Hedonic *and* Ergonomic Quality Aspects Determine a Software's Appeal

Marc Hassenzahl, Axel Platz, Michael Burmester and Katrin Lehner
Corporate Technology - User Interface Design, Siemens AG
81730 Munich, Germany
+49 (0) 89 636-49653
marc.hassenzahl@mchp.siemens.de

"we continue to see [...] the prospect of a decade of research analysis of usability possibly failing to provide the leverage it could on designing systems people will really want to use by ignoring what could be a very potent determination of subjective judgements of usability - fun" (p. 23) [6].

ABSTRACT

The present study examines the role of subjectively perceived *ergonomic quality* (e.g. simplicity, controllability) and *hedonic quality* (e.g. novelty, originality) of a software system in forming a judgement of appeal. A hypothesised research model is presented. The two main research question are: (1) Are ergonomic and hedonic quality subjectively different quality aspects that can be independently perceived by the users? and (2) Is the judgement of appeal formed by combining and weighting ergonomic and hedonic quality and which weights are assigned?

The results suggest that both quality aspects can be independently perceived by users. Moreover, they almost equally contributed to the appeal of the tested software prototypes. A simple averaging model implies that both quality aspects will compensate each other.

Limitations and practical implication of the results are discussed.

Keywords

perceived software quality, emotional usability, hedonic components, joy of use

INTRODUCTION

In 1988, Carroll and Thomas [6] admonished us not to confuse the concepts *easy to use* and *fun to use* when talking about software quality. They argued that ease of use implies simplicity, which in turn is partly incompatible with fun. By making something as simple as possible, there is a good chance to make it boring as well. On the other hand, fun requires a subtle balance of not being too simple and not being too challenging (see [7]). As Carroll and Thomas put it: "we do not necessarily want to merely make things as simple as possible. We ought to make them fun" (p. 22). They call for the scientific study of fun and specialised design methods for promoting fun.

Since then *fun of use* did not receive that much attention in the field of Human-Computer Interaction. It plays a minor role in Technology Acceptance literature. For example, Davis et al. [8] showed that perceived fun (defined as "the extent to which using a software system is enjoyable in its own right") can accelerate usage intention if the software system is already perceived as useful ("the extent to which a person believes that using a particular software system would enhance his or her job performance"). Fun had no effect on usage intentions if a software system was not regarded as useful.

Igbaria, Schiffman and Wieckowski [12] studied the impact of perceived usefulness and perceived fun on both system usage and user satisfaction in a work context. Their results showed an almost equal effect of perceived fun and perceived usefulness on system usage. Perceived fun had even a stronger effect on user satisfaction than perceived usefulness. User satisfaction in turn had a clear effect on system usage. It can be concluded that enhancing perceived fun will primarily lead to increased time spent with a software system. This in turn may lead to a better understanding and/or a more productive use of the system. Furthermore, enhancing perceived fun should be a valuable road to directly increase user satisfaction.

Both studies succeed in demonstrating the positive impact of perceived fun - a factor independent of the efficiency and effectiveness of a software system - on usage intentions and user satisfaction. What these studies do not provide is an idea which design features will increase perceived fun.

One approach to find design principles which have the power to promote fun/enjoyment of a software system is to analyse what makes computer games fun [6;19;20]. Malone [19] identifies the following three broad design categories: "Challenge", "Fantasy" and "Curiosity". Each consists of several principles and recommendations for designing an appealing computer game. Among those, principles both consistent and contradictory to the notion of usability can be found. For example, the recommendation "provide a fantasy" is generally consistent with the idea of using a "metaphor" to increase familiarity and thereby the usability of a system.

However, besides the cognitive aspects (increased familiarity) Malone stresses the emotional aspects of a metaphor used, i.e. its power to satisfy the user's emotional needs. Mainly contradictory to usability is the principle to foster curiosity by designing a system that is novel and surprising (but not completely incomprehensible). In order to work, novelty and surprise must impair at least the external consistency of the software, a core principle of usability. In usability design, this might conflict with the task-related efficiency and effectiveness of the software system. This raises the question, whether it is appropriate to call for enjoyment and fun when it comes to work-related software systems (e.g. [11]).

