The Fulfillment of User Needs and the Course of Time in Field Investigation

Claudia Nass

Fraunhofer Institute for
Experimental Software Engineering
Fraunhofer-Platz 1
67663 Kaiserslautern, Germany
claudia.nass@iese.fhg.de

Daniel Kerkow

Fraunhofer Institute for
Experimental Software Engineering
Fraunhofer-Platz 1
67663 Kaiserslautern, Germany
daniel.kerkow@iese.fhg.de

Jessica Jung

Fraunhofer Institute for
Experimental Software Engineering
Fraunhofer-Platz 1
67663 Kaiserslautern, Germany
jessica.jung@iese.fhg.de

Abstract

Business contexts represent a big challenge for software development, specifically in terms of finding a balance between business goals and users' goals. This context determines the utility of an application, but good user experience (UX) with business applications is only achieved if the software supports the fulfillment of users' goals and needs. This article presents the efforts realized in a call-center of a German telephone company aimed at enhancing UX and hence creating a positive influence on the emotional state of the users/employees. It describes a method applied for the elicitation of user needs as well as ideas for improving UX. Beyond that, the results indicate that software properties can influence the emotional state of the user if they support the fulfillment of human needs and thus positively affect the achievement of business goals.

Keywords

User experience, fun, user needs, emotion, call-center, field study, evaluation

ACM Classification Keywords

H5.2. [Information Interfaces and Presentation]: User Interfaces - Evaluation/methodology; Theory and methods.

General Terms
Design, Experimentation, Human Factors, Theory

Introduction

The development of business applications calls for inspiration and know-how on the part of the designers to find the balance between business and user goals. If an ideal balance is achieved between these goals, the software can support the achievement of the expected results for the company and also generate a positive user experience (UX) for its users. For many business domains, positive user experience is crucial for success, such as in the call-center context, where the work can involve confusion or frustration as well as joy and satisfaction in solving customers' problems [13]. Software in the context of a call-center means knowledge management systems that enable their users, the call agents, to communicate with customers, resolve problems, and develop relationships in order to achieve the business goals. Just making this product usable does not necessarily ensure that it is beneficial for the people [14], but it is also necessary to satisfy users' goals. Millard (2005) states that the call agents need to perceive the goodness of the software. One way to achieve this is to enhance emotional factors such as enjoyment, satisfaction, and fun. These emotional factors influence the usage of software; therefore, their effect can trigger a good or bad UX.

In this sense, Hassenzahl (2008) defines UX "as a momentary, primarily evaluative feeling (good-bad) while interacting with a product or service. By that, UX shifts attention from the products and materials (i.e., content, function, presentation, interaction) to humans and feeling – the subjective side of product use." The origin of this feeling, specifically of good UX, is the consequence of the achievement of user goals and also the fulfillment of human needs [8]. The term human need originates from motivation psychology. Some well-known theories are by McDougall (1908), Murray (1938), Hull (1952), Maslow (1954), and McClelland (1987) [9]. A need describes the cause of human behavior. For some theorists, it is the basis for a person's personality's development. These theories mention a large number of human needs, but Hassenzahl (2008) thinks that good UX is influenced by five needs: autonomy, competency, stimulation (self-oriented), relatedness, and popularity (others-oriented) [8].

Another interesting factor is presented by Forlizzi and Battarbee (2004). According to them, an experience is characterized as much by an explicit beginning and end as by the events that happen during this period. Hassenzahl (2007) argues that this assumption gives UX a temporal dimension that influences the evaluation of an interactive product, because it emphasizes that the user has to spend more time with a product to be able to build a reasoned opinion about his experience. The user's judgment may change after the initial contact with the software, and these changes are rarely explored in practice [6].

This article presents a study performed at a call-center of a German telephone company. By altering software properties, the study aimed at improving the user experience based on the fulfillment of the call agents' human needs as well as at achieving the business goals that should improve the quality of internal business processes. The work was executed in three phases. First, a UX Method, called KREA-FUN, was applied to identify the goals of the stakeholders and to generate

ideas for achieving these goals. Then, some of these ideas were implemented and integrated into a knowledge management system. Finally, the new application was evaluated in a case study with real users in the original working context. These steps are described in the first three sections of this article. The last section presents the results of the field investigation.

