

The Influence of Aesthetics in Usability Testing: the Case of Dual-Domain Products

Andreas Sonderegger, Andreas Uebelbacher, Manuela Pugliese and Juergen Sauer

University of Fribourg, Department of Psychology
Rue Faucigny 2, 1700 Fribourg
Contact: andreas.sonderegger@unifr.ch
+41 26 300 76 30

ABSTRACT

An experimental study examined whether the effects of aesthetic appeal on usability test outcomes are moderated by usage domain. The aesthetic appeal of a cell phone was experimentally manipulated in both home- and work-based usage domains. The two usage domains were modeled in a usability laboratory. 60 participants completed a series of typical cell phone user tasks. Dependent measures such as performance, perceived usability, and emotion were taken. The results showed that aesthetic appeal had a positive effect on perceived usability but a negative effect on performance. The effects of aesthetic appeal on usability test outcomes were not moderated by usage domain. The results of this study imply that it may be sufficient to test dual-domain products in only one of their usage domains.

Author Keywords

Usability test; Product aesthetics; Usage domain; Perceived usability; User performance

ACM Classification Keywords

H.1.2

INTRODUCTION

Aesthetics in product design

The role of aesthetics in the design of interactive products has not only attracted increasing interest in the product design community, but also in the field of human factors and ergonomics [27,60]. This is reflected in an increasing number of empirical studies that have investigated the influence of aesthetically pleasing designs on outcome variables such as perceived usability, performance and emotion. The findings of these studies showed aesthetics had a positive influence on perceived usability [e.g. 4,23,30,41,43,47,48,49,51]. Although there is some work

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

CHI 2014, April 26 - May 01 2014, Toronto, ON, Canada
Copyright 2014 ACM 978-1-4503-2473-1/14/04...\$15.00.
<http://dx.doi.org/10.1145/2556288.2557419>

that did not report such effects [e.g. 54,56,58], most studies having investigated aesthetics and usability found that users rate usability of an appliance higher when the design is more aesthetically pleasing. Even when appliances had exactly the same inherent usability, users still perceived the more aesthetically pleasing appliance as having a higher usability. To explain this intriguing observation, the halo effect has been proposed as an underlying mechanism. Aesthetics, as a predominant (and mainly visual) feature, influences the perception of other less prominent features of the product such as inherent usability [52]. The "what is beautiful is good" stereotype, known from the field of social psychology [13] also provides explanation to this phenomenon. Recent studies addressing the link between perceived attractiveness and perceived usability claim that a general evaluation of the product ('goodness') might act as moderating factor between perceived attractiveness and perceived usability [16,57]. Other work however put forward different moderators (e.g. affect, [49]) and there is currently no set of influencing factors upon which there is general agreement.

In addition to the relationship between aesthetic appeal and perceived usability, some studies also examined the effect of aesthetics on user performance. The findings are however highly inconsistent. To our knowledge, there are two studies that found a performance decrease when users operated aesthetically pleasing products [2,41], while two other studies revealed an increase in performance [29,44]. Finally, in several other studies aesthetic appeal was unrelated to performance [7,15,49,52]. To explain these conflicting findings, it has been postulated that there may be two opposing effects on performance [44]. Performance may decrease because of the occurrence of a "prolongation of joyful experience"-effect, that is, users enjoy the aesthetically pleasing design of the appliance and therefore wish to increase the time of operating the product, thus resulting in longer task completion times. On the other hand, performance may improve because of an "increased motivation"-effect. Based on Norman's 'positive affect mediation model' [32], it can be assumed that user interaction with an aesthetically appealing device is mediated by positive affect and may have a positive effect on user performance. Furthermore, it may even put the user in a state of "flow" [10], which may result in shorter task

completion times. Currently, it is not clear under what circumstances either effect occurs but contextual factors such as work- or leisure-based usage may have a moderating function. However, no direct empirical comparison between these usage contexts has yet been made to test this assumption.

