

The Relevance of UX Models and Measures

Asbjørn Følstad

SINTEF

Po. Box 124, Blindern

0314 Oslo

+47 22067515

asf@sintef.no

ABSTRACT

Two approaches to research on UX models and measures are discussed on basis of experiences from the field of usability research and an ongoing case of user involvement in software development (SD) by way of social media. It is suggested that simple measures and ad-hoc models, rather than complex models and measures, may be beneficial to the relevance of UX research for SD practice.

Categories and Subject Descriptors

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Keywords

User experience, software development, model, measure

INTRODUCTION

A key objective of research on user experience (UX) models and measures should be to influence the practice of software development (SD). However, to reach this objective, UX research need to address research relevance as well as rigor.

A useful discussion of rigor and relevance in research was made by Lee [1] within the context of information systems (IS). He argued that knowledge produced by IS research emulating the rigorous natural sciences does not by necessity imply relevance. Rather, for the research to be relevant it needs to generate “*knowledge about how to intervene in the world and change it in order to satisfy real world needs*” (p. 29).

In this paper, I will discuss two distinct approaches to research on UX models and measures, and their implications for research rigor and relevance. The first approach, *complex models and measures*, is what I see as a likely trend given that relevance is not prioritized in UX research. The second approach, *simple measures and ad-hoc models*, is suggested as an alternative.

The first approach is paralleled in previous work of the related field of usability research. In order to learn from the experiences made in this more mature field, my argument will be supported by reference to research on usability models and methods.

The second approach represent an intermediary position to what has been referred to as *phenomenological/pragmatist* vs. *inspired by experimental psychology* [2]. This approach is exemplified by a case from an ongoing research project.

The intended contribution of the paper is to serve as a starting point for discussions on the relevance of UX research.

APPROACH 1: COMPLEX MODELS AND MEASURES

A recurring theme of UX research discussions is the components of UX [2]. Which model components are needed for measurement and systematic UX improvement? Given the comprehensiveness of the UX concept [3] the set of model components is likely to be voluminous, indicating that future UX models well may be complex. Suggested components include, for instance, motivation, trust, aversion, hedonics, and fun [2].

This trend towards complexity mirrors parts of the usability research during the 80'es and 90'es. A large number of usability components were suggested, including Nielsen's efficiency, satisfaction, learnability, memorability, and error rate [4]. A similar complexity is found in measures such as SUMI [5]. ISO 9241-11 [6] decomposes usability in effectiveness, efficiency and satisfaction.

The complexity of usability models and measures is beneficial for the rigor of usability research. However, complexity may be detrimental to the relevance for usability models in SD. For SD practitioners, complex generic models may be impractical. The generic model may include several components that seem irrelevant to the development project at hand. Also, seemingly important aspects may not be handled by the model. One may speculate that the prevalence of homegrown usability measures at the expense of standardized measures [7] may be a consequence of the complexity of the underlying usability models and their mismatch with the SD practitioner's understanding of which aspects of usability that are relevant.

Sauro and Kindlund [8], a key critic of the relevance of usability research, argued that to increase the practical impact of usability data “*usability metrics need to be easier to use*” (p. 401). To this end they suggested a single summative usability metric.

The position of Sauro and Kindlund is controversial [9] though it recently has been underpinned by substantial empirical evidence from industry projects [10]. Controversial or not, an important lesson may be learnt here for research on UX models and metrics: Complex models, underpinning complex measures, are likely to be valuable to the advancement of UX theory. However, in order to advance the interplay between UX and SD, simplified models and measures may be required.

APPROACH 2: SIMPLE MEASURES AND AD-HOC MODELS

An alternative approach to UX models and measures may be pursued through simple measures and ad-hoc models.

A simple UX measure is a single rating scale, common in social software. Consider for instance the book ratings of Amazon (1-5 stars) or video ratings at YouTube (thumbs up / down). Such scales are typically shunned for the scientific measurement of experiences or attitudes, due to reliability issues. Similarly, Sauro and Lewis argue for composite rather than single item measures [9]. Even so, such measures seem to serve their purpose as practical social navigation tools. Also, research within the field of marketing indicates that single-item measures may hold similar predictive validity as multi-item measures for concrete constructs such as *ad liking* and *brand attitude* [11].

By ad-hoc UX models I mean models developed in response to a given concept, prototype or running system. Instead of utilizing a general UX model as basis for a generic UX measure, relevant UX components may be established on basis of users' responses.

