Emotional Experiences and Quality Perceptions of Interactive Products

Sascha Mahlke[1] and Gitte Lindgaard[2]

[1] Centre of Human-Machine Systems, Berlin University of Technology,
Franklinstr. 28/29 – FR2-7/2, 10587 Berlin, Germany
sascha.mahlke@zmms.tu-berlin.de
[2] Human-Oriented Technology Lab, Carleton University,
1125 Colonel By Drive, Ottawa, K1S 5B6, Canada
gitte_lindgaard@carleton.ca

Abstract. Over the past few years, various novel approaches have been applied to the evaluation of interactive systems. Particularly, the importance of two categories of concepts has been emphasized: non-instrumental qualities and emotions. In this paper we present an application of an integrative approach to the experimental study of instrumental and non-instrumental quality perceptions as well as emotional user reactions as three central components of the user experience. A study is presented that investigates the influence of system properties and context parameters on these three components. The results show that specific system properties independently influence the perception of instrumental (i.e. usability) and non-instrumental qualities (i.e. visual aesthetics). Especially the perception of instrumental qualities was shown to have an impact on the users' emotional reactions (subjective feelings as well as cognitive appraisals). There was also evidence suggesting that context parameters influenced emotional user reactions.

1 Introduction

To date, approaches to the evaluation of interactive systems have mainly focused on tasks and goals, their efficient achievement, and the cognitive information processing involved [1]. In the past few years, various ideas have been discussed that go beyond the notion of efficiency and that aim to better understand how people experience technology. In this regard, two important concepts have been explored: non-instrumental qualities and emotions. In the next sections we will discuss these two concepts and their relations.

1.1 Non-instrumental Quality Perceptions

Traditionally, evaluations of technology have focused on instrumental aspects of interactive systems, predominantly the concepts of usefulness and usability. Non-instrumental qualities on the other hand can be described as quality aspects that address user needs that go beyond tasks, goals and their efficient achievement.

J. Jacko (Ed.): Human-Computer Interaction, Part I, HCII 2007, LNCS 4550, pp. 164–173, 2007.

Mahlke [2] reviewed various approaches to the study of non-instrumental quality aspects. Briefly, he argued that two distinct categories of non-instrumental qualities have been differentiated in most approaches. On the one hand, aesthetic aspects have been discussed. These contain first and foremost visual aspects of product appearance, but can also imply other sensory experiences like haptic or auditory aspects of product use, as for example discussed by Jordan [3] and captured in his definition of physio-pleasure. The other category refers to a symbolic dimension of product appearance. The concept of hedonic quality discussed by Hassenzahl [4] belongs to this category, which is similar to what Jordan [3] calls socio- and ideo-pleasure.

Although much is being said about non-instrumental quality aspects and their application to design, only a few empirical studies actually measuring these have been reported. In a study of the interplay of non-instrumental quality perceptions with other concepts, Tractinsky, Katz and Ikar [5] highlighted the connection between aesthetics and usability. They argue that users' aesthetic judgment made before using an interactive system affects their perceived usability even after using it. Lindgaard & Dudek [6] found a more complex relationship between these two concepts. Hassenzahl [4] studied the interplay between usability and hedonic quality in forming overall judgments concerning beauty and goodness. He found that judgments of beauty are more influenced by the user's perception of the hedonic qualities, while judgments of goodness - as a more general evaluative construct - are affected by both hedonic quality and usability.

Although a few empirical studies do exist that contribute to a better understanding of the role of non-instrumental qualities and their interplay with other relevant aspects of technology use, many questions remain to be addressed. In particular, the relationships between quality perceptions and emotional experiences have barely been explored.

1.2 Emotions as Part of the User Experience

Rafaeli and Vilnai-Yavetz [7] attempted to link quality perceptions and emotional experience. They suggested that artifacts should be analyzed in terms of three conceptually distinct quality dimensions: instrumentality, aesthetics, and symbolism. They conducted a qualitative study in a non-interactive product domain to better understand the influence of these three quality dimensions on emotional responses. All three categories contributed significantly to the emergence of emotion. Tractinsky and Zmiri [8] applied this idea to an interactive domain by studying various existing websites which yielded similar results, and Mahlke's [9] study on actual audio players showed that various instrumental and non-instrumental quality perceptions influenced users' emotional responses.