The strong focus on task-related efficiency and effectiveness arises from the implicit notion that the computer is a tool - or even stronger - *has to be* a tool when applied at the workplace. This somewhat conservative perspective where "things must be taken seriously" is predominant in business and might be at least partly wrong considering the role of perceived fun/enjoyment in the context of technology acceptance at the workplace. The somewhat narrow focus of usability on task-related efficiency and effectiveness is eventually criticised in studies concerning consumer products [1;14;18].

What is needed is an expanded concept of usability which adopts enjoyment and satisfaction of the user as the major design goal[1]. Obviously, in a work context designing for efficiency and effectiveness (i.e. traditional usability) will be still a valuable way to reach the goal. Nevertheless, something should be added which makes the software system interesting, novel, surprising etc. Being both usable and interesting, a software system might be regarded as appealing and as a consequence the user may enjoy using it. Such an expanded perspective on usability would take us a step further toward designing *user experiences* [16] instead of merely making a software usable.

An expanded concept of usability was suggested by Logan [17]. He has developed a two-component usability concept that consists of behavioural and emotional usability. Behavioural usability refers to more or less traditional usability. Emotional usability refers "to the degree to which a product is desirable or serves a need(s) beyond the traditional functional objective" (p. 61). Although important, the concept of emotional usability still leaves many questions unanswered. For example, Logan provides no model or data about whether and how behavioural and emotional usability influence each other.

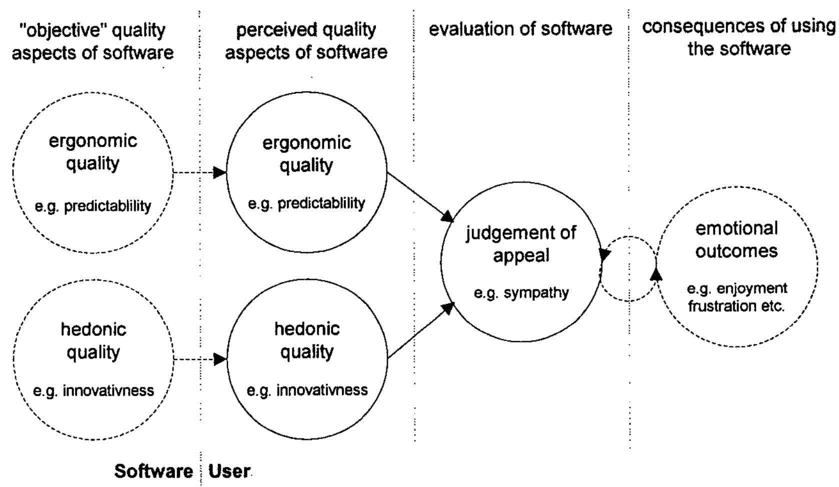
Another line of research to broaden the traditional usability concept mainly considers the impact of visual design on the expected (apparent) usability of a system *before* actually using it [5;15;22]. By demonstrating the impact of visual design on expected (apparent) usability, these studies succeed in establishing visual design as a key design factor. However, this is rather done by regressing visual design on the traditional usability concept than expanding usability itself.

This paper attempts to provide and test a model for an expanded concept of usability that incorporates key factors for designing appealing, enjoyable software interfaces and systems.

APPEALING SOFTWARE SYSTEMS: A HYPOTHESIZED MODEL

Figure 1 shows the elements of the hypothesised model for appealing software systems, i.e. quality aspects of the software system, evaluation of the system and consequences of using it.

Fig. 1: Research Model



A software system can be described on different quality dimensions such as its predictability, controllability etc. The model distinguishes between two groups of quality dimensions (i.e. aspects): *ergonomic* and *hedonic* quality.

