ISBN: 978-972-8924-39-3 © 2007 IADIS

ON THE DISCREPANCIES BETWEEN DESIGNERS' AND USERS' PERCEPTIONS AS ANTECEDENTS OF FAILURES IN MOTIVATING USE

Evangelos Karapanos
Eindhoven University of Technology
P.O. Box 513, 5600 MB, Eindhoven, The Netherlands
E.Karapanos@tue.nl

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

ABSTRACT

Motivational aspects of product use have recently attracted an increasing interest in the field of Human-Computer Interaction. This paper presents a study that explored the discrepancies between designers' and users' perceptions as antecedents of failures in motivating use. The study corroborates prior findings suggesting that designers often fall short in accounting for users' views on the product and foreseeing their preferences. Designers seemed to ascribe more weight to efficiency and effectiveness while users seemed to have a more balanced view with a tendency to ascribe more weight to the emotional aspects of the products. Thus, the need for a holistic understanding of the product qualities that motivate use becomes apparent. Finally, a methodological approach for the study of motivational aspects of product use is proposed, based on a relatively unexplored structured interview technique, the Repertory Grid Technique.

KEYWORDS

Designer, user perceptions, emotion, Repertory Grid.

1. INTRODUCTION

Empirical studies on usability have been traditionally focusing on the functional (i.e. "how to") aspects of product use. Motivational (i.e. "why to") aspects though, have been recently receiving increasing attention. Cockton (2004) argues for a value-driven approach in product quality assessment. Such value may be derived from products' instrumental qualities (e.g. perceived ease of use; Keinonen, 1997) but also from the aesthetic (Lavie and Tractinsky, 2004) and symbolic (Hassenzahl, 2004) qualities of the product (cf. Mahlke, 2006; for a recent review). Overbeeke and colleagues (2002) go even a step further to correlate motivation with the aesthetic quality of the interaction per se. They argue that "there is more to usability than ease of use. A user may choose to work with a product despite it being difficult to use, because it is challenging, seductive, playful, surprising, memorable or rewarding, resulting in enjoyment of the experience". Albeit from different perspectives, all these approaches object to the narrow focus of product quality assessment on efficiency and ease-of-use.

From an industrial point of view, such an interest on the motivational aspects of product use is clearly supported. A recent analysis of the reasons of product returns in a multinational Consumer Electronics company demonstrated that almost 50% of product returns were not due to violations of product specifications (Ouden et al., 2006). Reasons why products did not meet customers' expectations and therefore were returned, were mostly traced back to decisions made in the concept design phase. The question raised then is: to what extend are motivational failures caused by a discrepancy between designers' and users' perceptions of the product.

This paper describes a first study in a research line that aims at understanding the reasons for motivation failures in product use. In this study designers' views on users' perceptions were compared to actual users' perceptions, using a mixture of qualitative and quantitative techniques. First, designers' and users' idiosyncratic views were elicited in a structured interview approach using the *Repertory Grid Technique (RGT)* (Fransella et al., 2003). The RGT originates from Kelly's Personal Construct Theory (Kelly, 1955) which suggests that people form idiosyncratic interpretations of reality based on a number of dichotomous variables (e.g. good-bad), referred to as *personal constructs* or *attributes*. In the RGT one tries to elicit the user perceived attributes by asking for similarities and dissimilarities within a set of products.

Perceptual distances between designers and users were then derived from their dissimilarity ratings by means of *Multi-dimensional Scaling (MDS)* (Martens, 2003, Green et al., 1989). In MDS, an n-dimensional object configuration is built where the distance between any two of the plotted objects relates to their dissimilarity as rated by the subjects.

Finally, designers' and users' perceptions, as elicited from RGT, were employed in a comparison between two products in order to obtain insight into the ways in which the two groups differ.