While Rafaeli and Vilnai-Yavetz [7] used interviews, Tracinksy and Zmiri [8] and Mahlke [9] applied questionnaires to assess users' emotional responses. All these studies focused on the subjective feelings that arise when perceiving or using the relevant products. Much research has been conducted on measurements of emotion during interaction with technical devices, and different methods have been proposed to measure emotions in interactive contexts. Mahlke, Minge and Thüring [10] used Scherer's [11] multi component model of emotion to structure a range of relevant emotion-measurement methods and relating them to the five components of emotion:

subjective feelings, facial expressions, physiological reactions, cognitive appraisals and behavioral tendencies.

Taken together, there are two major problems with the interpretation of results emerging from the studies reported above that relate emotional experiences during the interaction with users' quality perceptions [7, 8, 9]:

1. They took a quasi-experimental approach by using existing products. As it was not discussed which properties of the stimuli or other variables influenced quality perceptions and the emotional experience, this question remains unanswered.
2. Rather than measuring all the five components of Scherer's [11] model, only subjective feelings were measured as indicators of emotions.

1.3 Research Approach

Mahlke and Thüring [12] describe an integrated research approach to the experimental study of emotional user reactions considering both instrumental and non-instrumental quality perceptions of interactive systems. Their model defines instrumental and non-instrumental quality perceptions as well as emotional reactions as three central components of the user experience, claiming that characteristic of the interaction affect all three of these. These characteristics primarily depend on system properties, but both user characteristics and context parameters like aspects of the tasks and the situation can play an important role. The outcomes of the users' interactive experience as expressed in overall judgments of a product, usage behavior or choices of alternatives are shown to involve all three components, namely emotional user reaction as well as instrumental and non-instrumental quality perceptions.

This model has been applied to study the influence of system properties on the three user experience components and users' overall appraisal of the system [12]. In an effort to affect the perception of instrumental qualities as well as user performance, the level of usability was systematically varied as were other system properties modified expected to affect perception of visual aesthetics. Emotions were measured in terms of subjective feelings, motor expressions and physiological responses. The results confirmed that the manipulations had the predicted impact on the perception of both instrumental and non-instrumental qualities. Prototypes high in usability and attractiveness were significantly rated more highly than those that were low in both aspects. The results of the questionnaire assessing subjective feelings showed an effect of both factors. They revealed that the effect of variations in usability was greater than variations in visual aesthetics on both valence and arousal measures. Consequently, the high-usability/high-aesthetics prototype was experienced as most satisfying, while the low-usability/low-aesthetics was found to be most annoying. Since no statistical interaction of usability and aesthetics was found, both factors contributed additively to these emotions. EMG data of facial muscle sites and other physiological measures (dermal activity and heart rate) supported this interpretation.

The following study is based on the same research approach, but differs in two aspects. First, the measurement of emotions focuses on subjective feelings and cognitive appraisals to learn more about another component of emotions defined by Scherer [11], and second, task demands were varied as an example for contextual parameters. Hassenzahl, Kekez and Burmester [13] found that the influence of instrumental and non-instrumental quality perceptions on overall judgments differs

depending on whether users are in a goal- or action-mode. In the goal-mode participants were required to accomplish given tasks, while they had the same amount of time to explore the system on their own in the action-mode. This variation was applied to investigate the effect of context parameters on emotional responses. The following predictions were made:

1. The versions with higher levels of usability and/or visual aesthetics would lead to higher instrumental and/or non-instrumental quality ratings.
2. Quality ratings would not be influenced by the usage mode [13].
3. The versions with higher levels of usability and/or visual aesthetics would lead to differences in the cognitive appraisal of the usage situation and more positive subjective feelings.
4. In goal-mode, the correlation between instrumental quality perceptions and subjective feelings would be higher than between non-instrumental quality perceptions and subjective feelings. In action-mode the opposite would be found.

