User Experience Over Time

Evangelos Karapanos
User-Centered Engineering
Department of Industrial Design
Eindhoven University of Technology
P.O. Box 513, 5600 MB, Eindhoven
The Netherlands
E.Karapanos@tue.nl

Marc Hassenzahl
Economic Psychology and Human-Computer Interaction
University of Koblenz-Landau
Fortstraße 7, 76829
Landau, Germany
hassenzahl@uni-landau.de

Jean-Bernard Martens
User-Centered Engineering
Department of Industrial Design
Eindhoven University of Technology
P.O. Box 513, 5600 MB, Eindhoven
The Netherlands
J.B.O.S.Martens@tue.nl

Copyright is held by the author/owner(s).
CHI 2008, April 5 – April 10, 2008, Florence, Italy
ACM 978-1-60558-012-8/08/04.

Abstract
The way we experience and evaluate interactive products develops over time. An exploratory study aimed at understanding how users form evaluative judgments during the first experiences with a product as well as after four weeks of use. *Goodness*, an evaluative judgment related to the overall satisfaction with the product, was largely formed on the basis of pragmatic aspects (i.e. utility and usability) during the first experiences; after four weeks of use identification (i.e. what the products expresses about its owner) became a dominant aspect of how good a product is. Surprisingly, *beauty* judgments were largely affected by stimulation (e.g. novelty) during the first experiences. Over time stimulation lost its power to make the product beautiful in the users' eyes.

ACM Classification Keywords
H5.2. User Interfaces: Evaluation/methodology.

Keywords:
User Experience, longitudinal, product appraisal

Introduction
With the advent of user experience, a number of qualities that go beyond the instrumental aspects of product use have been introduced. Nowadays, it seems widely accepted that a product needs not only be useful and

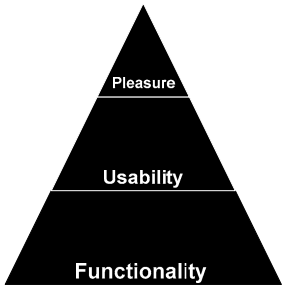

Figure 1 Jordan's [5] fixed hierarchy of consumer needs

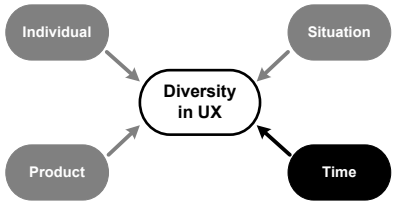

Figure 2 Four sources of diversity in User eXperience

usable (i.e., pragmatic), but also stimulating, beautiful and pleasurable (i.e., hedonic).

A question closely related to the introduction of hedonic quality to complement pragmatic quality is their relative importance [2]. Jordan suggested a fixed hierarchical structure of qualities that contribute to positive experience [5] (figure 1). According to this, a product has to provide useful and usable functionality before hedonic aspects, such as beauty and stimulation, can take effect.

In contrast to Jordan, we assume the importance of these different qualities to vary with several contextual factors (figure 2).

First, individual differences (e.g. human values [13]) will moderate the importance individuals attach to different qualities of an interactive product [7, 1]; while some might prefer playful and stimulating products, others might value simplicity and austerity in products. Second, the type of the product will matter [6]; while a playful interaction might be crucial for the success of a computer game, the same quality might be perceived as inadequate for professional software. Third, even for the same product, the way we use it will differ across situations and this will impact the importance we attach to different qualities [4]; the same mobile phone could be used for exploring the available ring tones or to make an emergency call. All these factors, the individual, the product and the situation will modify the importance of the qualities for a satisfying experience with an interactive product. In this view, Jordan's [5] hierarchy of consumer needs could be seen as a *particular, context-dependent* prioritization of needs [2].

An aspect mostly overlooked so far is the systematic change of experience over time. As we use a product, our perception of the qualities of the product will change [14]. For example, we get used to it, which eventually changes our perception of its usability; at the same time it excites us much less than in our first moments with it. Even more interestingly, at different phases of use we will evidently attach different weights to different qualities. In our first interactions with a product we may focus on its usability and the stimulation that it provides to us. After we use it for some time, we might become less concerned about its usability, and other aspects of the product such as novel functionality or communication of a favorable identity to others become more important.

Beauty, goodness and time

Hassenzahl [3] suggested two distinct overall evaluative judgments of the quality of interactive products: *beauty* and *goodness*. He found goodness to be affected primarily by *pragmatic* aspects (i.e. utility and usability). On the contrary he found beauty to be a rather social aspect, largely affected by a construct he termed *identification* (i.e. the product's ability to address the need of expressing one's self through objects one owns). Surprisingly, *stimulation* (i.e. the product's ability to address the human need of stimulation, novelty and challenge) did not seem to have a significant effect on either overall evaluative judgment [3, 10]. While goodness was affected by experience, with utility gaining importance after the first interaction, beauty was not affected by this initial usage experience.

