

To Measure or Not to Measure UX: An Interview Study

Effie Lai-Chong Law
University of Leicester
Dept. of Computer Science
LE1 7RH Leicester, UK
elaw@mcs.le.ac.uk

Paul van Schaik
Teesside University
School of Psychology
TS1 3BA Middlebrough, UK
P.Van-Schaik@tees.ac.uk

ABSTRACT

The fundamental problem of defining what UX is (or is not) has a significant influence on another challenging question: *to measure or not to measure UX constructs*. The answer of most, if not all, UX researchers and practitioners, would probably be “It depends!” As we were motivated to find out “depending on what”, we conducted semi-structured interviews with eleven UX professionals where a set of questions in relation to UX measurement were explored. Participants expressed scepticism as well as ambivalence towards UX measures and shared anecdotes related to such measures in different contexts. To improve the interplay between UX evaluation and system development, a clear definition of UX, combining various data types, and robust education in UX concepts are deemed essential.

Author Keywords

User experience; Measurement; Interview; Feedback loop

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Design; Evaluation; Measurement

INTRODUCTION & BACKGROUND

To measure or not to measure concepts of interest? A clear cut “*Yes!*” if this question is raised in the context of physical sciences whereas an ambiguous “*It depends!*” when it is addressed in the context of social sciences in general and the emerging research area of User Experience (UX) in particular. We aimed to explore such stipulations (i.e. ‘depending on what’) for UX measures and their implications to design and evaluation of interactive systems. To meet this purpose, we conducted an empirical study in which eleven UX researchers and practitioners were interviewed. In this paper we report some main findings of the study that are particularly relevant to understanding the interplay between UX measurement and iterative system redesign. Specifically, we adopt Hand’s ([4], p.3) definition of measurement “*quantification: the assignment of numbers to represent the magnitude of attributes of a system we are studying or which we wish to describe.*”

The exploration of the issue of UX measurement was embarked on (e.g. [6]) after another, if not more, thorny issue of UX - its multiple definitions - had been examined [7]. In principle these two foundational issues should be

solved in tandem. However, the definitional issue on UX remains unresolved, UX researchers and practitioners tend to select and adapt one of the many definitions out there to serve their particular goals and needs. The recent efforts of deepening the understanding of the theoretical roots of UX [10] can complement the earlier work on UX evaluation methods on the one hand [13] and the current operationalisation work for UX measurement on the other hand (e.g. [11]). As UX research studies have hitherto relied heavily on qualitative methods [1], the progress on UX measures has thus been slow. A plausible reason is the scepticism about the measurability of UX.

The field of HCI in which UX is rooted has inherited theoretical concepts, epistemological assumptions, values, and methodologies from a diversity of disciplines, ranging from engineering where measures are strongly embraced (cf. William Thomson’s [14] dictum ‘to measure is to know’) to humanities where measures can be regarded as naïve or over-simplistic, especially when the concepts to be measured are ill-defined, leaving (too) much for interpretation [2]. As UX subsumes a range of fuzzy experiential qualities such as happiness, disgust, surprise and love, controversies and doubts about the measurability of UX are inevitable. A main divergence between two major camps of UX researchers is the legitimacy of breaking down experiential qualities into components, rendering them to be measured; it is rooted in the age-old philosophical debate on reductionism versus holism.

INTERVIEW ON UX MEASUREMENT

Instrument

The interviews were semi-structured with 12 questions grouped into three main parts. Part A comprises four background questions (Table 1).

Q1. Gender: Female, Male
Q2. Age: <=20, 21-30, 31-40, 41-50, >50
Q3. I am a: Practitioner, Researcher, Student, Other
Q4. How long have you worked in the area of UX? (Never, <1 year, 1-3 year, 3-5 year, >5 year). Please describe the topic and related work.