2 Method

The variables investigated concerned the influence of system properties associated with usability and aesthetics of the system and task demands, that is, goal- versus action-mode, on the perception of instrumental and non-instrumental qualities and emotional user reactions. These included subjective feelings and cognitive appraisals.

2.1 Participants

Eighty undergraduate students (48 women, 32 men) participated in the study. They were between 18 and 54 years old (average 21.3 years) and received course credit for participation in the study. Most of the participants (n = 72) owned a portable audio player and used it regularly. Almost all (n = 78) used computers daily.

2.2 Material

Portable audio players were chosen as the domain of study and different versions were simulated on a computer. The aim of the variation of system attributes was to influence perceived usability and aesthetics of the system independently.

To produce two versions with different levels of usability, three system features were varied: the number of menu lines shown (five versus two), a scrollbar indicating available but hidden menu items (given or not), a cue about the present position in the menu hierarchy (given or not). These variations had been used in a previous experiment [12] in which the effect of these on usability varied in the direction one would predict, that is, the most usable version resulted in the highest usability ratings.

With respect to system features designed to influence the perception of visual aesthetics, two different body designs were used in the earlier experiment [12] varying in symmetry (high or low), color combination (high or low color differences) and shape (round or square). Because these manipulations resulted only in small differences in perceived aesthetics between the two versions, an attempt was made here to improve the high-aesthetic version by consulting a professional designer.

The prototypes were presented on a 7" TFT-display with touch screen functionality that participants could hold in their hands for providing input. The display was connected to a computer which ran the simulation of the audio player.

2.3 Design

Three independent variables were manipulated: 'usability', 'visual aesthetics', and 'mode' (goal- vs. action-mode). Since each of the variations of 'usability' and 'visual aesthetics' had two levels ('high' and 'low'), four prototypes were created: (a) 'high-usability' and 'high-aesthetics', (b) 'high-usability' and 'low-aesthetics', (c) 'low-usability' and 'high-aesthetics', (d) 'low-usability' and 'low-aesthetics'. In the goal-mode participants were required to accomplish a set of tasks, and in the action-mode they were freely browsing the system for the same amount of time. All three variables were between-subjects factors.

2.4 Measures

Two types of behavioral data were recorded in the goal-mode condition to ensure that versions of assumed high or low usability differed as planned: task completion rates and time on task.

Questionnaires were employed to assess the user's perception of instrumental and non-instrumental qualities. Selected sub-dimensions (controllability, effectiveness, helpfulness, learnability) of the Subjective Usability Measurement Inventory (SUMI) [14] served to rate usability. The dimension 'classical visual aesthetics' of a questionnaire developed by Lavie and Tractinsky [15] was used to measure visual aesthetics.

Subjective emotional data were obtained via the Self-Assessment Manikin (SAM) [16] which captures the quality, or valence (positive/negative), and intensity (arousal) of emotions.

Cognitive appraisals were obtained via a questionnaires based on the Geneva Appraisal Questionnaire [17]. It measures five appraisal dimensions: intrinsic pleasantness, novelty, goal/need conduciveness, coping potential, and norm/self compatibility. Novelty is a measure of familiarity and predictability of the occurrence of a stimulus, while intrinsic pleasantness describes whether a stimulus event is likely to result in a positive or negative emotion. A goal conduciveness check establishes the importance of a stimulus for the current goals or needs. Coping potential refers to the extent to which an event can be controlled or influenced. Norm/self compatibility describes the extent a stimulus satisfies external and internal standards.

2.5 Procedure

The experiment took roughly 30 minutes on average. Participants were given instructions describing the experimental procedure and the use of SAM. They were then asked to rate their subjective feelings as a baseline measure. Then, depending on the experimental condition to which they were assigned at random, the relevant player was presented and participants rated its visual aesthetics. Next, they read a short text describing how to use the system.

Participants were then asked either to complete the set of five tasks or to explore the system for a certain amount of time. In the goal-mode condition a limit of two minutes was set for each task. Typical tasks were 'Please have a look which songs you find on the player in the Genre POP' or 'Please change the sound setting of the player to CLASSIC'. However, participants actually completed the five tasks in five minutes on average. Therefore, a five-minute time limit was also set for the browsing participants.