UX Method KREA-FUN

There are different methods for eliciting user needs and goals in the field of UX [12]. In order to enhance emotional factors such as enjoyment, satisfaction, and fun, we chose a method called KREA-FUN. KREA-FUN is "a systematic method to facilitate the elicitation of ideas for new and innovative ways how users of software might enjoy more fun when working with the software" [11]. The method was applied in a workshop where the different stakeholders were brought together to discuss their goals and ideas.

The software presented, Excalibur, was custom developed for this company and has been used by the employees for two years. Excalibur helps the call agents to capture the customers' personal data and their problems. It also guides the agent through the "best solution", like a wizard assistant. The solutions of Excalibur have been developed previously by the process department and are constantly updated based on the feedback given by the agents after a call.

Seven people participated in the workshop, including two expert call agents, two call-center managers, two software developers, and one user experience expert.

The first step of the workshop was to explore and define the user and business goals that are to be achieved by the implementation of new interaction concepts in the software. In this case, the following goals were selected and refined from an initial list of aims:

User goals

- Fun: The call agents do not perceive the image of Excalibur as attractive and pleasant; they do not have fun when using the software. The agents wanted to use Excalibur not only for working, but as a means to amuse themselves.

- Acceptance: The agents wanted to use Excalibur like a social platform to support the work relationship between the different groups within the company.

- Utility: The agents wanted additional functionalities to increase the use of Excalibur for secondary tasks, such as work break sensor or online chat.

Business Goals

- Increase the quality of the documentation: Excalibur offers ready-made solutions for previously detected problems. These solutions are called Troubleshoots. They guide the call agents to the best solution when attending a call, like a wizard application. This should be an interactive process; the agent has to click through this wizard function at each call. The problem is that the more experienced agents do not use this wizard function during a call, but rather give the clients a known solution and pick the shortest Troubleshoot

in Excalibur, even if this solution does not match the real problem. This generates a serious problem with the documentation about the clients' issues. The company wanted to reduce the gap between registered and real problems and therefore wanted to motivate the agents to capture the real problem in Excalibur in order to achieve a balance between the quantity and the quality of their work.

- Explore new solutions: The expert agents no longer use the wizard function to attend to the clients. They resolve the client questions based on their knowledge and experience gained in the past. They have their "favorite Troubleshoots" and use only these to attend the calls. This behavior is problematic because new and revised Troubleshoots remain unknown and unused. The objective in this case was to inspire and stimulate the agents to become acquainted with new Troubleshoots and learn new ways to treat the calls.

The second phase was characterized by finding ideas to improve the software used in the call-center and as a means for achieving the goals defined during the first part of the workshop. Using triggers from motivational theory, the participants were encouraged to generate innovative solutions based on free association. For this workshop part, two creativity techniques were used: Brainstorming, for generating initial concepts/topics, and the Lotus Blossom Technique, for routing the participants further away from controlled thinking, provoke divergent thinking, and elicit more ideas [11].

The third part of KREA-FUN was the prioritization of the ideas generated. Only the most important and interesting (for the participants) topics were refined. The ideas were assessed as being questionable or reliable; the reliable ones should offer more possibilities when implemented and tested as well as a stronger effect.

At the workshop, forty-four ideas were generated to improve Excalibur and achieve the goals described above. Out of these ideas, eleven were selected as candidates for the implementation. Four of these were assessed as being questionable and the other seven were assessed as being reliable. Later, these seven were fully implemented.

Implementation
The seven ideas were classified into two groups: The first group contained patterns from the social or computer-mediated interaction domain and the other group comprised functionalities related to Excalibur and the workflow of the call agents.

Excalibur is a web application and the new functionalities were embedded into its homepage as web widgets [18]. This homepage stays open as long as an agent has not answered a call. When a call is answered, this page is substituted by the wizard application of Excalibur. Figure 1 shows the prototype used by the agents during the test phase.

The following patterns from social domains were implemented in the application:

Figure 1. Homepage of released Excalibur.