2. METHOD

Eleven "designers" and eleven potential end-users participated in the study. Designers were employees of the R&D department of an international company developing document systems. They were all involved in the conception and realization of *TouchToPrint*, which is a new way of personalizing the use of a shared printer by means of fingerprint recognition. They ranged from usability experts and market experts to software engineers and documentation experts. We refer to them as *'designers'* since they were all stakeholders in the concept design phase. Users were researchers from our department who had no prior knowledge of the product under development.

The TouchToPrint concept and five alternative ways of interacting with a shared printer were presented to the participants in the form of posters. Each poster described a usage scenario of the relevant concept. First, the six products were combined in three triads. The order in which the three products were presented was counterbalanced between participants. For every triad, participants were asked to "think of a property or quality that makes two of the products alike and discriminates them from the third". The same procedure was repeated until a point where no new attributes were arising. Though, a limit of six attributes per triad was set (in total up to 18 attributes for the three triads). While users were instructed to express their own perceptions, designers were instructed to express their views on the perceptions of this particular user group. This process resulted in a list of *attributes*, product qualities that users perceive and use while forming overall evaluations of a product. After attribute elicitation, subjects were asked to remove duplicate attributes and rank the remaining attributes according to their importance. Finally, subjects were asked to rate all products on their personal attributes, as well as on *preference* and *dissimilarity*.

In contrast to the traditional Repertory Grid approach, we employed paired comparisons instead of semantic differentials, as this was a priori expected to deliver more stable results (Martens, 2003). While in the semantic differential technique only one product is being rated and thus being compared to an implicit reference point, in paired comparison two products are being compared on a specific attribute. Out of the six products one can form up to $n(n-1)/2 = 15$ pairs. To reduce the number of pairs we employed a balanced incomplete design (Furlan and Corradetti, 2006, Sandt Van de, 1970) with 9 total pairs and every of the six products participating in $\lambda=3$ pairs ($n\lambda/2$).

3. RESULTS

A total of 81 attributes for designers and 95 for users were obtained in the study (6 to 11 per participant). To enable interpersonal analysis, all attributes need to be classified in shared attribute categories. Two rounds of analysis were performed: an exploratory semantic classification of attributes, followed by a confirmatory analysis of the classification. For the first round, sixteen semantically unique attribute categories were first formed out of the data. To minimize the researcher's bias, the naming of the attribute categories was restricted to choosing one of the attribute names that reflect this semantic value. Subsequently, the first

author and two additional experimenters independently classified every attribute to one of the sixteen categories (Table 1). Interrater agreement (Fleiss et al., 2003) of the initial classification was satisfactory (k=0.72). All sixteen categories were then classified to three overall classes: Effectiveness, Efficiency, and Emotional Appreciation (interrater agreement, k= 0.80).

During the confirmatory analysis of the classification, statistical consistency across attributes within the same category was being sought. Attribute scores were submitted to a cluster algorithm where Euclidean distances between attributes were calculated and visualized in two or three dimensions. Outlier attributes were identified. The prospect of transferring the attribute to one if its statistically-neighbor categories was explored. If there was no argument for a transfer to another category, the attribute was deleted.

Table 1. Attribute categories and overall classes

Effectiveness (32%)	Efficiency (51%)	Emotional Appreciation (17%)
1. Secure	5. Fast	12. Personal
2. Reliable	6. Error-prone	13. Modern
3. Powerful	7. Brainless	14. Privacy
4. Multi-user	8. Cognitive load	15. Hygienic
	9. Effort-searching-documents	16. Environmental friendly
	10. Fast (total time)	
	11. Easy-to-learn	

The Repertory Grid process resulted in three kinds of data: *dissimilarity*, *preference* and *attribute* data. Dissimilarity between two products can be assumed to be an overall function of the degree to which the two products differ in a set of perceived attributes, and the weighting of the importance of every attribute. Every individual provided a dissimilarity score for every of the k=9 pairs of products.

