UX Graph Tool" for Evaluating the User Satisfaction, International Journal of Computer Science Issues, 13(5): 86-93.
3. Evangelos Karapanos, John Zimmerman, Jodi Forlizzi and Jean-Bernard Martens. 2010. Measuring the Dynamics of Remembered Experience Over Time, Interacting with Computers, doi: 10.1016/j.intcom.2010.04.003
4. Reed Larson and Mihaly Csikszentmihalyi, 1983, The experience sampling method, Jossey-Bass, 41-56.
5. Sari Kujala, Virpi Roto, Kaisa Vaananen-Vainio-Mattila, Evangelos Karapanos and Arto Sinnela. 2011. UX Curve: A Method for Evaluating Long-term User Experience, Interacting with Computers, doi:10.1016/j.intcom.2011.06.005
6. Masaaki Kurosu, 2015, Satisfaction as a Measure of Quality in Use for Kansei Experience, International Symposium on Affective Science and Engineering 2015 Proceedings
7. Masaaki Kurosu, 2016, Theory of User Engineering, CRC Press
8. Masaaki Kurosu, Ayako Hashizume, Yuuki Ueno, Tuyoshi Tomida and Hirotohi Suzuki, 2016, UX Graph and ERM as Tools for Measuring Kansei Experience, HCI International 2016 Proceedings
9. Tonya J. Tidline, 2005, Dervin's Sense-making, in Karen E. Fisher, Sanda Erdelez, and Lynne McKechnie (eds.), Theories of Information Behavior (Asist Monograph), Information Today Inc.
10. Margeritta von Wilamowitz-Moellendorff, Marc Hassenzahl, and Axel Platz, 2006, Dynamics of User Experience: How the Perceived Quality of Mobile Phones Changes Over Time". pp, 74–78.