As an exemplification of simple measures and ad-hoc models, I will briefly present an ongoing development case addressing new functionality for mobile phone e-mail clients. Design typically involves the "*simultaneous investigation of multiple alternatives by the same designer or team*" [12, p. 1243]. In the present case 22 ideas were generated and then refined as six early concepts; all across a working period of 60 hours. Following this, the concepts were made available for user feedback.

User feedback was collected through a social software application for sharing audio-visual content, modified for the purposes of design feedback. The participants were 212 regular users of e-mail-clients on mobile phones (use several times a week or more). They were presented for six concepts in sequence. For each concept they were asked to make a rating (1-5 stars) and one or more comments. The comments were made in response to open questions: *How would you use the suggested function? How may the suggested function be improved?*

As part of a research design not to be detailed here, half of the participants were allowed to see the other participants'

comments prior to making their own and half of them were not. All participants were allowed to comment on each other's comments. None were allowed to see the others' ratings prior to making their own.

The user feedback provided the following key information:

- **Differentiation between the concepts:** Three of the concepts were rated low, three rated high. The ratings corresponded closely to the frequency of positive vs. negative comments for each concept.
- **Establishment of concept specific issues that may serve as basis for an ad-hoc UX model:** For one of the concepts, detailing functionality for 'Postponed sending of messages', relevant issues were mainly targeting utility. For another, 'Reading aid for long e-mails on small screens', relevant issues were lack of comfort and utility. For a third, 'Context-dependent e-mail receipt', relevant issues were privacy and utility.

The case thus illustrates that simple measures and ad-hoc models may serve as basis for choosing which design alternative to pursue, and enable the establishment of ad-hoc models that may be used to control the development process. In the present case, a UX model for the first concept would need to include only utility in addition to usability, whereas the UX models for the second and third concepts would also need to include comfort and privacy/trust respectively.

CONCLUSION

By this paper, I hope to contribute to a discussion on how research on UX models and measurements should be approached in order to obtain relevance. I hold that complex models and methods will indeed be beneficial for the establishment of UX theory. However, research relevance may require a different approach. Possibly, such an approach may be *simple measures and ad-hoc models*.

ACKNOWLEDGMENTS

This paper was written as part of the research project RECORD (www.recordproject.org), supported by the VERDIKT programme of the Norwegian Research Council.

REFERENCES

- [1] Lee, A. S. 1999. Rigor and relevance in MIS research: Beyond the approach of positivism alone. *MIS Q.* 23, 1, 29-33.
- [2] Law, E. L.-C., Vermeeren, A., Hassenzahl, M. and Blythe, M. 2007. Towards a UX Manifesto. Proceedings of Towards a UX Manifesto. COST294-MAUSE affiliated workshop.
- [3] ISO. 2010. *ISO 9241-210: Human-centred design of interactive systems*. Geneva: International standards organization.

- [4] Nielsen, J. 1993. *Usability engineering*. Academic Press, San Diego, CA.
- [5] Kirakowski, J. *SUMI Background Reading: The Use of Questionnaire Methods for Usability Assessment*. Available at: <http://sumi.ucc.ie/sumipapp.html> (downloaded July 6, 2010)
- [6] ISO. 1997. *ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs). Part 11 – guidelines for specifying and measuring usability*. Geneva: International standards organization.
- [7] Hornbæk, K. 2006. Current Practice in Measuring Usability: Challenges to Usability Studies and Research. *International Journal of Human-Computer Studies*, 64, 79-102.
- [8] Sauro, J., Kindlund, E. 2005. A method to standardize usability metrics into a single score. *Proceedings of CHI '05*. ACM, New York, NY, 401-409.
- [9] Hornbæk, K., & Law, E. (2007). Meta-analysis of correlations among usability measures. *Proceedings of CHI'07*. ACM, New York, NY, 617-626.
- [10] Sauro, J., Lewis, J. R. 2009. Correlations among Prototypical Usability Metrics: Evidence for the Construct of Usability. *Proceedings of CHI'09*. ACM, New York, NY, 1609-1618
- [11] Bergkvist, L., and Rossiter, J.R. 2007. The predictive validity of multi-item versus single-item measurements of the same constructs. *Journal of marketing research*, 44, 175-184.
- [12] Tohid, M., Buxton, W., Baecker, R., and Sellen, A. 2006. Getting the right design and the design right. *Proceedings of CHI'06*. ACM, New York, NY, 1243-1252.