While perceived usability and, to a somewhat lesser extent, performance were the principal outcome measures in this research area, emotion was measured only in a small number of studies. This is rather astonishing considering the fact that different models and theories of aesthetic design and evaluation regard affect as being a central aspect in the process of aesthetic perception [25,32]. Furthermore, some empirical work in the domain of human factors and ergonomics has indicated that emotions might be an important mediator or moderator for the interplay of design aesthetics and outcome variables of usability tests, such as perceived usability and performance [e.g. 33,49]. Overall, emotion as an outcome measure has gained in importance in the field of consumer ergonomics, which is demonstrated by the introduction of concepts such as fun, pleasure and joy into the research literature [20,32,61]. This indicates that the impact of using technology needs to be measured at multiple levels, including the users' emotional response as an important aspect of user experience, as it is suggested by ISO standards (ISO 9241-210). An increasing number of publications have already addressed the issue of emotion in the design of interactive consumer products [e.g. 5,19]. Concerning the impact of product aesthetics, there have been several studies that measured user emotions. While some studies did not find a relationship between the two concepts [e.g. 46], other work showed that emotions were more positively affected by the operation of more aesthetically appealing products [29,41,47,49]. All this work suggests that emotion needs to be considered as an important aspect of product design and should be measured in usability testing.

When reviewing the research literature on aesthetics, some conceptual confusion surrounding the term aesthetics is noticeable. Apart from aesthetics, there are several other terms that denote exterior properties of a product-like appearance and attractiveness [e.g. 6,18]. These terms are not consistently employed across research fields. For example, some authors employed the term aesthetics when referring to a user's response to the appearance of the product [9,55]. Other authors distinguish between different forms of aesthetics such as expressive and classical aesthetics [24]. Expressive aesthetics is characterized by original and creative design. In contrast, classical aesthetics represents a more traditional form of design with orderliness and clarity as important elements. This also includes the use of harmonious colors and symmetric patterns [31,34]. To avoid conceptual confusion, in the present study the term aesthetics is used to refer to the objective exterior properties of the product, while the term attractiveness is employed when the user's subjective

response to these properties is concerned (i.e. measuring user perception of the product).

Usage domain

The context in which a product is used is expected to influence the way it is operated and perceived [3,28]. While there are many aspects in which the context may differ (e.g. presence of co-users or observers, testing in lab or field), of particular relevance appears to be the so-called usage domain. This refers to a classification of product use according to location (e.g. work, domestic and public domain) or according to the primary activity (e.g. work, leisure). While the work environment is undoubtedly the domain that has most frequently been the subject of ergonomic research, other domains have also benefited from such research efforts, partly fuelled by the increasing complexity of technical systems in non-work domains. Examples for research in the design of products for the domestic domain are central heating systems [40] and washing machines [36]. In the public domain, automated teller machines [52] and ticket vending machines [8] represent examples of products that have been examined. Finally, a portable digital audio player [49] represents products that belong to the leisure domain.

Usability requirements may differ as a function of these domains. In contrast to work environments, non-work domains are more restricted with regard to the possibilities of modifying user behavior [40]. At work, humans are typically not only selected for their competencies in operating a device, but they also receive training, and they are often given feedback about their performance. By using these measures, deficiencies in the usability of a device may be compensated to some extent. This is in contrast to non-work domains, where such measures are usually not available. In a similar vein, user expectations with regard to aesthetics may also be different between the domains concerned. This may be related to the fact that, unlike work environments, users in the leisure domain often own the product they operate. Some of the interactive consumer products owned by users also convey a message of prestige and status (e.g. the brand of a cell phone [11,35]), that is, product features that are unrelated to usability may have an influence on usability variables (e.g. aesthetic design features, price, product value, and life style associated with a product [45]). This may suggest a higher significance of such factors in the leisure domain than at work, where efficiency and performance appear to play a dominant role. Indeed, there is evidence that humans experience work and leisure activities in fundamentally different ways [50]. The former is associated with attributes such as performance, goal orientation and external rewards while for the latter intrinsic satisfaction, relaxation and aesthetic appreciation represent important issues. A study by Rheinberg et al. [38] also pointed out the differences between the two domains. At work, individuals experience lower levels of happiness and satisfaction but, at the same time, higher levels of flow (i.e. being fully immersed in an activity) than during leisure

activities. These differences were attributed to the strong goal orientation of work-related activities.