In Hassenzahl's study however, users' experience with the products was limited to short interaction episodes, where users were asked to carry out a number of pre-

Table 1. Bipolar attributes measuring pragmatics, stimulation and identification.

Pragmatics
Technical—human
Complicated—simple
Impractical—practical
Cumbersome—direct
Unpredictable—predictable
Confusing—clear
Unruly—manageable
Stimulation
Typical—original
Standard—creative
Cautious—courageous
Conservative—innovative
Lame - exciting
Easy—challenging
Commonplace—new
Identification
Isolating—integrating
Amateurish—professional
Gaudy—classy
Cheap—valuable
Noninclusive—inclusive
Takes me distant from people—brings me closer to people
Unpresentable—presentable

defined tasks. Our interest here was to explore how the user experience and the subsequent evaluative judgments develop over a longer period of time, and in less controlled circumstances.

The present study

A total of ten individuals (four female) participated in a four weeks study of an Interactive TV set-top box (STB). They all responded to an invitation that was placed on the website of a multinational consumer electronics company. Their age varied from 22 to 35 years (mean 26y). Their likelihood to recommend the brand ranged from 3 to 9 on a 10-point scale (mean 7.8, std. 2.3). Participants were classified to respective market segments based on demographic information. A bias towards innovator consumers was observed, as it was expected. The study focused on a particular part of the set-top box, uWand. uWand is a novel pointing device for interactive TV contexts. It uses led technology to identify where the user points at within the content that appears on the TV.

During the four week testing period participants were asked to rate uWand at two different times, during the first week of use as well as at the end of the 4th week. The AttracDiff 2 questionnaire [3] was employed for the assessment of three distinct aspects of the quality of interactive products: *pragmatics* (e.g. simple, practical, clear), *stimulation* (e.g. innovative, exciting, new) and *identification* (e.g. inclusive, classy, presentable). Each quality aspect is measured with seven bi-polar attributes, employed in semantic differential scales. Beauty and goodness were measured with single items (taken from AttracDiffs appeal construct). Both are evaluative judgments with Goodness focusing on the complete product, while Beauty is rather restricted to visual features. Note that for evaluative, high level summary judgments single item measurements are appropriate and commonly used (e.g., to measure subjective well-being).

Results

Since we were interested in a detailed picture of the relationship between product attributes, we decided to analyze every attribute (in Table 1) separately. The distance $D_{ij} = 1 - |R_{ij}|$ between the individual attributes *i* and *j* was derived from the correlation R_{ij} between the ratings on the two attributes. The obtained distances were subsequently visualized in three dimensions (Stress value S=0.19 1st week; S=0.15 4th week) using the Multidimensional Scaling tool XGms [11]. Hierarchical clustering (with minimum variance) was applied to the predicted distances in the three-dimensional space. Figure 3 illustrates a 2D projection of the 3-dimensional visualization of the distances between the quality attributes. The obtained clusters are denoted by the connecting lines. The left figure reflects the users' ratings during the first week of use while the right figure reflects the users' ratings after four weeks of use. All in all, clusters derived from the measurement in the first week reflect the assumed underlying constructs, with a close knit groups of mainly pragmatic and hedonic stimulation attributes and a looser rest of hedonic identification attributes After four weeks, the perceptions seem much more differentiated, and relationships among attributes lost some of their strength.

During the first week, *Beauty* judgments relate mostly to attributes reflecting the quality of stimulation (i.e. original, creative, new, innovative) and to one attribute reflecting identification (i.e. classy). This relation between stimulation and beauty seemed to disappear af-