Table 1. Background questions

Part B comprises five questions on the measurability of UX qualities (Table 2). The inclusion of Q5 is to know if the respondent’s understanding aligns with any of the existing definitions of measurement. For Q6, the rationale underpinning each statement varies. The first one was derived from the classic justification for measurement

advocated by Thomson [14]. The second and third ones were two rather extreme views against UX measures expressed in some informal contexts (e.g. group discussion in a workshop). They were aimed to stimulate thoughts and should not be treated as scientific claims. In contrast, the fourth and fifth statements represent views on the potential uses of UX measures. They were deliberately broad in scope to stimulate discussions.

Q5. What is a 'measure'?
Q6. (a) Please rate your agreement with each of the following statements (5-point Likert scale); (b) Explain your ratings <ul style="list-style-type: none"> ▪ UX measures lead to increase of knowledge ▪ UX measures are insane ▪ UX measures are a pain ▪ UX measures are important for design ▪ UX measures are important for evaluation
Q7. (a) Name a specific experiential quality (e.g., fun, surprise) that is most relevant to your work; (b) Explain the relevance; (c) Do you think the named quality can be measured: If 'yes', describe how; If 'no', describe why.
Q8. (a) Name an experiential quality that you are (almost) certain is measurable; (b) How can it be measured and when (before/during/after interaction)? (c) Why are you so (almost) certain about its measurability? What is your reservation, if any?
Q9. (a) Name an experiential quality that you think (almost) impossible to measure; (b) Why do you think so? What is your reservation, if any?

Table 2. Five main questions on UX measures

The notion of “*experiential qualities*” is central for Q7, Q8 and Q9. In the simplest sense, they are referred to as feelings. In the broadest sense, they are related to the concept of emotional responses, as defined in the Components of User Experience (CUE) model [15], which are influenced by instrumental (i.e. usability) and non-instrumental qualities (i.e. aesthetic, symbolic and motivational). While CUE focuses more on evaluation, in the context of the design the notion of experiential qualities is defined as articulations of key qualities in the use of a certain type of digital artefact intended for designers to appropriate in order to develop their own work [8]. Note that in order to enable open discussion no definition was provided to the interviewees unless requests for clarification were solicited. Part C comprises three questions aimed to simulate in-depth discussion (Table 3).

Q10. Which theoretical arguments (e.g. reductionism) are for or against UX measurement?
Q11. Which methodological arguments (e.g. validity) are for or against UX measurement?
Q12. Which practical arguments (e.g. cost) are for or against UX measurement?

Table 3. Questions for in-depth discussions

Participant and Procedure

An invitation to the interview was circulated in the intranet of a university. Eight participants volunteered to take part in it. The other three participants were recruited by the first author via personal invitation. Their participations were also voluntary. They were designated as P1, P2 and so on.

Seven of them were female, five aged between 31 and 40, another five between 41 and 50 and one above 50. All were researchers except P5, who was a practitioner. The job of eight of the participants was predominantly design-oriented, be it practical or theoretical, such as empathic design for house renovation, co-design for persuasive games, and design theories. The other three focused more on UX evaluation of interactive products such as mobile phone. Two of them have worked in UX for less than 1 year, three 1-3 years, five 3-5 years and one for than 5 years. All the interviews were conducted on an individual basis in English, audio-taped and transcribed subsequently.

RESULTS AND DISCUSSIONS

For analysing the data, we developed coding schemes for individual interview questions by applying thematic analysis [3] and the CUE model [15]. Due to limited space, here we do not report results of Q5 (What is a 'measure'?).

Statements on UX Measures

Given the small sample size, no inferential statistics of the ratings are computed. Justifications for the ratings are of higher relevance and the analyses are presented below.

UX measures lead to increase of knowledge (mean = 4.0, range: 2-5). When prompted to specify which kinds of knowledge would be increased, several were mentioned,

- references against which products can be compared;
- the extent to which the development goals is achieved;
- values to be delivered by certain design methods;
- information helpful for future projects;
- experience per se;

Ambivalence was observed, for instance: “There are ways to get knowledge about UX in a more meaningful way rather than using measures, but I still think that they are important.” (P6). Besides, the need for including qualitative data as complementary knowledge was emphasized: “We should have both... qualitative is to know what the reason is for user experience and for the related design issue.” (P8). Furthermore, conditions for benefiting from UX measures were specified: “It requires people using the measure, understand the measure and what it actually means... There might be people who are not trained to use UX measures, no matter how well we define the measures.” (P5). This observation highlights the need for enhancing education and training in UX.