Ergonomic quality (EQ) comprises quality dimensions that are related to traditional usability, i.e. efficiency and effectiveness (e.g. [13]). EQ focuses on task-related functions or design issues.

Hedonic quality (HQ) comprises quality dimensions with no obvious relation to the task the user wants to accomplish with the system, such as originality, innovativeness, beauty etc. Although not task-related, the users may regard HQ as an important quality aspect for its own sake.

[1] We are aware that user satisfaction is a part of the usability concept provided by ISO 9241-11. However, it seems as if satisfaction is conceived as a consequence of user experienced effectiveness and efficiency rather than a design goal in itself (see [10] for an example). This implies that assuring efficiency and effectiveness alone guarantees user satisfaction.

Fig. 2: Prototypes

Whether a software system is viewed as appealing by the users or not depends heavily on the users' perception of the quality aspects. For example, "objective" usability (i.e. inherent or intended by the software designer) is not always also perceived by the users [15]. Clearly, a software system designed to be as simple as possible fails if this simplicity can not be perceived or experienced by the users. For this reason, we focus on the subjective perceptions of EQ and HQ provided by the users.

The user's *judgement of appeal* (APPEAL) differs from the mere quality aspects of an additional evaluation. The judgement may be formed by weighting and combining different perceptions of the system's quality aspects (EQ and HQ, respectively).

Emotional outcomes are conceptualised as a consequence of using a software system, rather than as a design goal in itself. If the software system is appealing, the user may experience enjoyment or even fun (if not, she may experience frustration or anger). For this reason, the emotional outcome is not explicitly included in the present study. However, it remains important and should be included in further studies.

The present paper will see to the following research questions.

Q1: Are ergonomic (EQ) and hedonic quality (HQ) subjectively different quality aspects that can be independently perceived by the users?

Q2: Is the judgement of appeal (APPEAL) formed by combining and weighting EQ and HQ? Which weights are assigned to EQ and HQ?

Q3: The research concerning apparent usability [5;15;22] emphasises user expectations concerning the quality aspects of a software system *before* actually using it. Based on this general idea, we additionally investigated EQ, HQ and APPEAL under both conditions - before and after using a software system. Furthermore, the stability of EQ, HQ and APPEAL will be examined.

METHOD

Participants

Twenty individuals (6 women, 14 men) participated in the study. They were recruited among the employees of Siemens Corporate Technology in Munich. The sample's mean age was 32.5 years (Min 25, Max 57). Computer expertise varied from moderate to high.

Prototypes

Seven software prototypes were designed and implemented (see Figure 2). The participant's task was to switch off a pump in an assumed industry plant. It was a very simple, but realistic task from a domain unknown to the participants. Thereby, unwanted comparison to existing software systems should be reduced. Each prototype allowed the user to accomplish the same task but strongly varied in design and interaction style. The prototypes were designed and implemented by students of visual, industrial and ergonomic design. Multiple design dimensions were varied (e.g. colours, design style). Six out of the seven prototypes had animated parts. The predominant design principle was heterogeneity, i.e. maximising variance. Hence, flaws in the visual and ergonomic design were not corrected.

Semantic differential

A semantic differential (see [9] p. 73) was constructed to measure EQ, HQ and APPEAL. It consists of 23 seven-point scale items with bipolar verbal anchors (see Table 1).

Each scale item was selected beforehand to represent a facet of the quality aspect to be measured (EQ, HQ) and the judgement of appeal (APPEAL) according to the definitions given above. Informal expert reviews were conducted to assure measurement quality of the differential's initial version. The scales were presented in random order.

Pre and Post Usage

All measurements were obtained twice: before and after using the software system. EQ, HQ and APPEAL values obtained *before* actual use are named *expected* (i.e. expected by the user); EQ, HQ and APPEAL values obtained *after* using the system are called *experienced*.