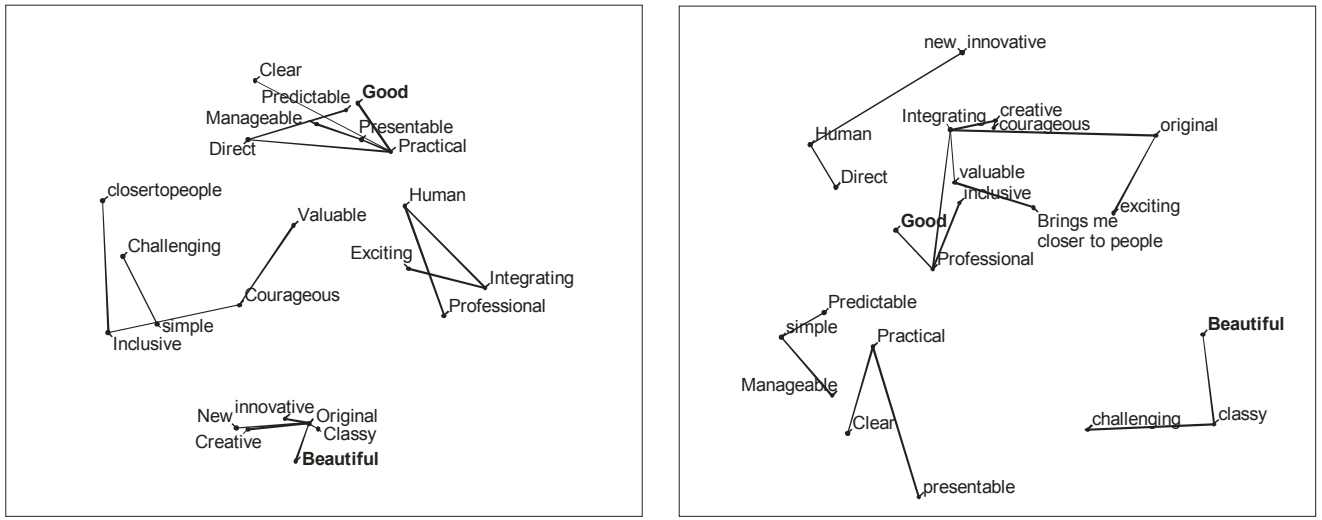

Figure 3 2D view on the 3-dimensional visualization of distances between quality attributes, beauty and goodness. Users' perceptions during the 1st week of use (left) and after 4 weeks of use (right).

ter four weeks of use; beauty now seems to be a disjoint judgment only related to one identification (i.e. classy) and one stimulation (i.e. challenging) attribute. While the relation between "classy" and beauty remained over the period of four weeks, stimulation seemed to influence beauty judgments mostly during the first experiences. Surprisingly, stimulation seemed to be a more important determinant of beauty than identification in the first experiences.

A similar pattern can be observed for judgments of *Goodness*. During the first week judgments of goodness relate mostly to pragmatic quality attributes (i.e. practical, direct, manageable, predictable, clear) and to one attribute reflecting identification (i.e. presentable). After four weeks of use goodness appears to be related mostly to identification (i.e. professional, inclusive, valuable, integrating, brings me closer to people) while a weaker relation can be observed with attributes relating to stimulation (i.e. creative, courageous, original, creative). Pragmatic aspects seem to be relevant for goodness judgments only for the first experiences with a product. Over time, identification (i.e. what the product expresses about its owner) becomes a more prominent aspect of the goodness of the product.

Discussion

Two questions were of interest in the current study: what makes a product good or beautiful, and how does this develop over time.

As far as goodness judgments are concerned, we partially replicated Hassenzahl's [3] results. During the first experiences pragmatic aspects were the most prominent determinants of goodness. Here, users are still exploring the functionality of the product, trying out new things and experiencing usability problems. As people get used to using the product they learn to handle usability problems; at the same time they restrain themselves only to part of the product's functionality that is most attractive to them. The value of the product is now derived on a different basis, being ownership-based rather than usage-based. Social aspects (i.e. identification) became more prominent here.

For beauty judgments however, the results seemed more divergent. While Hassenzahl [3] found identification to be the major determinant of beauty judgments, we found stimulation to be even more prominent than identification in the first experiences. At least, beauty is related to hedonic rather than pragmatic aspects.

This however might be affected by the product sample. While the current study employed a novel consumer electronics product, Hassenzahl's study focused on different variations of mp3 player skins; one could doubt whether variations in aesthetics and the usability of the same product (i.e. virtual mp3 player) can truly affect its perceived stimulation (i.e. innovation, novelty). Mahlke [10] found a quality called expressive aesthetics [8] (arguably comparable to the quality here called stimulation) to have an impact on goodness but not on beauty judgments, during the first experiences with tangible mp3 players. Further, the nature of the experience differed significantly in this study. Both Hassenzahl [3] and Mahlke [10] asked participants to carry out a number of tasks in a laboratory context; in this study participants had the opportunity to use the product at their homes over an extended period and the first evaluation took place during the first days of use.

After four weeks of use, stimulation seemed to loose its dominance on beauty judgments. Eventually, users were not any more surprised by the product's stimulating character and the product's novelty lost its power to make the product more beautiful in the users' eyes.