UX measures are insane (mean = 2.0, range: 1-4). A common view was that the insanity lies not in UX measures but rather in what claims to be made about them, especially when people do not understand such measure, intentionally misuse them, are unaware of their inherent limitations (e.g. incompleteness) or over-formalize them. There were also concerns whether UX measures can explain why people experience something or have any use for design, as remarked by P11 (a designer):

“... for the purpose of design, measuring variables up to a very high degree and intricate level of measurement might not be that

purposeful because you have to translate the numbers back to design requirements, and I am not sure whether that works.”

UX measures are a pain (mean = 3.27, range: 1 – 5). Pain inflicted was psychological rather than physical. Reasons for such pain varied with the phase of UX measurement. In the preparation phase, defining valid and meaningful metrics, which entailed deep and wide knowledge of various matters, was cognitively taxing and thus painful. For data collection, participant recruitment and time constraint were a pain for researchers, as illustrated by P4’s remark: “We would not use half-an-hour to measure something but rather get some qualitative data out of participants.” On the other hand, the intrusiveness and lengthiness of the procedure could be pain for users. For data analysis, statistical analysis was deemed challenging by four participants. This again is a clear implication for the training of UX. Interpretation of UX measures was another common concern: it could be an issue of lack of knowledge, confirmation bias, and attempts to draw implications from exact measures for design.

UX measures are important for design (mean = 4.0, range: 2-5). Participants’ stance on this claim was ambivalent. They recognized that UX measures could help identify design constraints and justify design decisions by convincing developers and management, given that numbers could convey a sense of reliability. However, they stipulated the importance of UX measures in design with the need of combining with qualitative data, for instance:

“I mean they are important, but I’d *not* base my design solely on UX measures... there are lot of things that I don’t think that we can measure properly enough yet... it would cause too much work to get really really good measurement that would be our main basis for design... [UX measurement] would only be second; the first being an overall understanding of qualitative views we have found out from users.” (P4)

“If UX measures are clusters that are described through numbers or questionnaires, then they are *not* important for design, whereas if UX measures are, for instance, clusters of qualitative data and users’ accounts, then they are important for design” (P11)

Some participants explicitly expressed their doubt about the role of UX measures in design, for instance:

“I can see relatively little value of applying UX measures, because they don’t really link to the product’s attributes in most cases... they link it at an abstract level... it is hard to trace what the underlying causes for certain response. It is almost impossible if we just use UX measures without combining them with qualitative data” (P1)

Furthermore, one participant pointed out the differences between usability and UX measures:

“... sometimes it is difficult to explain why we design like this even when we provide evidence. From usability point of view we can more easily give this measurement that it is better, but designing for UX is problematic. People with technical backgrounds have problems making the difference between UI and UX. They think they are the same thing.” (P3)

In summary, the interplay between UX measures, which are common evaluation outcomes, and (re)design is ambiguous.

UX measures are important for evaluation (mean = 4.6, range: 2-5). On this claim the participants were somewhat less ambivalent. Supporting arguments such as justifying decisions, validating design goal, and giving reliability (cf. P2’s remark: “If you only use the designer intuition, only use empathic interpretation, it is not very reliable for the rest of the world”) were given. Some participants pointed out the time issue: in which development phase UX measures are taken and how much time the process of measuring is allowed, for instance:

“... in industry-led cases they are more keen on fast phenomenon ... the industrial people want to improve the design but not really want to provide input for the academic world in general” (P4)

There are also reservations about the role of UX measures in evaluation, for instance:

“it’s not been proven yet that [UX measures] can make any difference to outcomes.... I mean, they *could* be; certainly if you include traditional usability measures, then persistent task failure for many designs is going to be something you want to know about. But I don’t think they’re automatically important; they’re all hinges around design objects” (P11)