In the task condition participants filled in SAM scales after the first, third and fifth task. In the browsing condition, they were asked to rate their current subjective feeling after one, three and five minutes of exploration. At the end of this, the cognitive appraisal questionnaire was completed and usability ratings were obtained.

3 Results

A 2x2 ANOVA for 'usability' and 'visual aesthetics' was performed on the goal-mode data only, assessing task-completion rates and task-completion time. There was a significant main effect for 'usability' only, for both task-completion rates, $F(1,38)=9.20$, $p < .01$, and task-completion time, $F(1,38)=13.10$, $p < .01$. Thus, high usability led to better performance on both measures.

3.1 Instrumental and Non-instrumental Quality Perception

Table 1 summarizes the average usability and visual aesthetics ratings for each condition. The ratings were transformed to values between 0 and 1 because the range of ratings differed between the variables. The Table shows that the average ratings were comparatively high even in the low-usability and the low-aesthetics conditions.

Table 1. The first number in each cell represents the average usability rating and the second number the average visual aesthetics rating for each condition (ratings are transformed to values between 0 and 1)

	Goal-mode (tasks)		Action-mode (exploration)	
	Usability high	Usability low	Usability high	Usability low
Aesthetics high	0.74 / 0.80	0.59 / 0.78	0.78 / 0.82	0.59 / 0.72
Aesthetics low	0.77 / 0.58	0.55 / 0.57	0.73 / 0.58	0.54 / 0.50

A 2x2x2 ANOVA for 'usability', 'visual aesthetics' and 'mode' performed on the usability ratings revealed a significant main effect for 'usability' only, $F(1,72)=9.0$, $p < .01$. A similar 2x2x2 ANOVA carried out on the visual aesthetics ratings showed a significant main effect for 'visual aesthetics' only, $F(1,72)=34.3$, $p < .001$. Consistent with hypotheses 1 and 2, this suggests that the system properties affected the perception of both instrumental (i.e. usability) and non-instrumental qualities (i.e. visual aesthetics), and that quality perceptions were not influenced by usage mode.

3.2 Emotional User Reactions

A series of 2x2x2 ANOVAs for 'usability', 'visual aesthetics' and 'mode' on each of the five cognitive appraisal dimensions showed that participants rated the intrinsic pleasantness of the interaction higher for the high-usability than for the low-usability version, $F(1,72)=3.9$, $p < .05$. Furthermore, the experience with the low-usable system was rated as more novel, $F(1,72)=5.6$, $p < .05$, and self/norm compatibility was higher for the high-usability version, $F(1,72)=5.2$, $p < .05$. Neither 'visual aesthetics' nor 'mode' influenced intrinsic pleasantness, novelty or self/norm compatibility, and goal conduciveness as well as coping potential showed no significant effect for any of the independent variables. In summary then, we found partial support for hypothesis 3: differences in cognitive appraisals for three of the appraisal dimensions and only the factor 'usability' had a significant influence.

For the analysis of subjective feelings we calculated the changes from the baseline value obtained at the beginning of the experiment to the three values assessed during the interaction for each participant. For the changes from the baseline to the first two assessments of subjective feelings the 2x2x2 ANOVAs with 'usability', 'visual aesthetics' and 'mode' as independent variables revealed no significant effects for either the dimensions valence or arousal. Figure 1 shows the average subjective feeling changes to the third data point at the end of the interaction for the four prototypes. A 2x2x2 ANOVAs for 'usability', 'visual aesthetics' and 'mode' and the changes in valence as dependent variable revealed a significant effect for 'usability' only, $F(1,72)=25.5$, $p < .05$. The ANOVA for arousal as dependent variable showed no significant effects. Thus, only 'usability' affected the valence of subjective feelings, what again only partially supported hypothesis 3.