The question remains: is beauty solely an appearance-based judgment, largely affected by the products' ability to communicate a favorable image about its owner, as Hassenzahl [3] suggested, or does it encompass other aspects of the products, e.g. its perceived novelty or a stimulating interaction [12]?

Limitations and future work

A number of limitations of the study have to be noted however before proceeding to general conclusions. First, an inherent limitation of the study was the evaluation of only one product and only by a small number (10) of subjects. Note however the exploratory character of the study and the lack of "longitudinal" studies in the field of user experience. We hope that this can be the start of a body of studies taking time systematically into account. Second, the two measurements (1st and 4th week of use) assessed users' perceptions at each time (current state) rather than directly assessing how their perceptions changed over time. One could be concerned about the sensitivity as well as the reliability of such absolute measures where judgments do not take place in contrast to some specified reference point.

Despite the exploratory character of the study, it seems that we came across some interesting results. The proposition that novelty and stimulation impact beauty judgments resonates with Berlyne's work on stimulation and surprise as arousal-heightening attributes and their impact on the appraisal of art [9]. Furthermore, time seems to have an impact on the importance we attach to different qualities of the experience with interactive products. For example, despite the crucial importance of usability in a product's initial acceptance, aspects of product ownership (and not use) are even more crucial for a user to resonate with a product and value it in the long term. A future study that is currently being planned will try to get further insight into the reasons of change of our evaluative judgments over time. It will further attempt to correlate subjective evaluations with behavioral measures as well as with emotional reactions during reported experience incidents.

Acknowledgements

We thank Nico van den Gaarden, Serge Stevenart and Ronald Geerlings from Philips Consumer Electronics for facilitating and commenting on this study. This work is being carried out as part of the "Soft Reliability" project, sponsored by the Dutch Ministry of Economic Affairs under the IOP-IPCR program.

References

[1] Desmet, P. M. A., Hekkert, P. and Hillen, M. G., *Values and Emotions*, *in proceedings of Fifth European academy of design conference*, Barcelona, Spain, (in press).

[2] Hassenzahl, M., *Hedonic, emotional, and experiential perspectives on product quality*, in C. Ghaoui, ed., *Encyclopedia of Human Computer Interaction*, Idea Group, 2006, pp. 266-272.

[3] Hassenzahl, M., *The interplay of beauty, goodness, and usability in interactive products*, Human-Computer Interaction, 19 (2004), pp. 319-349.

[4] Hassenzahl, M. and Ullrich, D., *To do or not to do: Differences in user experience and retrospective judgements depending on the presence or absence of instrumental goals*, Interacting with Computers, 19 (2007), pp. 429-437.

[5] Jordan, P. W., *Designing Pleasurable Products: An Introduction to New Human Factors*, Taylor & Francis, London, 2000.

[6] Jordan, P. W. and Persson, S., *Exploring users' product constructs: how people think about different types of product*, International Journal of CoCreation in Design and the Arts, 3 (2007), pp. 97-106.

[7] Karapanos, E. and Martens, J.-B., *Characterizing the diversity in users' perceptions*, in C. Baranauskas, ed., *Human-Computer Interaction - INTERACT 2007*, Springer, 2007, pp. 515-518.

[8] Lavie, T. and Tractinsky, N., *Assessing dimensions of perceived visual aesthetics of web sites*, International Journal of Human-Computer Studies, 60 (2004), pp. 269-298.

[9] Machotka, P., *Daniel Berlyne's contributions to empirical aesthetics*, Motivation and Emotion, 4 (1980).

[10] Mahlke, S., *Aesthetic and Symbolic Qualities as Antecedents of Overall Judgements of Interactive Products*, *People and Computers XX - Engage, Proceedings of HCI '06*, Springer, London, 2006.

[11] Martens, J.-B., *Image technology design: A perceptual approach*, Kluwer Academic Publisher, Boston, 2003.

[12] Overbeeke, C. J., Djajadiningrat, J. P., Hummels, C. C. M. and Wensveen, S. A. G., *Beauty in Usability: Forget about Ease of Use!*, Taylor & Francis, London, 2002.

[13] Schwartz, S. H., *Universals in the content and structure of values: theoritical advances and empirical tests in 20 countries.*, in M. P. Zanna, ed., *Advances in Experimental Psychology*, Academic Press, San Diego, CA, 1992, pp. 1-65.

[14] von Wilamowitz Moellendorff, M., Hassenzahl, M. and Platz, A., *Dynamics of user experience: How the perceived quality of mobile phones changes over time*, *User Experience - Towards a unified view, Workshop at the 4th Nordic Conference on Human-Computer Interaction*, 2006.