Procedure

The study was carried out in the usability laboratory of the User Interface Design department of Siemens Corporate Technology. Each participant was separately led into the lab by the experimenter. After a short instruction, the experimenter showed the first prototype to the participant for approximately one minute. The prototypes were presented in random order. The participant was then asked to make an initial assessment with the semantic differential before actually using the prototype (*expected* values).

At a signal of the experimenter, the participant had to select the running pump on the screen with the mouse. Depending on the prototype, a question appeared asking whether the participant really wanted to switch off the pump. After a confirmation and a safety check the pump could be turned off. The participant then had to wait until the pump stopped. The end of the task was reached when the participant indicated that s/he was convinced that the pump came to a halt. The whole interaction per prototype took about two minutes.

After finishing the task, the participant was requested to revise his/her initial assessment (*experienced* values). This procedure was repeated for all seven prototypes.

Questionnaires concerning demographics, computer expertise and computer anxiety were handed out. S/He was then shortly debriefed and led out of the laboratory. A session took about an hour.

Tab. 1: Bipolar verbal scale anchors per factor

Scale Item	Anchors	
EQ 1	comprehensible	incomprehensible
EQ 2	supporting	obstructing
EQ 3	simple	complex
EQ 4	predictable	unpredictable
EQ 5	clear	confusing
EQ 6	trustworthy	shady
EQ 7	controllable	uncontrollable
EQ 8	familiar	strange
HQ 1	interesting	boring
HQ 2	costly	cheap
HQ 3	exciting	dull
HQ 4	exclusive	standard
HQ 5	impressive	nondescript
HQ 6	original	ordinary
HQ 7	innovative	conservative
APPEAL 1	pleasant	unpleasant
APPEAL 2	good	bad
APPEAL 3	aesthetic	unaesthetic
APPEAL 4	inviting	rejecting
APPEAL 5	attractive	unattractive
APPEAL 6	sympathetic	unsympathetic
APPEAL 7	motivating	discouraging
APPEAL 8	desirable	undesirable

Note: verbal anchors of the differential are originally in German

RESULTS

Factorial validity and scale reliability of the semantic differential (Q1)

The first question to be answered is whether ergonomic (EQ) and hedonic quality (HQ) are independently perceived by the participants (Q1).

A factor analysis (Principal Components, Varimax rotation) of the *experienced* EQ and HQ items of the semantic differential extracted two factors with an Eigenvalue higher than 1 (see Table 2). Together both factors explain approx. 68% percent of the total variance. This analysis confirms the initial version of the differential. It shows high consistency with the theoretically assumed factors EQ and HQ. Furthermore, the successfully attained simple structure (e.g. [21]) by applying Varimax rotation shows that EQ and HQ are perceived as independent quality concepts by the participants. For the remainder of the paper, a mean ergonomic and hedonic quality value is computed from the respective single scale item values.

Tab. 2: Factorial validity of experienced EQ and HQ

Scale Item	Principal Components with Varimax	
	Factor 1	Factor 2
EQ 1	.783	
EQ 2	.816	
EQ 3	.715	-.233
EQ 4	.880	
EQ 5	.893	
EQ 6	.824	
EQ 7	.885	
EQ 8	.679	
HQ 1		.805
HQ 2		.731
HQ 3		.722
HQ 4		.854
HQ 5		.892
HQ 6	-.226	.818
HQ 7	-.218	.831
Eigenvalue	*5.44*	*4.76*
% Variance explained	*36.27*	*31.72*

Note: N=140 (20 participants x 7 prototypes), EQ: ergonomic quality, HQ: hedonic quality, factor loadings < .20 are omitted

An analysis of the *expected* EQ and HQ items revealed similar results.

A factor analysis (Principal Components) of the *experienced* APPEAL items of the semantic differential extracted only one factor with an Eigenvalue higher than 1 (see Table 3).

This confirms the initial selection of items for the APPEAL scale, its univariate character and internal consistency (see scale reliability). For the remainder of the paper a mean judgement of appeal is computed from the values of the single scale items.