Some interactive consumer products may be used in several domains. For example, a telephone can be employed for work purposes in an office or at home in a leisure context. These products may be termed dual-domain products. Using the same product in different domains raises the question of whether the influence of the domain is of such significance that it may affect the outcomes of a usability test. Put differently, would it make a difference whether the product was tested in an office environment or at home? To answer this question an empirical study is needed that allows for a direct comparison of the effects of operating such a dual-domain product in each domain by using equivalent usage scenarios.

The present study

Based on our literature review on aesthetics and usage domain, the study addressed a number of research questions. (a) To what extent is the well-established positive effect of aesthetics on perceived usability domain-specific? (b) To what extent can the inconsistent pattern of findings with regard to the impact of aesthetics on performance (there is evidence for positive as well as negative effects) be explained by specific domain characteristics? (c). Is there a need to test dual-domain products separately in each domain?

The research questions are examined by comparing two appliances that are different in their aesthetic appeal but are identical in all technical features, which means that objective (or inherent) usability is the same for both. Each appliance is tested in two usage domains (work vs. leisure). A cell phone was chosen as the product to be tested since for this type of product not only functionality and usability are relevant but also aesthetic properties. Furthermore, the cell phone represents an excellent dual-domain product.

The research questions were addressed in an experimental study that manipulated two independent variables: aesthetic appeal (high vs. low) and usage domain (work vs. leisure). Using a device that was manipulated in terms of its aesthetic appeal, users carried out a number of typical tasks of cell phone usage. Half of the users completed these tasks in an office context while the other half did so in a leisure context. Perceived usability and performance were taken as primary measures in the experiment, due to their prominence in the literature review. Finally, the emotional response was measured.

The following predictions were made: (a) The aesthetically appealing appliance will result in higher perceived usability and more positive emotion than the aesthetically less pleasing one. (b) This effect will be stronger in the leisure domain than in the work domain, resulting in a significant interaction between the two independent factors. (c) Aesthetically appealing appliances will lead to increased performance in the work domain (increased motivation) but will decrease performance in the leisure domain

(“prolongation of joyful experience”-effect) than aesthetically less pleasing appliances.

METHOD

Participants

Sixty participants (68.3% female) took part in the experiment, with their ages ranging from 19 to 43 years ($M = 25.9$ yrs). They were students and research assistants. All participants were regular users of a cell phone but none of them owned the specific device to be used in the experiment or had employed it previously. They were not paid for their participation but six book vouchers of CHF 20 each (approximately \$ 20) were given out to participants in a draw.

Experimental design

A 2×2 between-participants design was used, with aesthetic appeal of the product (high vs. low) and usage domain (work vs. leisure) as independent variables. Participants were randomly assigned to one of the four experimental conditions. Half of the participants used their respective phones in a simulated work context whereas the other half carried out the tasks in a simulated leisure context. Usage domain was manipulated by using different set-ups of the laboratory environment and different task instructions. The task instruction either described a leisure or a work domain (c.f. experimental task section below). Both groups were tested in a usability laboratory. In the work set-up, a personal computer, a desk and stationery were placed in the lab. In the leisure set-up, none of these items were present, but the lab was equipped with a TV-set, a table, a sofa, pictures, books and a plant, instead.