1. *Profile*. This is a virtual representation of the agents that was seen by other employees on their new homepage. This pattern is related to the human need for popularity, because it helps the user to show his characteristics to the group and defines his position in the social net [16].

2. *Microblogging*. This pattern allows the call agents to publish short messages on a kind of personal web-site and to satisfy their communication need in this way [3].

3. *Collection.* The call agents observe the number of different Troubleshoots used. This number is a coherent metric of their own performance. This collection need is explained by Steven Reiss (2000) as a basic desire that drives our existence [15].

4. *Ranking.* With this pattern a call agent compares his own performance to that of other agents. The agents compare their own quantity of different Troubleshoots with the quantity of their colleagues.

This pattern is used in the context of games because it offers a basis for competition [1].

In addition, three functionalities were implemented in the software:

- *Top-hitter.* This widget shows the Troubleshoots used most often during the last few hours.
- *Excalibur update.* This is a list on the homepage of the software with the new and updated Troubleshoots in Excalibur.
- *Problem ticker.* This is a list recommending a specific Troubleshoot for a current problem.

Evaluation

The case study took place in a real environment, at a call-center of a German telephone company. Three groups of call agents participated. Group 1 (G1) consisted of seven participants who were extensively informed about the new functionalities. Group 2 (G2) also had seven agents, but they only got the new functionalities without any further explanation. The third group (CG) was the control group, which did not receive any of the new functionalities. The five agents of this team continued working with the old application.

Several studies present results about the first impression that a user has with a new system [6], but in our case we wanted to examine the effect of the new version of Excalibur for a longer period of time. Because of this, the call-agents were observed during a ten-week period and answered weekly surveys about their experience with the software. The employees did not receive any kind of compensation for their participation in the study.

Four exploratory questions provided directions for the data collection and analysis. The first interesting issue was to know whether the new functionalities could fulfill the users' needs, and in this way positively affect their emotional state during work. For this purpose, we utilized two instruments, an adaption from Hassenzahl (2008) of the Needs questionnaire by Sheldon et al. (2001) and the Self Assessment Manikin (SAM) by Bradley and Lang (1994).

With the Needs questionnaire, the users were asked about the fulfillment of five needs (competence, popularity, relatedness, stimulation, and autonomy). Each need had three items and a five-level Likert scale (from "not at all" to "very much"). The questionnaire's factor structure has been confirmed in three different studies (Sheldon et al. 2001), even though data about the reliability and validity of the scales is not available [17]. The users' emotional state was recorded by Bradley and Lang's SAM. The instrument presents two dimensions of emotion, valence and arousal. These are measured by pictograms, similar to a human figure, in combination with a nine-point rating scale. The valence scale consists of pictograms that show a happy/proud person on the one end and a sad/unsatisfied person on the other end. The other dimension, arousal, is represented by a calm and relaxed figure on the one end and an excited and tense human shape on the other end [2].

Our second question dealt with the perception of the product qualities. We wanted to know how strong the new functionalities influence user perception in terms of the quality of joyfulness. For this, we applied the questionnaire AMUSE, which focuses on recording the perceived quality of product aspects, namely effectiveness, productivity, joy of use, trust, and