An analysis of the *expected* APPEAL items revealed similar results.

Tab. 3: Factorial validity of experienced APPEAL

Scale Item	Principal Components
	Factor 1
APPEAL 1	.868
APPEAL 2	.794
APPEAL 3	.683
APPEAL 4	.878
APPEAL 5	.865
APPEAL 6	.903
APPEAL 7	.800
APPEAL 8	.775
Eigenvalue	*5.43*
% Variance explained	*68.20*

Note: N=140 (20 participants x 7 prototypes), APPEAL: judgement of appeal

Table 4 summarises the characteristics of the computed scales.

Tab. 4: Scale characteristics

Scale	Cronbach's Alpha	Mean	Stand. Dev.	Min	Max
expected (pre usage)					
EQ	.93	0.87	1.37	-2.88	3.00
HQ	.93	-0.19	1.47	-3.00	2.75
APPEAL	.95	0.38	1.40	-3.00	2.86
experienced(post usage)					
EQ	.93	0.46	1.50	-2.75	3.00
HQ	.92	0.02	1.40	-2.88	3.00
APPEAL	.93	0.31	1.37	-3.00	3.00

Note: N=140 (20 participants x 7 prototypes), EQ: ergonomic quality, HQ: hedonic quality, APPEAL: judgement of appeal

Predicting the participant's judgement of appeal (Q2)

The judgement of appeal is conceptualised as being formed on the basis of the individual's perceptions of EQ and HQ. To check this assumption two regression analysis were performed in order to predict expected APPEAL from expected EQ and HQ and experienced APPEAL from experienced EQ and HQ (see Table 5).

In both analyses, the variables succeed in predicting APPEAL: EQ and HQ almost equally contribute to the judgement of appeal. No interaction of EQ and HQ was found. This pattern implies an averaging model, with the final judgement of appeal depending on both, the perception of ergonomic *and* hedonic quality.

Tab. 5: Regression analysis of EQ and HQ on APPEAL (expected and experienced)

Criterion	Adjusted R^2	Predictors	Beta	Std. Error	Sig.
expected (pre usage)					
APPEAL	.74***	EQ***	.51	.05	.000
		HQ***	.62	.04	.000
APPEAL	.74***	EQ***	.40	.14	.003
		HQ***	.53	.10	.000
		EQ x HQ	.16	.03	n.s.
experienced (post usage)					
APPEAL	.67***	EQ***	.64	.05	.000
		HQ***	.58	.05	.000
APPEAL	.67***	EQ***	.64	.12	.000
		HQ***	.72	.14	.000
		EQ x HQ	-.09	.03	n.s.

Note: N=140 (20 participants x 7 prototypes), EQ: ergonomic quality, HQ: hedonic quality, APPEAL: judgement of appeal, ***p<.000

Expected vs. Experienced EQ, HQ and APPEAL (Q3)

Figure 3 shows expected and experienced EQ, HQ and APPEAL (i.e. pre and post usage, see also Table 4 for scale characteristics).

A repeated measurements analysis of variance with the within-subject factors "time of measurement" (pre usage, post usage) and "type of measurement" (EQ, HQ, APPEAL) revealed a highly significant main effect of "type" (F=11.30, d.f.=2, p=.000) and a highly significant "type" and "time" interaction (F=19.92, d.f.=2, p=.000). No main effect was found for "time" (F=1.71, d.f.=1, n.s.).