Measurable and Non-measurable Experiential Qualities

In response to Q7, Q8 and Q9 (Table 2), participants identified different experiential qualities (EQ), which we categorized by the adapted CUE model [15]:

- *Instrumental qualities (NQ)* – “the experienced amount of support the system provides and the ease of use” (e.g. controllability, learnability, effectiveness);
- *Non-instrumental qualities (NIQ)* – “the look and feel of the system”, including aesthetic, symbolic and motivational qualities ([15], p. 916; [9]);
- *Affective responses (AR)* – subjective feelings, motor expressions, and physiological reactions [12] arising from interacting with the system (NB: It broadens the scope implied by original notion of ‘emotional reactions’ to accommodate mildly affective responses with an artefact).
- *Evaluation* (cf. system appraisal) – long-term effects of interacting with the system on user affect, attitude and cognition;

Several interesting observations are noted:

- i) All three EQs considered as non-measurable fall into the category of Evaluation; it seems implying that long-term effects of interaction are considered not amenable to measurement;
- ii) No non-measurable instrumental and non-instrumental qualities were identified by the participants; this is not surprising as instrumental qualities are closely related to traditional software attributes that have explicitly been operationalised and operationising non-instrumental qualities such as aesthetic and symbolic has been endeavoured in recent UX research efforts (e.g. [5]);
- iii) Fun is the EQ that was dually considered as measurable as well as non-measurable. This is somewhat surprising

because game experiences of which fun is an integral part have been one of the hot topics in UX research where different attempts to measure fun have been undertaken (see the review in [1]). This observation underpinned P11's argument for the measurability of fun as it is a well-defined concept. In contrast, P1's counterargument referred to the complexity and multidimensionality of fun; reporting on overall fun after interaction seemed more plausible than on individual sub-constructs;

- iv) Several high-level concepts were mentioned: 'hedonic quality' for measurability and 'long-term experience' and 'deep [sub]-conscious experience'; they do not fit into any of the categories.

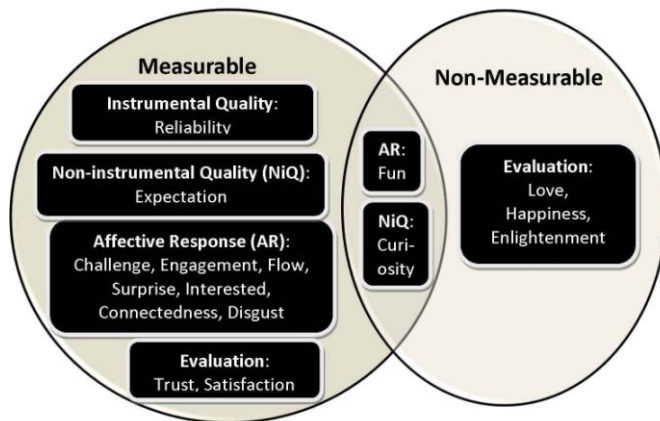


Figure 1: Categorisation of experiential qualities in terms of measurability

Furthermore, the main argument for measurability is that the EQs of interest are well defined and documented in the literature. Two participants, however, could not name any certainly measurable EQ because they considered that qualitative data were better for understanding feelings and that experiential concepts were in general fairly vague. In contrast, the key arguments for non-measurability are the epistemological assumption about the nature of certain experiences and lack of a unified agreement on what UX is. The five participants could not name any certainly non-measurable EQ. They, while assuming that everything can be measured, had the reservations for the validity, impact and completeness of UX measures. Specifically, P9 pointed out the issue of conflating meaningfulness with relevance:

“I think anything can be measured in a meaningful way; it depends who the audience is... the issues with measurement ... are well understood in the psychometric system whether you are really measuring what you think you are measuring. So, and, again you need to distinguish between meaningfulness and relevance... there are things that are irrelevant ... but I don't think it's possible for things in this world to have no meaning... people are natural interpreters.