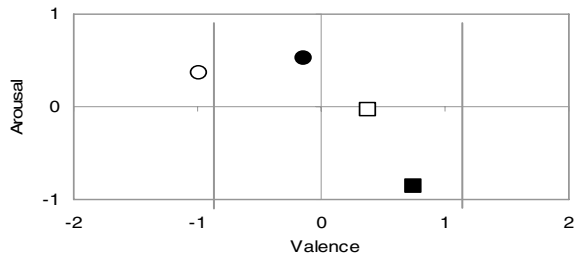

Fig. 1. Changes of subjective feeling ratings from the beginning of the experiment to the third assessment during the interaction with the system for the four systems (squared high vs. round low usability; filled high vs. unfilled low aesthetics; SAM ratings were between 0 and 8)

In order to test prediction 4 we conducted partial correlations to assess the correlation of usability and visual aesthetic ratings and subjective feelings in the two usage situations. As shown in Table 2 we found a high correlation for perceived usability and valence in the goal-mode, but not for perceived aesthetics and valence. For arousal none of the correlations was significant. For the action-mode the results yielded a moderately significant correlation with perceived usability and also with perceived aesthetics. For arousal again none of the correlations was significant.

Table 2. Correlation coefficients between quality ratings (usability and visual aesthetics) and subjective feelings (valence and arousal)

	Goal-mode (tasks)	Action-mode (exploration)
perceived usability – valence	.66 [a) **]	.35 [a) *]
perceived aesthetics – valence	-.01 [b)]	.35 [b) *]
perceived usability – arousal	-.16 [a)]	-.19 [a)]
perceived aesthetics – arousal	.04 [b)]	.22 [b)]

Partial correlation coefficients with [a)] visual aesthetics controlled and [b)] usability controlled
[*] $p < .05$; [**] $p < .01$

4 Discussion

As stated in hypothesis 1, system properties did independently influence instrumental as well as non-instrumental quality perceptions. Both usability and aesthetics manipulations affected subjective predictions in the predicted directions. In comparison to other studies [5, 18], we did not find any influence of the visual aesthetics variation on perceived usability. One reason may be that in other studies an overall usability rating was used, while we applied a detailed measure for usability. No effect of the factor 'mode' was found on quality perceptions (prediction 2) as one would have expected based on Hassenzahl et al.'s [13] findings.

The integration of cognitive appraisals as another component of emotions followed the recommendations by Mahlke et al. [10] to consider different components of emotions. We found an influence of the factor 'usability' on cognitive appraisals. The interaction with the low-usability system was experienced as less intrinsically pleasant, which corresponds to the findings regarding the subjective feelings. Furthermore, participants rated it as more novel or unusual, which may have led to more negative subjective feelings. The low-usability system was also rated as less self/norm compatible. Although this experiment is another step to the study of cognitive appraisals in interactive contexts, further research is clearly needed on this topic.

In terms of the users' subjective feelings, these were only affected by variations in usability. Furthermore, only the valence dimension was influenced. Participants' subjective feelings were more positive in the high usability condition towards the end of the experiment compared to the beginning. Surprisingly, we did not find an effect of 'visual aesthetics', although we tried to improve the differences in visual aesthetics in comparison to a previous experiment [12].

The variation of usage mode revealed differences in the connections between quality perceptions and participants' subjective feelings. These differences were most pronounced for the subjective feeling dimension of valence. While there was a high correlation between the valence of users' subjective feelings and the perceived usability of a system and no correlation with the perceived visual aesthetics when participants focused on the given tasks in the goal-mode, we found moderate correlations between valence and both perceived usability and aesthetics when participants were merely exploring the system. These results indicate that context

parameters like usage mode influence both the specific quality dimensions for overall judgments [13], and also the quality of the emotional experience.

However, more research is needed on these relationships, especially with respect to the subjective feeling dimension of arousal. In future studies the influence of user characteristics should also be studied in addition to system properties and context parameters. Furthermore, the variation of system properties that influence non-instrumental qualities other than visual aesthetics (e.g. haptic and acoustic quality) may reveal important insight especially for the domain of consumer electronic products.