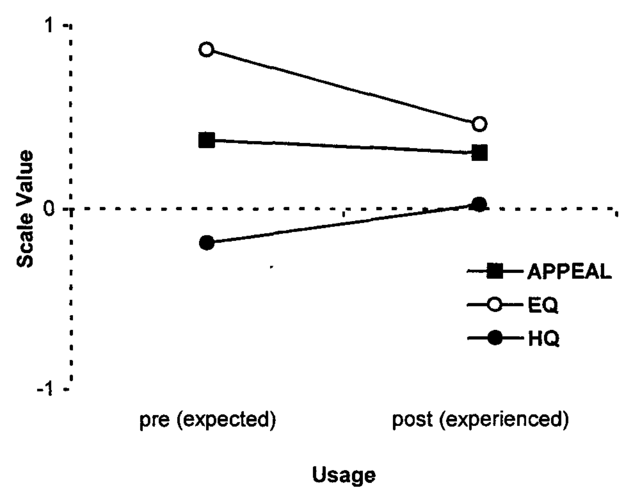


Fig. 2: Expected and experienced EQ, HQ and APPEAL

Single paired sample t-tests for each measure showed a highly significant increase of HQ (diff=.21 scale units, t=3.66, p=.000) and a significant decrease of EQ (diff.=-.41 scale units, t=-3.94, p=.000) after using the prototypes. APPEAL remained stable (diff.=-0.07 scale units, t=-0.87, n.s.).

Values of expected and experienced EQ, HQ and APPEAL were all significantly correlated (expected EQ - experienced EQ: r=.64, N=139, p=.000; expected HQ - experienced HQ: r=.89, N=139, p=.000; expected APPEAL - experienced APPEAL: r=.78, N=139, p=.000).

DISCUSSION

The results of the factor analysis show that the two assumed quality aspects EQ and HQ can be perceived consistently and independently by users. This is true for both expected and experienced values. In other words, users are able to distinguish task-related aspects from non task-related aspects.

Carroll and Thomas' [6] argument that ease of use implies simplicity, which in turn is partly incompatible with fun, is apparent in the data (see Table 2). The items EQ 3 "simple - complex", HQ 6 "original - ordinary" and HQ 7 "innovative - conservative" are negatively correlated with the opposing aspect (factor). This points to the fact that from a software design perspective it might be impossible to have both aspects maximised. In other words, making it innovative (increase HQ) may result in increased perceived complexity (decreased EQ) and making it simple (increase EQ) may lead to a boring software system (decreased HQ). These findings are consistent with the idea that curiosity can by stimulated by providing an optimal level of informational complexity (e.g. [3]). If the software system is either too simple or too complex, the user will feel bored or overloaded, respectively.

The judgement of appeal (APPEAL) demonstrates to be a highly consistent construct. Even such different dimensions as motivation and aesthetics have the evaluational aspect in common, which is taken into account by the users. The regression analysis of EQ and HQ on APPEAL showed that both quality aspects play an almost equal role in forming the judgement of appeal. Together with the missing interaction, this points at a simple averaging model with EQ and HQ compensating each other. This finding is consistent with the averaging model of Information Integration Theory [2]. This theory proposes that different pieces of information in a multi-attribute judgement are integrated by an averaging process. Taking APPEAL as a multi-attribute judgement and EQ and HQ as internally generated pieces of information (based on expectations or experiences with the software system), the observed averaging is simply a consequence of a cognitive process.

From these results we may cautiously conclude that the hypothesised model for the appeal of software is valuable for guiding future research.

No substantial differences can be found when comparing the way the judgement of appeal is formed before and after using the software. It strikes the eye that the expected APPEAL seems to be more based on HQ, whereas the experienced APPEAL seems to be based on both HQ and EQ. This may be due to the fact that the HQ quality aspect is much easier to perceive without using the software than EQ.

A comparison of expected and experienced EQ and HQ shows that in general HQ increases whereas EQ decreases. The suggestion of this result is two-fold: First, it shows that HQ and EQ are based on more than simply the appearance of the software system. Both can be influenced by the experiences individuals have with the system. Second, an increase in one quality aspect may always lead to a decrease in the other. This may again be due to the partly incompatibility of HQ and EQ already discussed above.

The fact that APPEAL remains stable might again be due to the averaging process where the decrease in EQ is compensated for by the increase in HQ.

Looking at the correlations of expected with experienced values we find relatively high correlations. This may be a consequence of the relatively short interaction time in our study (see the following section for a discussion).