With regard to the question on how to measure EQ, the participants identified a range of known HCI methods, which can be categorized into three major types: *overt behaviour* (e.g., time-on-task, number of trials to goal); *self-*

reporting (e.g. diary, interview, scale); and *psycho-physiological* (e.g. eye-tracking, heart rate). Obstacles for implementing measurement were also mentioned, including various forms of validity, individual differences, cultural factors, confidence in interpreting non-verbal behaviour, translating abstract concepts into concrete design property, and consistency of observed behaviour

Anecdotal Descriptions on the Interplay between Evaluation and Development

In responding to the interview questions, some participants described intriguing cases that can well illustrate the challenges of enhancing the interplay between UX evaluation and system development. Subsequently we highlight the challenges and related anecdotes, which are grouped as theoretical (Q10), methodological (Q11) and practical issues (Q12).

Theoretical issues

- *Problem of measuring UX in a holistic way and breaking down into components seems not an ideal solution.*

P3: When we go through the issues with uses, we observe the whole expression, their comments on certain issues. If we have a lot of things to study, it is more difficult to run this kind of a holistic study; in a lab test where we only study some specific items. In an evaluation session when we study several issues, we can show users some of them and then the whole one. Holistic approach is the way to go, but measures about some specific details help as well.

P4: I'd say UX is holistic in nature, it is difficult to break it down into very small pieces. From the traditional scientific perspective, the way to measure something, to break it down and separate different factors ... The value of the measurement gets lower if you break it down to small pieces... My colleague studied 3D video. She was able to measure objectively some aspects in lab by breaking things down, but when she went to realistic context for certain kinds of arrangement, the results are really different.... Your experience may change dramatically.

- *Memorized experiences prone to fading and fabrication*

P5: the actual intensity of the moment fades very fast... So it is interesting to see how to recall and how we change the memory of the experience. When we ask people whether they like something or not it depends on the moment you are asking. iPhone, there is so much positive information of that product out there that even if you did not like it, your environment is so positive about it that you are positive as well. It is the same as with reconstructing the memories. ... Most people as well as I myself are sure I have memories where I cannot make a difference between the reconstructed and actual memory.

- *UX measures are highly sensitive to timing and nature of tasks*

P2: When to measure depends the duration and complexity of the task. For a small task, we can let people complete it and take measures at the end. For the longer one may need to be interrupted.... I am thinking a lot how much I am manipulating everything when I am organizing a workshop with some tasks how everything would be different if the tasks would be

different....

P8: Different measures in different phases of the use they complement each other if we need long-term evaluation. Sometimes you can get details out of these supporting design. They are more for prioritising the essential issues.... You don't have exact measures for evaluating emotions at the moment. Very momentary info can be useful, but you also need other measures. Even though you can capture all the momentary emotional measures, you don't know how the user interprets the emotion. The interpretation of the person is very important a negative experience can be interpreted as a positive experience later on.

Methodological Issues

▪ *Different preferences for qualitative and quantitative data by design- and engineering-oriented stakeholders*

P7: ... we are not fond of measures ... we have smart design work, something we have emphasized more on qualitative and inspirational aspect of UX. We have something to do with design perspective; kind of measurement only gives basic constraints and do not give directions. It depends where you apply the methods; how they should be interpreted and position the methods. Measures are good background knowledge but we have more unpredictable, qualitative data.

P8: Qualitative data could cover everything, but then how to convince the engineers, that's why we need numbers. Also for research purpose, it could be interesting to find the relationships between factors. I have to measure somehow to find out which is more influential, hedonic or pragmatic quality, on customer loyalty... quantitative data are more convincing, but developers need qualitative data as well because they want to understand the reason for frustration... the developers like videos because they can describe very lively the situation. They can also believe textual descriptions. ... It is important to measure both immediate experience and memorable experience. Practitioners are very thrilled by the idea that you can do it afterwards because it is so easy. So the companies are very interested in long-term UX or this kind of retrospective evaluation, they don't mind that, because they are convinced that memories are very important because they are telling stories to other customers; they are loyal to the companies based on the memories. Only the reviewers are criticising the validity of retrospective methods. Practitioners are very interested in it and like the idea.