Acknowledgements. This research was supported by the German Research Foundation (DFG) as part of the Research Training Group 'Prospective Engineering of Human-Technology Interaction' (no. 1013) and by the German Academic Exchange Service (DAAD) with a travel grant. We would like to thank Lucienne Blessing, Manfred Thüring and various colleagues at the Center on Human-Machine-Systems in Berlin and the Human-Oriented Technology Lab in Ottawa for the discussions on the study.

References

1. ISO: ISO 9241: Ergonomic requirements for office work with visual display terminals. Part 11: Guidance on usability. ISO, Genf. (1998)
2. Mahlke, S.: Aesthetic and Symbolic Qualities as Antecedents of Overall Judgements of Interactive Products. In: Bryan-Kinns, N., Blanford, A., Cruzon, P., Nigay, L. (eds.) People and Computers XX - Engage, pp. 57–64. Springer, Heidelberg (2006)
3. Jordan, P.W.: Designing pleasurable products. Taylor & Francis, London (2000)
4. Hassenzahl, M.: The Interplay of Beauty, Goodness, and Usability in Interactive Products. Human-Computer Interaction 19, 319–349 (2004)
5. Tractinsky, N., Katz, A.S., Ikar, D.: What is beautiful is usable. Interacting with Computers 13, 127–145 (2000)
6. Lindgaard, G., Dudek, C.: What is the evasive beast we call user satisfaction? Interacting with Computers 15(3), 429–452 (2003)
7. Rafaeli, A., Vilnai-Yavetz, I.: Instrumentality, aesthetics and symbolism of physical artifacts as triggers of emotion. Theoretical Issues in Ergonomics Science 5, 91–112 (2004)
8. Tractinsky, N., Zmiri, D.: Exploring Attributes of Skins as Potential Antecedents of Emotion in HCI. In: Fishwick, P. (ed.) Aesthetic Computing, MIT Press, Cambridge (2006)
9. Mahlke, S.: Studying user experience with digital audio players. In: Harper, R., Rauterberg, M., Combetto, M. (eds.) ICEC 2006. LNCS, vol. 4161, pp. 358–361. Springer, Heidelberg (2006)
10. Mahlke, S., Minge, M., Thüring, M.: Measuring multiple components of emotions in interactive contexts. In: CHI '06 extended abstracts on human factors in computing systems, pp. 1061–1066. ACM Press, New York (2006)
11. Scherer, K.R.: What are emotions? And how can they be measured? Social Science Information 44, 693–727 (2005)

12. Mahlke, S., Thüring, M.: Antecedents of Emotional Experiences in Interactive Contexts. In: CHI '06 proceedings on human factors in computing, ACM Press, New York (2007)
13. Hassenzahl, M., Kekez, R., Burmester, M.: The importance of a software's pragmatic quality depends on usage modes. In: Lucsak, H., Cakir, A.E., Cakir, G. (eds.) (WWDU2002). Proceedings of the 6th international conference on Work With Display Units, pp. 275–276. ERGONOMIC Institut für Arbeits- und Sozialforschung, Berlin (2002)
14. Kirakowski, J.: The software usability measurement inventory: Background and usage. In: Jordan, P.W., et al. (eds.) Usability Evaluation in Industry, pp. 169–178. Taylor & Francis, London (1996)
15. Lavie, T., Tractinsky, N.: Assessing dimensions of perceived visual aesthetics of web sites. International Journal of Human-Computer Studies 60, 269–298 (2004)
16. Lang, P.J.: Behavioral treatment and bio-behavioral assessment: Computer applications. In: Sidowski, J., Johnson, H., Williams, T. (eds.) Technology in Mental Health Care Delivery Systems, pp. 119–137. Ablex Publishing, Greenwich (1980)
17. Scherer, K.R.: Appraisal considered as a process of multi-level sequential checking. In: Scherer, K.R., Schorr, A., Johnstone, T. (eds.) Appraisal processes in emotion: Theory, methods, research, pp. 92–120. Oxford University Press, New York, Oxford (2001)
18. Ben-Bassat, T., Meyer, J., Tractinsky, N.: Economic and Subjective Measures of the Perceived Value of Aesthetics and Usability. ACM Transaction on Computer-Human Interaction 2, 210–234 (2006)