competence [4]. AMUSE presents for each dimension four items and a seven-level Likert scale (from "strongly disagree" to "strongly agree"). For our study, we used only the scale joy of use, because of its relevance in terms of the users' goals.
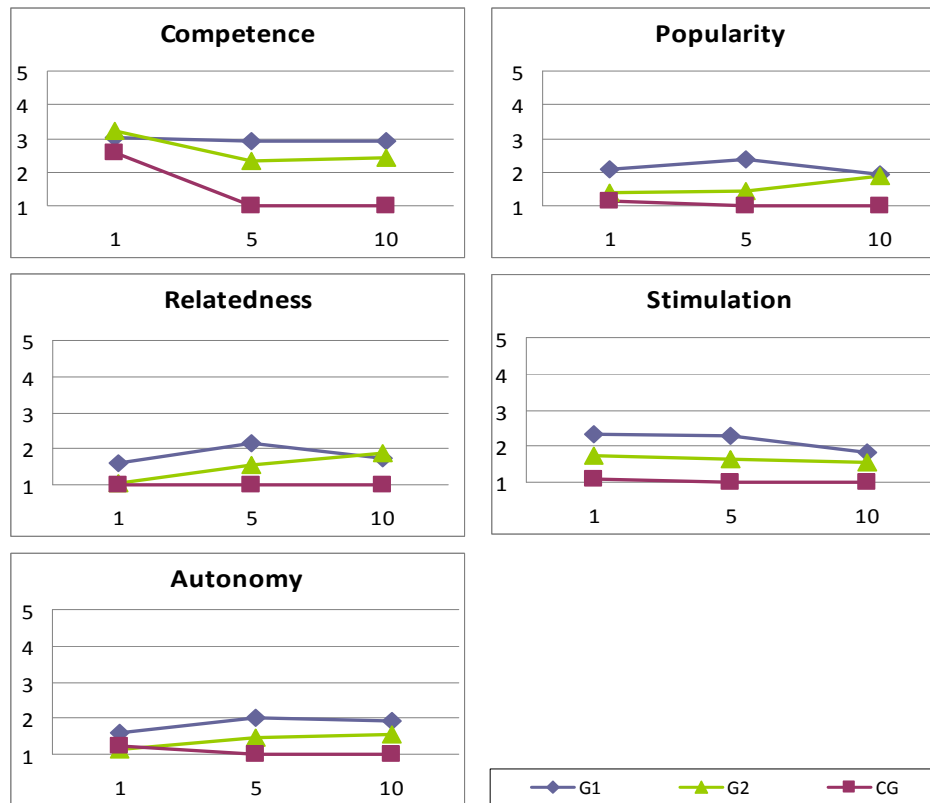

Figure 2. Perceived fulfillment of human needs
Observations: Graph shows mean of measurement during the 1st, 5th and 10th week.
Scale 1= not at all to 5= very much.

From the business goal point of view, two issues were important and built our third and fourth exploratory questions. Assuming that the new functionalities had a positive effect on the call agents' mood and their working habits, we expected an influence on the customers' satisfaction with the employees on the one hand, and, on the other hand, an improvement in the documentation quality of the calls.

To record the information regarding customer satisfaction, during the 10 weeks a randomized sample of customers was asked after their calls how satisfied they were with the call agents. Satisfaction was estimated in percentages (between 0% and 100%). Moreover, as objective measure, the use of Troubleshoots was recorded by log data integrated into the system for the whole period of time.

At the beginning, in the middle, and at the end of the 10-week period, the participants had to answer all questionnaires (Needs, SAM, and AMUSE). In the meantime, a condensed version was handed out, resulting in a total of seven measuring points (Needs with two items for relatedness, stimulation, and self-actualization; SAM Valence). The condensed version was necessary due to economic reasons (e.g., time consumed by answering the questionnaires). All questionnaires were available in German.

Results

19 call agents participated in the study, 11 females and 8 males, with an average age of 33.9 years (Min=22, Max=50, SD=8.5).

One part of our first question examined the fulfillment of user needs when the released Excalibur was used.

The development during the study period is shown in Figure 2. Competence was the need people evaluated highest during the work with Excalibur. But it was apparent that in G1 and G2 the competence need was more satisfied than in CG. Another experiment with a similar condition, or usage mode, also showed that resolving a task addresses the human need for competence [10]. These results serve as evidence for the prediction by Hassenzahl (2003). He argues that the usage mode impacts the user's experience with an interactive product and his retrospective judgment [7], meaning in our case that the task-oriented context strongly influences the fulfillment of the human need for competence.

The other needs were rated substantially lower with values around 2 ("a little") and 1 ("not at all"). It is clear that G1 rated the fulfillment of needs higher than the other two groups for all five needs; however, the statistical analysis (ANOVA) conducted showed no significant result, neither in comparing the groups nor over time.

The visually observable constant increase in the perceived fulfillment of the two needs popularity and relatedness in G2 is also interesting. G2 did not receive detailed information about the released Excalibur. This may indicate that the call agents had to take more time to explore the new functionalities and consequently felt the positive effect of their use more than the users of G1.