Limitations

In the following some possible limitations of the present study will be discussed:

Stimulus dependency: The observed effects may depend on the stimulus material provided (the prototypes), hence generalisation may be limited. To reduce this problem beforehand, the prototypes were designed in a way to elicit a wide variety of positive and negative reactions. The scale characteristics (see Table 4) lend support to the notion that this strategy succeeded: minimum and maximum values show that the whole scale was used, the overall means tend to be around the theoretical mean of the scale and the standard deviations for each scale are similar. Nevertheless, there remains the possibility that the presented results can not be generalised to other stimuli. Future studies should address this problem.

Validation of the semantic differential: Internal consistency and factorial validity are important indicators of the reliability of the semantic differential. However, whether the scales really measure the hypothesised quality aspects and the user's evaluation (i.e. the validity of the scales) remains unanswered. Future studies should attempt to validate the scales. For EQ this can be easily achieved by correlating available usability questionnaires. For HQ and APPEAL other ways have to be found.

Lack of judgmental context: The procedure employed in the present study gives no explicit information about the context for the judgement of appeal. The question on hand is whether individuals are able to form a judgement about a software system without thinking of a context, i.e. where to use it. If a context-free judgement is not possible, the participants might simply have induced their own contexts. For example, a group of participants may have thought of how it would be like to use a software system akin to presented prototypes in their daily work. These participants may base

APPEAL more strongly on EQ than on HQ. On the other hand, a second group of participants may have thought that there is a low probability of using a software system akin to prototypes in the future or in their daily work. These participants may emphasise the importance of HQ. The overall result of the two possible strategies would be similar to the observed averaging model with both HQ and EQ being of equal importance.

Future studies should seek to control the context by providing an explicit situation where to apply the software system (e.g. at work). This would change the context-free judgement of appeal to an evaluation that takes the software system's intended "context of use" [4] into account.

Short interaction time: The direction of the expected vs. experienced effect for HQ and EQ (increase of HQ vs. decrease of EQ) seems to be counter-intuitive. We expected an increase in EQ stemming from a sense of reduced complexity and increased familiarity induced by getting to know the system and a decrease in HQ stemming from reduced novelty. The actual results may point at a limitation of the study: the short interaction time. The interaction time of about two minutes might have been too short to change much in the participants' perception and evaluation. The same explanation holds for the relatively strong pre and post usage correlation of HQ, EQ and APPEAL. Further studies should provide more complex systems and should make a longer interaction time (maybe with several measurements over time) possible. This will help to understand how perceptions and evaluations of users change over time.

Practical implications for the design of appealing software systems

Since hedonic quality is perceived by the users and plays a substantial role in forming their judgement of appeal, it should be explicitly taken into account when designing a software system. Without considering hedonic quality a potential source for increased software quality is neglected.

Due to the partial incompatibility of hedonic and ergonomic quality, software designers should try to find a subtle balance of both quality aspects rather than to independently maximise them. Especially interface designers must identify ways to introduce novelty and surprise with their interfaces (and the behaviour of the software system) without sacrificing to much ergonomic quality (e.g. familiarity). From this perspective the impact of hedonic quality on the appeal of a software system may be the rationale for introducing new interface elements (or even completely new metaphors) and to justify the risk of impaired ergonomic quality.

The averaging model implies that a lack of hedonic quality can be compensated by increased ergonomic quality and vice versa. This means also that good usability can be cancelled out by a lack of hedonic quality. Therefore, the primary strategy should be to have both aspects covered. If this is not possible, the designer may concentrate on maximising only one quality aspect. Which one may depend on the software systems purpose and context of use.

CONCLUSION

Despite the possible limitations of the presented study the results look promising and should stimulate further research. The quantitative research approach presented herein should be complemented by a qualitative approach that seeks to identify specific design factors that are able to stimulate the perception of hedonic quality. The major goal should be to provide a better understanding of what makes a software system appealing to people. Such an understanding may guide software design and thus may have positive effects on the acceptance of software, its creative use and the well-being of the users among us who have to spend a large proportion of their professional lives in front of computer screens.

