Position paper for Cultural Impact on User Experience Design and Evaluation, NordiCHI2012 – One Day Workshop – draft only, pls do not cite or share

User experience measurement in the wild

Zhengjie Liu[1] & Torkil Clemmensen[2]

1 Dalian Maritime University, Dalian, China
2 Copenhagen Business School, Denmark

1   Introduction

User experience measurement is important for HCI research and for the industry that produce interactive services and products. We want to develop a context-aware user experience measurement. In this paper, we propose a user experience measurement scale, and discuss how to validate the scale in different contexts, implement the scale in context aware prototypes, and evaluate the prototypes in real-life user contexts.

The main contribution of this paper is to combine existing concepts of user experience and context-aware computing in one context aware user experience measure. We are not going to develop new definitions of neither UX nor context-awareness, but simply combine the two.

The background for our paper is the importance of enabling organizations to provide high quality interactive experiences to customers and citizens; to empower citizens in our emerging digital societies to assess and be critical towards interactive solutions presented to them; and to support HCI researchers and others with reliable and valid instruments to score, compare, and analyse user experience.

We focus on measurement of UX in the shopping context(s).

1.1   General concepts and definitions of UX

System usage can be conceptualized as for example duration, frequency, and intensity of the use of a particular computer system (Venkatesh, Brown,

Maruping, & Bala, 2008), which however, seems too narrow to capture the many different relations between people and computers in real life, see e.g., (Rogers, 2011). In contrast, UX has emerged as a broad concept that covers a more complex set of human-computer relations.

UX has been described as a positive emotion (an emotional appraisal) related to a specific interaction event, but it would also involve the user's background, their experience, the criticality of their decision to use a given interactive system, as well as the content, the usability, and the aesthetics of the interaction with the computer. All of these would vary between users, between organizations, and between work domains.

Thus the many definitions of UX range from a focus on the ways people retrospectively judge their use of interactive products, to include all aspects of users' interaction with product vendors' services and products (see e.g. Law, Roto, Vermeeren et al., 2008; Law, Roto, Hassenzahl et al., 2009). Despite this lack of clarity concerning what UX is and what it comprises, there is some consensus that it captures the dynamics of experience, and how interactive products, person characteristics, and context blend to shape "the experience of use" (Bargas-Avila & Hornbæk, 2011). Law et al.'s (2009) survey of 275 UX researchers and practitioners led them to suggest that the term UX be "scoped to [interactive] products, systems, services, and objects that a person interacts with through a user interface" (p. 722). Their results agree with Bargas-Avila and Hornbæk's (2011) assertion that UX is dynamic, context-dependent, and subjective, "which stems from a broad range of potential benefits users may derive from a product" (Law et al., 2009, p. 722). This dynamic, subjective and context-dependent UX is what we want to help capture.

1.2   Needs for UX measurement

So far existing measures concern usability, se fx (Brinkman, Haakma, & Bouwhuis, 2009; Brooke, 1996). However, user experience is different from usability. Usability and user experience are related but different concepts. Usability has a functional flavour, emphasizing work actions and business effectiveness in the use of IT. User experience is more

closely related to users' aesthetics preferences and decision making, and on interaction design. There is a need for a simple way to assess the broadly defined user experience (Porat & Tractinsky, 2011).

1.3   UX is attached to contexts, so is the measurement

UX is attached to contexts, so the measurement also should be context-based. Because UX is about the users' experience of the interaction with different technologies over time (Karapanos, Zimmerman, Forlizzi, & Martens, 2009), it is decisive that the UX measurement is done "in the wild"(Law & Sun, 2012; Roto, Väätäjä, Jumisko-Pyykkö, & Väänänen-Vainio-Mattila, 2011), that is, in different contexts that people encounter in life and all across the products lifecycle (Kujala, Roto, Väänänen-Vainio-Mattila, Karapanos, & Sinnelä, 2011).

2   What is context? (related work and our definition)

Context as a concept has been studied in many relevant disciplines and areas. Context is often thought of as "out there" or "in the wild", that is, outside the researchers laboratory, out in peoples everyday life (Rogers, 2011).

Context can be temporal context – over time – or in terms of the users task, in terms of use scenarios, in terms of different technologies (embodied etc), in social context (fx social media), and more. Context can be user related, task related and environment (physical, social, technical etc.) related. We want to use the concept of context in a way similar to how it has been used by Activity theory research (Bardram, 2009), that is, context is the overall human activity within which the action and operations are carried out.

A particular context of interest in this project is national cultural context, that is, the overall human activity compared across countries. In addition, different stakeholder groups and nationalities may construe usability and user experience slightly differently (Hertzum & Clemmensen,

2012; Hertzum et al., 2011), which increase the requirement for a truly international and adaptive measure of user experience.