P10: You have to interpret psycho-physiological data and map these data to one of these experiential concepts and it is very hard to know whether you get it right. You can have a high heart rate because you really love it or you hate it. So may be it also depends on how many categories you have; the more categories you have, the more difficult to find a good mapping.

P11: To see the impact of the goal of the system, how people perceive it. I think that's fine. For the purpose of design, quantitative measures do not make sense. It is a wrong method for the purpose of design.

▪ *Resource-demanding evaluation with a large number of heterogeneous users*

P4: Our perspective is very design-oriented. My experience in measuring UX in design process is not so much. It is so easy and fast to make the participants fill out AttrakDiff, it really would not make sense *not* to do it. How we analyse the results and get out of it, that's still to be seen. We don't have so many participants that we could see what the different ways of using those results are. Like a backup, we get a general understanding of the situation to compare for making the second prototype, what things to change. When we have the second prototype and we use the same measurement, we can see where the design is going. As measurement depending so heavily on individual participants, it is difficult to make conclusion about the measurements... it is hard to say why there is a difference in the results because of different social groups.

▪ *Need of sophisticated prototypes for eliciting authentic user experiences*

P7: Difficult, especially housing business ... we cannot build only one prototype and then ask people experience it, get feedback and then do it... we need good examples, media we can use to produce our tools, social media, TV, etc to show what kind of solution we might have.. the storytelling method like movie; I'd like to see sophisticated level like what would be done with professional actors, directors, writers, like real life, feeling like real life with different natural mistakes.

Practical Issues

▪ *Lack of knowledge in exploiting feedback on UX for future system development*

P5: Most people in industry, whether they have backgrounds in economics, engineers or marketing, for them handling qualitative information is very difficult and they even don't know how to use that or they would need that.... We've been criticising the UX evaluation, not about how we measure UX, but how we use the information it in industry. ... But there is so much information that people don't bother to read or follow them. We need to make things simple and easy so that people don't have backgrounds they can understand. In fact, the majority of usability people, at least in Finland, have engineering or computer science background but have little about psychology. There are a lot of things natural for psychologists or sociologists during the study handling control vs. experiment. They don't necessarily come to think of; there are experts in company talking about human beings, but they have certain views. It is challenging. This area of UX has the good side of interdisciplinary as well as the negative ones.

P4: Quite often field experiments lead to straightforward results that can be exploited in their design work right away. One project quite a while ago... We had purely lab experiments. We were doing lab test applying Fitt's law with different input devices, we were creating some constants that could be used for evaluating early stages of design to see if input device Design A is better than Design B. The partners were really excited about the results. They were well done, theoretically and practically validated and applicable... Industrial people were quite lost when we were not there. They needed our guidance. Unfortunately we had no choice. We had good results, but no real exploitation of the results since the customer did not know what to do with the results.

- *Lack of standard UX metrics renders redesign decisions prone to personal biases*

P5: People make decisions based on their personal beliefs. They just pick from the UX measures the ones that support their existing belief, and ignore the other results that don't support. ... They don't even realize it themselves that they are manipulating the results. ... People don't know how to use information on human beings. ... we had noticed that the same icon did not work for various kinds of notification... We got feedback the people were annoyed... there was a very strong personality in the design lead who said that he did not want the design changes because they look ugly... It is problematic that UX have no commonly agreed definition or no commonly agreed metrics. It allows people to use this kind of argumentation that "I believe that it is better UX". You don't need to justify, it can be a personal opinion even though there are tons of user feedback.

- *Packaging UX measures for decision makers and speaking their language*

P4: ... social TV case we did Attrakdiff questionnaire and industry partner was very interested in that. They saw the potential in that when we had enough data, more convincing, more easily convince their superior of the organization to finance their projects, show the need for working on some aspects further; objective foundations.

P5: It is not meaningless to measure moment-to-moment experience, but the question is how you use this information... But how to pack the thing and sell the thing to people making product or legislation decisions. In this area we should talk about how we use the information in this domain for the legislation and guiding the decision makers of different countries... Even when I think about from the industry perspective. Strategy management what they are most interested in is that what are the elements that make users buy next devices from the same company as well and what can reduce the number of helpdesk contacts. The first one is related to the future revenue of the company and the second one is related to the cost saving. It is mostly transfer it to money. It is the language that the management understands.