ACKNOWLEDGMENTS

We like to thank the MediaPlant project group, especially Stefan Hofmann, Alard Weisscher, Jochen Klein and Tobias Komischke for designing and implementing the prototypes used in the present study.

REFERENCES

1. Adams, E. and Sanders, E. An evaluation of the fun factor for the Microsoft EasyBall Mouse, in *Proc. of the 39th Human Factors and Ergonomics Society Annual Meeting* (1995), 311-315.
2. Anderson, N. H. *Foundations of information integration theory*. Academic Press, New York, NY, 1981.
3. Berlyne, D. E. Curiosity and exploration. *Science 153* (1968), 25-33.
4. Bevan, N. and Macleod, M. Usability measurement in context. *Behaviour & Information Technology 13 1&2* (1994), 132-145.
5. Burmester, M., Platz, A., Rudolph, U., and Wild, B. Aesthetic design - just an add on? in *Proc. of the HCI '99* (1999), 671-675.
6. Carroll, J. M. and Thomas, J. C. Fun. *SIGCHI Bulletin 19 3* (1988), 21-24.
7. Csikszentmihalyi, M. *Beyond Boredom and Anxiety*. Jossey-Bass, San Francisco, 1975.
8. Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. *Journal of Applied Psychology 22 14* (1992), 1111-1132.
9. Fishbein, M. and Ajzen, I. *Belief, Attitude, Intention and Behavior*. Addison-Wesley, Reading, MA, 1975.
10. Harrison, A. W. and Rainer, R. K. A general measure of user computing satisfaction. *Computers in Human Behavior 12 1* (1996), 79-92.
11. Hollnagel, E. Keep cool: The value of affective computer interfaces in a rational world. In *Proc. of HCI International '99* (1999), 676-680.
12. Igbaria, M., Schiffman, S. J., and Wieckowski, T. J. The respective roles of perceived usefulness and perceived fun in the acceptance of microcomputer technology.

Behaviour & Information Technology 13 6 (1994), 349-361.

13. ISO. ISO 9241: Ergonomic requirements for office work with visual display terminals. Part 11: Guidance on usability (1996), International Organization for Standardization.

14. Kim, J. and Moon, J. Y. Designing towards emotional usability in customer interfaces - trustworthiness of cyber-banking system interfaces. *Interacting with Computers 10* (1998), 1-29.

15. Kurosu, M. and Kashimura, K.: Apparent usability vs. inherent usability. In *CHI '95 Conference Companion.* (1995), 292-293.

16. Laurel B. *Computers as Theatre.* Addison-Wesley, Reading, MA, 1993.

17. Logan, R. J. Behavioral and emotional usability: Thomson Consumer Electronics, in M. Wiklund (ed.) *Usability in Practice.* Academic Press, Cambridge, MA, 1994.

18. Logan, R. J., Augaitis, S., and Renk, T. Design of simplified television remote controls: a case for behavioral and emotional usability, in *Proc. of the 38th Human Factors and Ergonomics Society Annual Meeting* (1994), 365-369.

19. Malone, T. W. Toward a theory of intrinsically motivating instruction. *Cognitive Science 4* (1981), 333-369.

20. Malone, T. W. Heuristics for designing enjoyable user interfaces: Lessons from computer games, in J. C. Thomas and M. L. Schneider (eds.) *Human Factors in Computer Systems.* Ablex, Norwood, NJ, 1984.

21. Thurstone, L. L. Multiple factor analysis: A development and expansion of vectors of the mind. University of Chicago Press, Chicago, 1947.

22. Tractinsky, N. Aesthetics and Apparent Usability: Empirically Assessing Cultural and Methodological Issues, in *Proc. of the CHI '97* (1997), 115-122.