[Needs for tools – UX measurement in the wild needs tools] Traditional f2f user research methods cannot meet the vast demands for UX measurements in the wild. For industry, in comparison with f2f methods, tools to be used by users or to be used as remote evaluation, are less costly, can be used for more products and more participants, faster, and can provide quantitative data preferred by engineers. Furthermore, UX happens in the real world, not only in the lab, and mobile and ubiquitous technology allows us to explore new ways of UX measurement. For research, so far UX research has focused on mobile devices and websites, and our proposed tools will allow the researcher to focus on a larger diversity of technology, and cross-technology use. Furthermore, the tool will allow the researchers to collect reliable data longitudinally on UX. So we need to develop tools for assistance.

2.1 Tools need to be context aware

UX is context sensitive, and the UX scale-based measurement (ratings, rankings) data need to be collected in certain contexts, and analyzed and interpreted in relation to these contexts. When we analyzed the UX data, we need to go back to certain contexts, to know from which contexts the data are from. When we try to collect and analyze context sensitive UX data, we need certain tools to collect these data. Human observers will usually know the context they are in, but the tool can help make this context explicit. A context- aware tool can help collect (and analyze) observational data of user experience in meaningful ways.   Human researchers know well when to collect what, and analyze and interpret the data based on the contexts. This is done in order to ensure that the UX data is collected, analyzed and interpreted in a contextually meaningful way. The tools need to be able to do the same. It should be able to collect the data that is appropriate for the context at the moment and support the human researchers in data analysis and interpretation with the connections to the contexts.

3   Methodology: Context-aware UX measurement

We want to develop a context-aware user experience measurement. With this grant paper we suggest a user experience measurement scale, validate the scale in different contexts implement the scale in context aware prototypes, and evaluate the prototypes in real-life user contexts.

3.1   What is a UX measurement scale? (objectives)

An UX measurement scale is a psychometric scale that purports to measure the characteristics of the interactive use experience. This may include, but is not limited to, measuring the users positive emotions related to a specific interaction event, the relevance of the user's background and experience, the criticality of the use of this interactive system, as well as the content, the usability, and the aesthetics of the moment of interaction with the system. Each of these areas may require a separate subset of scales or questions to measure, just like intelligence or personality frequently are measured by a number of separate factors, each with a set of subscales. The aim of UX measurement is to identify a set of desired user experiences.

We will build on our UX measurement on the broad definition of UX given above and plan to design the scales based on our previous research on the words that different stakeholder groups and nationalities use to describe their interactive experiences (Hertzum & Clemmensen, 2012; Hertzum, et al., 2011). The development of UX sub-scales will follow psychometric scale development principles. The psychometric properties of the scales will be described and the validity and reliability of the scale compared to existing system usability scales such as (SUS), SUMI, and WAMMI. We plan to develop both long and short versions of the UX scale.

The aim is to let the UX scale cover different human-computer relations such as embodied use (like mobile phones), hermeneutic use (like tools, like laptop), alterity use (like social media), and background use (like interactive aircondition or heating systems). Furthermore, the UX scale should use small tour questions (e.g., "how was this interaction?") to cover concrete experi-

ences, and use grand tour questions (e.g., "how are such interactions generally?") to cover types of experience.

The UX scale will be adapted to different use contexts.

3.2 What is a context aware UX measurement? (objectives)

It is like tagging or adding a researcher on the back of the user – this researcher can then observe the users interactions during the day and can decide which contexts the user is now in, and collect the data that are meaningful for this context, and keep a trace to the context where these data are from. The tool should be like the researcher.

By context aware UX measurement, we mean to develop a kind of tool on the mobile platform that assists the researchers to collect UX data in a context aware manner. It runs on the participants' mobile devices and can sense and identify the contexts that are appropriate for UX measurement and collect the appropriate UX data from that context. While collecting and maintaining the data, it can keep a trace back to the contexts so as to support the following up data analysis and interpretation in a contextually meaningful way. The context aware UX measurement tool to be developed will be able to be used for cross context (measure all technology use during one day) or single context (measure one technology use only) situations. It will also work for different granularity (how fine do we want to measure (hour, minute, session, sampling). The context awareness technology will also help it to be non-intrusive or non-interrupting technology to the users by adapt measurement time to fit the user behavior patterns.