To explore the discrepancies between designers' and users' perceptions we calculated distances D_{ij} between participants i and j based on dot-product correlations R_{ij} of the k dissimilarity scores (1). Derived distances were then visualized in two dimensions (Figure 1) using the MDS tool XGms (Martens, 2002). The two dimensional visualization was judged as adequate (stress value S=0.18) (Clarke, 1993). Hierarchical clustering (with minimum variance) revealed two main clusters, one consisting entirely of users and the other consisting mostly of designers.

$$D_{ij} = 1 - R_{ij}^{\ 2} \qquad R_{ij} = \frac{\sum_k D_i(k) \cdot D_j(k)}{\sqrt{\sum_k D_i^{\ 2}(k) \cdot D_j^{\ 2}(k)}} \qquad (\mathbf{1})$$

$D_i(k)$: dissimilarity score given by subject i on k-*th* product comparison
(k sums over all (9) comparisons)

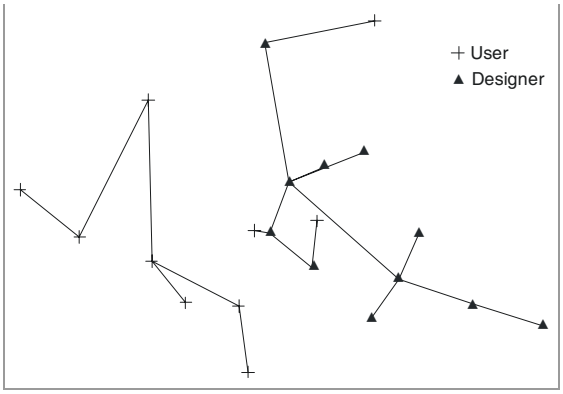

Figure 1. (Left) Two-dimensional visualization of dissimilarities between designers' and users' perceptions and hierarchical clustering (minimum variance)

To acquire richer insight into the ways in which designers and users differ we focused on a comparison between two of the six products: *TouchToPrint* and *Badge*. These two products differed only in the mechanism that was employed for the user identification process: a touch sensor and a sensor for an individual's badge. First, a comparison of preference between the two products was performed. Four

designers and eight users preferred TouchToPrint while six designers and two users had a preference for Badge. One designer and one user were neutral. One would expect designers to have a strong preference for TouchToPrint since they were recently involved in the development of this product. This was not evident though. Two possible explanations might be given: a) Badge is not yet implemented, therefore treated as future development by designers, therefore assigned greater value than the already existing TouchToPrint or b) potential drawbacks of TouchToPrint are only evident after use, and hence more obvious to the designers who have actually experienced the product. The users, who had no actual experience with TouchToPrint, seemed to value it higher than designers did.

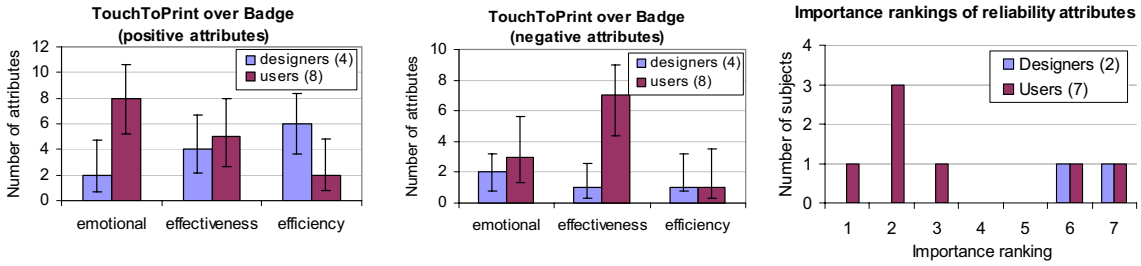

Figure 2. Attributes (a) positively and (b) negatively ranked when TouchToPrint is preferred (along with 95% exact confidence intervals). (c) Importance rankings for reliability attributes for designers and users.