Measures and instruments

Perceived usability

A German version of the Post-Study System Usability Questionnaire (PSSUQ [26]) was administered. This translated version has already been used in several experiments [e.g. 42,44]. The PSSUQ was originally developed for the evaluation of computer software. It consists of 19 items which are assigned to three subscales: system usefulness (8 items; e.g. ‘It was easy to learn to use this system.’), information quality (7 items; e.g. ‘The information provided by the system was clear.’), interface quality (3 items; e.g. ‘I liked using the interface of this system.’), and an additional item that is not assigned to any of the subscales (‘Overall, I am satisfied with this system.’). For the purpose of the present study, the term ‘system’ was replaced by ‘cell phone’ to make the items sound more relevant to the intended application area. The original version of the PSSUQ has very good psychometric properties (Cronbach’s $\alpha = 0.97$ for the overall scale and for the three subscales: system usefulness: $\alpha = 0.96$, information quality: $\alpha = 0.91$, interface quality: $\alpha = 0.91$ [26]). In the present study, the translated and modified scale had the following alpha coefficients: overall scale ($\alpha =$

0.93), system usefulness ($\alpha = 0.92$), information quality ($\alpha = 0.88$), and interface quality ($\alpha = 0.70$).

User performance

User performance was determined by measuring three parameters. (a) Task completion rate (%) was measured by calculating the percentage of users that successfully completed each task. (b) Task completion time (s) referred to the time users needed to accomplish each task. (c) Efficiency of user-product interaction refers to the number of commands that users entered to complete each task.

User emotion

A well-established questionnaire, PANAS (Positive Affect and Negative Affect Schedule), was used for measuring user emotion [59]. A German translation of this instrument is also available [22]. The instrument comprises 20 items, describing different positive and negative affective states with 10 items, respectively (e.g. active, excited, irritated, anxious). The intensity of each affective state is rated on a 5-point Likert scale (very slightly or not at all, a little, moderately, quite a bit, extremely). The psychometric properties of the German version of the PANAS are good (Cronbach's α was .84 for NA and .85 for PA [22]).

Materials

A cell phone from Sony Ericsson (c702i) was chosen for the study. The model was rather new on the market when the study was carried out, ensuring that the appliance was likely to be unknown to prospective participants. The exterior surface of the cell phone was manipulated to create an aesthetically appealing and an aesthetically unappealing version of the same appliance (see Fig. 1). Based on previous work on aesthetics [24,31,34,37], the manipulation of the unappealing device was mainly achieved by using asymmetrical and unclear patterns and less appealing color combinations. The color of the plastic casing of the cell phone was changed to the color combination yellow, green and black for the aesthetically unappealing phone and into black for the appealing one. In both cases, spray guns were used for changing the color. The manipulation of aesthetics was limited to the exterior surface of the appliance to ensure that the objective usability of the appliance would not be affected in any way (e.g. changing menu colors might be problematic since it may not only affect attractiveness but also readability).

Experimental tasks

The participants were asked to complete five tasks with the cell phones provided. All of them represented typical tasks for using this appliance. The wording of the tasks differed slightly as a function of the experimental condition of usage domain. The following five tasks were given to participants:



Fig 1. Two versions of the cell phone: (a) aesthetically appealing and (b) aesthetically unappealing

- (a) Make a phone call to ...
- (b) Send a text message to ... saying that you are going to be late for the meeting.
- (c) Change the date for a meeting with ... in the calendar of your mobile.
- (d) You received an email from ..., containing his new phone number. Please save the phone number in the phone book of your mobile.
- (e) Please save the phone number of ... as a short cut in the phone book of your mobile.

In the work domain, "Professor Smith from your university department" was added to the task instruction (i.e. the complete task instruction read 'Make a phone call to Professor Smith from your university department!'). In the leisure domain, "your friend Patricia" was added.