Steps for developing of a UX scale, based on (DeVellis, 2011):

1. Define clearly what we want to measure
   (a) Review and select among existing UX definitions
   (b) Base selection of previous UX construct research (Hertzum & Clemmensen, 2012; Hertzum, et al., 2011).
2. Create a set of draft questions.
   (a) From selected UX definition and previous research, create questions
3. Select media, format and answer options for the questions

(a) Implement in tool prototypes (done by DMU)

(b) Implement Chinese and Danish and English? language versions

4. Have experts review and revise the questions.

(a) Interviews and Think aloud usability tests of tool with experts

5. Consider questions for different contexts

(a) Use theory and tool-based context definition

6. Field test the questions with "real people."

(a) Test of scale-in-tool in defined contexts

7. Analyze the results of your field test.

(a) Statistical treatment of results

8. Decide how many questions--and which questions--to keep.

(a) Fine-tuning scale to fit tool and contexts

Steps in the development of context aware UX measurement app (user centered development):

1. study UX researchers'(developers, usability professionals, users, HCI researchers) practice to understand how they collect UX measurement data in contextually sensitive ways. Various possible relevant UX evaluation activities of varying granularity will be taken into consideration, like session related, task related, UI component related, physical environment related etc.
2. develop a definition of the context and a model for the context awareness that fit for the UX measurement practice
3. adapt the context awareness model to the UX measurement scale
4. develop a context aware UX measurement tool
5. iteratively evaluate and improve the tool in low fi to high fi prototypes
6. implement the tool in real technologies
7. field trials in real world

3.3   How to develop context aware UX measurement

We aim at user centered development of scale-app with low fidelity prototypes and high fidelity prototypes, and evaluation in four different use situa-

tions (embodied use, hermeneutic use, alterity use, background use) in Denmark and China. Furthermore, evaluation of a hi fi prototype in one use situtation in Denmark and China.

4   Expected results

We expect results of our future research to be a validated approach for context aware UX measurement. In particular we want to a) compare tool use with existing usability and UX scales, b) compare tool use with traditional f2f paper based user study, c) compare tool use with remote user study tools, and c) tool works in different cultural groups.

5   Conclusion

In conclusion, we want to propose research aimed at developing context aware UX measurements. The format of the measure might be an app.

We look forward to discussing the cultural aspect of our proposal with the workshop participants.

Bardram, J. E. (2009). Activity-based computing for medical work in hospitals. *ACM Transactions on Computer-Human Interaction (TOCHI), 16*(2), 10.

Brinkman, W. P., Haakma, R., & Bouwhuis, D. G. (2009). The theoretical foundation and validity of a component-based usability questionnaire. *Behaviour and Information Technology, 28*(2), 121-137.

Brooke, J. (1996). SUS-A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & A. L. McClelland (Eds.), *Usability Evaluation in Industry*. London: Taylor and Francis.

Clemmensen, T. (2012). Usability problem identification in culturally diverse settings. *Information Systems Journal, 22*(2), 151-175.

DeVellis, R. F. (2011). *Scale development: Theory and applications* (Vol. 26): Sage Publications, Inc.

Hertzum, M., & Clemmensen, T. (2012). How do usability professionals construe usability? *International Journal of Human-Computer Studies, 70*(1), 26-42.

Hertzum, M., Clemmensen, T., Hornbæk, K., Kumar, J., Shi, Q., & Yammiyavar, P. (2011). Personal Usability Constructs: How People Construe Usability Across Nationalities and Stakeholder Groups. *International Journal of Human-Computer Interaction, 27*(8), 729-761.

Karapanos, E., Zimmerman, J., Forlizzi, J., & Martens, J. B. (2009). User experience over time: an initial framework *CHI2009* (pp. 729-738): ACM.

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

Law, E. L. C., & Sun, X. (2012). Evaluating user experience of adaptive digital educational games with Activity Theory. *International Journal of Human-Computer Studies*.

Porat, T., & Tractinsky, N. (2011). It's a Pleasure Buying Here: The Effects of Web-Store Design on Consumers' Emotions and Attitudes. *Human–Computer Interaction*.

Rogers, Y. (2011). Interaction design gone wild: striving for wild theory. *Interactions-New York, 18*(4), 58.

Roto, V., Väätäjä, H., Jumisko-Pyykkö, S., & Väänänen-Vainio-Mattila, K. (2011). *Best practices for capturing context in user experience studies in the wild.* Paper presented at the MindTrek '11 Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments

Venkatesh, V., Brown, S. A., Maruping, L. M., & Bala, H. (2008). PREDICTING DIFFERENT CONCEPTUALIZATIONS OF SYSTEM USE: THE COMPETING ROLES OF BEHAVIORAL INTENTION, FACILITATING CONDITIONS, AND BEHAVIORAL EXPECTATION. [Article]. *MIS Quarterly, 32*(3), 483-502.