Regarding the emotional state of the user, the results showed that G1 had a more positive mood during the ten weeks of the study with less variation (see Figure 3). In the other groups, the variation was greater and

the call agents were in a worse mood than the agents of G1. While a t-Test shows no significant difference between G1 and G2 in the first week (t=-1.41; df=4; p=.23), a significant difference between these two groups in the fifth week (t=-7.00; df=4; p<.01) and also in the tenth week (t=-3.32; df=4; p=.03) can be reported. G1 showed over the course of time no explicit positive or negative trend (see figure 3), but rather relatively constant data, whereas G2 had an ambiguous time-series. The visually recognizable slight negative trend of CG could not be verified by the Neumann's trend test.
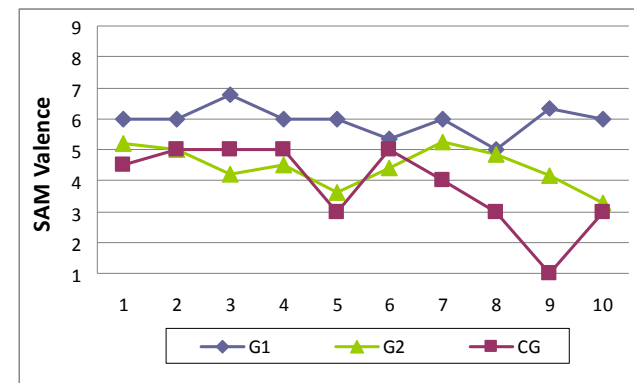

Figure 3. SAM Valence during the 10 weeks
Observations: Scale 1=sad/unsatisfied to 9=happy/proud.

Adding the SAM arousal scale to the interpretation of the data, it was observed that during the first week, the call agents were rather calm than tense in all groups. In the fifth week G1 was a bit more excited (valence: M=6.00, SD=1.00; arousal: M=5.00, SD=1.63) compared to the first week (valence: M=6.00, SD=1.41; arousal: M=3.57, SD=1.13), whereas CG was more unsatisfied and tense (valence:

M=3.00, SD=0.00; arousal: M= 5.00, SD=0.00). G2 was also more unsatisfied; unlike CG, however, the participants were unexcited (valence: M=3.60, SD=1.67; arousal: M=3.00, SD=2.00). Admittedly, these differences between the three groups for the three measuring points were too slight to generate any significant difference (ANOVA).

Need		SAM Valence		
		1st Week	5th Week	10th Week
Competence	1st Week	-.16		
	5th Week		.58	
	10th Week			.56
Popularity	1st Week	.27		
	5th Week		.48	
	10th Week			.35
Relatedness	1st Week	.53*		
	5th Week		.29	
	10th Week			.24
Stimulation	1st Week	.22		
	5th Week		.45	
	10th Week			.45
Autonomy	1st Week	.01		
	5th Week		.30	
	10th Week			.32

Table 1. Bivariate correlation between Needs and Valence
*Observations: Table shows correlation between measurements during the 1st, 5th and 10th week; * p < .05.*

A bivariate correlation analysis was used to evaluate the association between the two groups of variables. Table 1 shows the correlation between SAM Valence (positive and negative emotion) and the five human needs. There is a tendency towards a moderate correlation between the fulfillment of needs and the emotional state of the call agents. The different measuring points show that during the course of time, different needs were more correlated with the emotional state of the users than others. This could be evidence that in distinct situations, our needs are

satisfied differently and hence influence our mood positively or negatively.