CONCLUDING REMARKS

UX, as an immature research area, is still haunted by the challenges of defining the scope of UX in general and operationalising experiential qualities in particular. Apart from addressing these basic issues, it is necessary for UX professionals to identify plausible means for compromising the difficulties of evaluating UX in a holistic manner with the limitations of adopting the reductionist approaches. Deeper understandings about the relationship between experience and memory and about the temporality of UX are also required. While the utility and necessity of employing both quantitative and qualitative methods is commonly recognized, the concomitant issue of providing appropriate education and training in UX needs to be explored. Specifically, UX researchers and practitioners should be equipped with knowledge and skills to know why certain UX measures are taken and how to use and interpret them in order to inform design and development decisions.

Insights into the issues of UX measures have been gained from the interviews. The study has raised more questions

than it can answer. As the number of participants was relatively low with most of them originating from one country, namely, Finland, the views expressed might not be representative. Given this drawback, we have been motivated to expand the investigation on UX measurement with a larger scale survey of which results are documented elsewhere (under review). With a better understanding of the issues about UX measures, especially how they can be translated into new design requirements, insights into the interplay between UX evaluation and design can be gained.

ACKNOWLEDGEMENT

Many thanks should be given to Dr. Virpi Roto, Aalto University, Finland for her generous support in arranging the interviews when Effie Law had her short-term scientific mission in Helsinki funded by COST IC0904 TwinTide.

REFERENCES

1. Bargas-Avila, J.A., & Hornbæk, K. (2011). Old wine in new bottles or novel challenges? A critical analysis of empirical studies of user experience. In *Proc. CHI'11*
2. Bartholomew, D. J. (2006) (Ed). *Measurement (Sage Benchmarks in Social Research Methods)*. Volume 1. Sage.
3. Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Sage.
4. Hand, D.J. (2004). *Measurement theory and practice*. Wiley-Blackwell.
5. Hassenzahl, M., & Monk, A. (2010). The influence of perceived usability from beauty. *Human-Computer Interaction*, 25(3), 235-260.
6. Law, E.L-C. (2011). The measurability and predictability of user experience. In *Proc. of the 3rd ACM SIGCHI Symposium on Engineering Interactive Computing System (EICS 2011)*, Pisa, Italy, June 2011.
7. Law, E. L-C, Roto, V., Hassenzahl, M., Vermeeren, A., & Kort, J. (2009). Understanding, scoping and defining user experience: a survey approach. In *Proc. CHI '09*, 719-728. ACM, 2009.
8. Löwgren, J. (2007). Fluency as an experiential quality in augmented spaces. *International Journal of Design*, 1, 1-10.
9. Mahlke, S., Lemke, I., & Thüring, M.(2007). The diversity of non-instrumental qualities in human-technology interaction. *MMI-Interaktiv*, Nr. 13, Aug 2007, ISSN 1439-7854.
10. Obrist, M., Law, E.L-C., Väänänen-Vainio-Mattila, K., Roto, V., Vermeeren, A., & Kuutti, K. (2011). UX research- which theoretical roots do we build on – if any. In *Extended Abstract CHI'11*.
11. Schaik, van P., Hassenzahl, M., & Ling, J. (2012). User experience from an inference perspective. *ACM Transaction on Human-Computer Interaction*.
12. Scherer, K. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44(4), 695-729.
13. Vermeeren, A. P.O.S. Law, E. L-C., Roto, V., Obrist, M., Hoonhout, J., Väänänen-Vainio-Mattila, K. (2010). User experience evaluation methods: current state and development needs. In *Proc NordiCHI 2010* (pp. 521-530).
14. Thomson, W. (1891). *Popular Lectures and Addresses, Vol. 1*. (p.80). London: MacMillan.
15. Thüring, M., & Mahlke, S. (2007). Usability, aesthetics and emotions in human-technology interaction. *International Journal of Psychology*, 42(4), 253-264.