To further understand this discrepancy between designers' and users' preferences, we analyzed their perceptions for these two products. Figure 2a illustrates the reasons supporting preference for TouchToPrint over Badge, as it shows the number of attributes that are positively ranked when TouchToPrint is preferred. While users' most frequent reason for preference for TouchToPrint was emotional attributes, for designers it was efficiency attributes. All attributes in the effectiveness category were related to security. TouchToPrint was perceived as more secure than Badge, both by designers and users. Users' most frequent negative concerns, shown in Figure 2b, were related to reliability (5 out of the 7 effectiveness attributes had to do with reliability). This is also evident in Figure 2c where we can observe that only two designers expressed reliability concerns and ranked them as the 6th and 7th most important attributes while five users ranked reliability within their three most important concerns. Hence, although most users prefer TouchToPrint, they have some concerns that can potentially turn into failures to motivate use.

4. CONCLUSION

This paper has presented a study that aimed to account for failures in motivating product use, resulting from discrepancies between designers' and users' perceptions of the product. The results from the study corroborate prior findings (Kujala and Kauppinen, 2004) suggesting that designers often fall short in accounting for users' views on the product and foreseeing their preferences. The results also suggest that current design practices might underestimate the importance of emotional aspects of product use. One could argue that this reflects the goal-oriented paradigm largely promoted by usability engineering. It thus becomes evident that design should be more concerned about a holistic view on the product qualities that motivate use. Finally, this paper proposes a methodological approach for the study of motivational aspects of product use, based on a relatively unexplored structured interview technique, the Repertory Grid Technique. RGT's major advantage is its ability to account for users' idiosyncratic views in contrast to predefined questionnaires where a priori defined attributes are imposed in the evaluation (Hassenzahl and Wessler, 2000). Next steps will attempt to enrich RGT analysis with Multi-Dimensional Scaling (Martens, 2003) and develop a systematic process for analyzing users' perceptions and preferences in product use.

ACKNOWLEDGEMENT

We thank Fred de Jong from Océ Technologies for his help throughout the study, and all participants for their valuable input. 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

Clarke, K. R. (1993) Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology***,** 117-143.

Cockton, G. (2004) From quality in use to value in the world. *CHI '04 extended abstracts on Human factors in computing systems.* Vienna, Austria ACM Press.

Fleiss, J. L., Levin, B. & Paik, M. C. (2003) *Statistical methods for rates and proportions* Wiley-Interscience.

Fransella, F., Bell, R. & Bannister, D. (2003) *A Manual for Repertory Grid Technique* Wiley.

Furlan, R. & Corradetti, R. (2006) Reducing Conjoint Analysis Paired Comparisons Tasks by a Random Selection Procedure. *Quality and Reliability Engineering International,* 22**,** 603-612.

Green, P. E., Carmone Jr., F. J. & Smith, S. M. (1989) *Multidimensional Scaling, Concepts and Applications,* Boston, MA, Allyn & Bacon.

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

Hassenzahl, M. & Wessler, R. (2000) Capturing design space from a user perspective: The Repertory Grid Technique revisited. *International Journal of Human-Computer Interaction,* 12**,** 441-459.

Keinonen, T. (1997) Expected usability and product preference. *Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques.* Amsterdam, The Netherlands, ACM Press.

Kelly, G. A. (1955) *The Psychology of Personal Constructs* New York, Norton.

Kujala, S. & Kauppinen, M. (2004) Identifying and selecting users for user-centered design. *Proceedings of the third Nordic conference on Human-computer interaction.* Tampere, Finland, ACM Press.

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

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

Martens, J.-B. (2002) Multidimensional modeling of image quality. *Proceedings of the Ieee Software,* 90**,** 133-153.

Martens, J.-B. (2003) *Image technology design: A perceptual approach,* Boston, Kluwer Academic Publisher.

Ouden, E. d., Yuan, L., Sonnemans, P. J. M. & Brombacher, A. C. (2006) Quality and Reliability Problems from a Consumer's Perspective: an Increasing Problem Overlooked by Businesses? *Quality and Reliability Engineering International,* 22**,** 821-838.

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

Sandt Van de, U. (1970) Incomplete paired comparisons using balanced lattice designs. *Journal of Marketing Research,* VII**,** 246-248.