The call agents using Excalibur with the new functionalities and with the additional detailed information about the new advantages of the software (G1) evaluated the software as considerably more joyful than the other two groups (see Figure 4). It was also observed that in this group, the effect was lengthy and remained stable throughout the evaluation period. In G2 and CG, the perception of the quality "joy of use" sharply decreased during the fifth week. Moreover, an ANOVA showed that all three groups differ significantly from each other (F=33.61; df=2; p<.01).
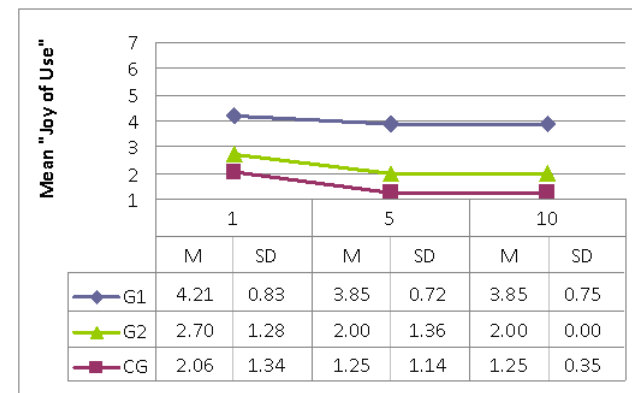

	1		5		10	
	M	SD	M	SD	M	SD
G1	4.21	0.83	3.85	0.72	3.85	0.75
G2	2.70	1.28	2.00	1.36	2.00	0.00
CG	2.06	1.34	1.25	1.14	1.25	0.35

Figure 4. Joy of Use
Observations: M, Mean - SD, Standard Deviation. Scale 1=strongly disagree to 7=strongly agree.

The averages of the 10-week study show that customers, attended by call agents from G1, were more satisfied than customers attended by agents from the other two groups. G1 had a customer satisfaction index of 65%, while G2 and CG had 60% and 57%,

respectively. Furthermore, the customers were significantly more satisfied with the support when the call agents use the new system in comparison to the support of the call agents using the former version of Excalibur (t=1.95; df=25; p=.03).
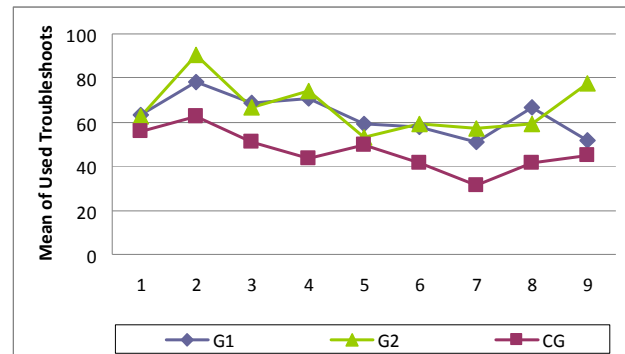

Figure 5. Mean of Troubleshoots used per week per group

Considering the use of different Troubleshoots, G2 had the best average during the ten weeks. The call agents in G2 had an average – rounded to the nearest full number – of 67 different Troubleshoots per week, while G1 had 63 and CG 47 Troubleshoots per week. In Figure 5, the development of the variation of Troubleshoots from the first until the ninth week can be observed. Although a slight negative trend for G1 and CG could be visually noticed, only for CG does the Neumann's trend test showed statistical significance (test statistic=.82, critical value=1.02 for p=.95). One reason for the greater use of Troubleshoots by G2 compared to G1 might be a strong concentration of the participants on the Ranking Pattern; this pattern is characterized by inciting competition and thus making the participants use a higher number of Troubleshoots.

Discussion and Conclusions

This study indicates that software properties can influence the emotional state of the user if they support the fulfillment of human needs. Consequently, choosing the right method for the elicitation of user needs is crucial for developing software that positively influences UX. KREA-FUN appears to be an appropriate method for eliciting these needs, especially when prioritization between business and user goals plays a role.

The fulfillment of human needs, on the other hand, was valued relatively low by the call agents. One reason for this result may be that the Needs questionnaire items present statements based on events that affect one's whole life, e.g. "during the use of Excalibur I felt that *I inspired others with my behaviour*", instead of questions that are specific for a business context. This indicates a current lack of methods for measuring the fulfillment of users' needs in such contexts.

The connection between the needs fulfillment and the enhancement of customer satisfaction is not transparent and we cannot deduce a causal effect between the two constructs. However, the difference between the customer satisfaction values in the respective groups calls for more extensive explorations in order to determine the origin of this improvement of the customer satisfaction index of G1.

The major difficulty in our field investigation was the data collection during the 10-week period. Context factors such as time pressure during the workday, vacation, illness, etc. negatively affected the data collection and hence influenced the internal validity of the study. Usually, field investigations possess a higher external validity compared to laboratory studies. In our

case, the external validity must be considered unsatisfying due to the low number of participants. Nevertheless, field investigation over a longer period of time provides rich information about people's behavior and the effects on the organizational processes and results, such as the enhancement of the customer satisfaction index in our study.

In terms of the difference between G1 and G2, one important factor refers to the presentation of new functionalities of a software solution in a business environment. The method in which new features are introduced to the users may accelerate the learning curve of the new functionalities and lead to earlier benefit offered by the system, and therefore should be considered carefully.

New studies are required to identify causal effects between the different constructs that are dealt with in this work, such as needs, emotion, and quality of use in business contexts. Furthermore it is necessary to develop methods and approaches that support the design of interactive products that will improve the fulfillment of user needs.

Acknowledgments

Citations

[1]   Bjork, S. and Holopainen, J. 2004. *Patterns in Game Design.* Charles River Media, Inc.

[2]   Bradley, M. M. and Lang, P. J. 1994. Measuring emotion: the self-assessment manikin and the semantic differential. In *Journal of Behaviour Therapy and Experimental Psychiatry*, 25, 49-59.

[3]   Crumlish, C. and Malone. E. 2009. *Designing Social Interfaces: Principles, Patterns, and Practices for Improving the User Experience.* O'Reilly Media, Inc.

[4]   Doerr, J., Hartkopf, S., Kerkow, D., Landmann, D., and Amthor, P. 2007. Built-in User Satisfaction - Feature Appraisal and Prioritization with AMUSE. In *Proc. RE 2007*. Los Alamitos IEEE Computer Society (2007), 101-110.

[5]   Forlizzi, J. and Battarbee, K. 2004. Understanding experience in interactive systems. In *Proceedings of the 5th Conference on Designing interactive Systems*. DIS '04. ACM, New York, NY, 261-268.

[6]   Hassenzahl, M. and Roto, V. 2007. Being and doing: A perspective on User Experience and its measurement. *Interfaces, 72,* 10-12.

[7]   Hassenzahl, M. 2003. The Thing and I: Understanding the Relationship Between User and Product. In M. A. Blythe, A. F. Monk, K. Overbeeke & P. C. Wright (Hrsg.), *Funology: From Usability to Enjoyment* (S. 31-42). Dordrecht: Kluwer Academic Publishers.

[8]   Hassenzahl, M. 2008, in press. User Experience (UX): Towards an experiential perspective on product quality. In: *Proc. IHM '08*.

[9]   Heckhausen, J. and Heckhausen, H. 2007. Motivation und Handeln (3. Aufl.). Heidelberg: Springer.

[10] Jung, J. 2009. User Experience und menschliche Bedürfnisse. Unpublished Diploma Thesis. University of Landau, Germany.

[11] Kerkow, D. and Graf, C. 2007. KREA-FUN: Systematic Creativity for Enjoyable Software Applications. In *Proc. FUN 2007*.

[12] Kuniavsky, M. 2003. Observing the User Experience: a Practitioner's Guide to User Research. Morgan Kaufmann Publishers Inc.

[13] Millard, N. 2005. Knowledge Weaving: Supporting Effective and Affective Interactions in the Call Centre. In: *12th International Conference on Artificial Intelligence in Education, The Netherlands*.

[14] Nielsen, J. and Levy, J. 1994. Measuring Usability: Preference vs. Performance. *Communications of the ACM*, 37:4, 66-75.

[15] Reiss, S. 2000. *Who am I: The 16 basic desires that motivate our actions and define our personalities?* New York: Tarcher/Putnam.

[16] Schuemmer, T. and Lukosch, S. 2007. *Patterns for Computer-Mediated Interaction.* Hoboken: John Wiley & Sons.

[17] Sheldon, K. M., Elliot, A. J., Kim, Y., and Kasser, T. 2001. What Is Satisfying About Satisfying Events? Testing 10 Candidate Psychological Needs. *Journal of Personality and Social Psychology*, 80, 325-339.

[18] Vora, P. 2009. *Web Application Design Patterns.* Morgan Kaufmann Publishers Inc.