Procedure

After welcoming the participants in the usability laboratory of the University of Fribourg, they were briefly informed about the purpose of the experiment. They were told that they would take part in a usability evaluation of a cell phone. Participants were asked whether they had had any previous experience with the cell phone to be used in the present study. If this had been the case, the participant concerned would have been excluded from the experiment. Participants were then told that they would need to complete five tasks with a cell phone which would enable them to provide feedback about the usability of the appliance. It was emphasized that the usability test aimed to evaluate the cell phone and not the technical competence of the user. Participants were also told that the experiment would take about 40 min to complete.

Participants were randomly assigned to one of the four experimental conditions.

After having been informed about the purpose and the procedure of the experiment, participants completed the PANAS to obtain a pre-test measure of emotion. The cell phone was then presented to the participants for the first time, permitting them to carry out a visual inspection of the appliance. Following the presentation of the appliance to be tested, participants were asked to complete two one-item scales measuring perceived usability and perceived attractiveness.

To strengthen the effectiveness of manipulating the usage domain, prior to receiving instructions, participants were requested to immerse mentally into their respective usage domain. This was achieved by asking them to think about some recent events at work or during their leisure time, respectively. After having been in this domain-specific immersion for about 30 s, participants received the instructions according to their experimental condition. (a) For the work domain, it read: "You have just started your new job as a research assistant at this university. You are looking forward to your new job and to the possibility of working independently. You are still on your probation period and want to do an excellent job to make a good impression with your boss. Since you have a lot of things to organize in your job, your employer has supplied you with this cell phone (investigator shows it to participant)". (b) For the leisure domain, it read: "You have just finished your last university class, with the vacation beginning soon. Now you can enjoy a wonderful evening with your friends, have a meal in a restaurant first before going to the movies. You usually communicate with your friends using a cell phone. You have just bought this cell phone (investigator shows it to participant) because your old one is broken."

The instruction was read out to participants as well as provided to them in a written version. Participants then completed the five tasks, which were presented to them orally by the investigator as well as in written form. After task completion, participants filled out the PANAS for the second time, followed by the PSSUQ. At the end of the experimental session, participants were debriefed and given the opportunity to provide feedback to the investigator about the testing procedure.

Pilot study, manipulation check and data analysis

In a pilot study comprising 10 participants (age: 23 - 65 yrs), the aesthetic appeal of the two appliances was evaluated by asking the participants to rate the attractiveness of both designs on a 10-point Likert scale (aesthetically appealing vs. aesthetically not appealing). Please note that the participants of the pilot study did not take part in the main study. The aesthetically pleasing design was clearly perceived as more attractive than the aesthetically less pleasing alternative ($M = 7.0$, $SD = 1.15$ vs. $M = 2.0$, $SD = 0.81$; $t = 12.6$, $df = 9$, $p < .001$).

In the main study, a manipulation check was used to examine whether the results of the pilot study could be replicated in a between-subjects design, this time using a 7-point Likert scale to make the scale consistent with the PSSUQ response format employed in the study. The findings confirmed that there was a significant difference in the aesthetic appeal of the two devices ($M_{\text{aesthetic}} = 3.04$, $M_{\text{unaesthetic}} = 2.33$; $t = 2.10$, $df = 53$, $p < .05$).

To check whether the participants assigned to the work domain condition experienced the lab setting differently from the participants in the leisure domain condition, a 100-mm visual analogue scale was used to measure the effectiveness of this experimental manipulation. The item was worded "Please indicate whether you experienced the situation in the laboratory as work-like or leisure-like.", with the labels at each end of the scale being: work-related and leisure-related. It was confirmed that the two usage domains were perceived differently in the expected direction ($M_{\text{work}} = 52.0$, $SD = 28.7$; $M_{\text{leisure}} = 26.6$, $SD = 17.4$; $t = 4.12$, $df = 58$, $p < .001$).

A two-factorial analysis of variance was carried out on all variables. For emotion, a baseline measure (taken prior to task completion) was used as a covariate. For all analyses, the alpha level was set to 5%.

RESULTS

Perceived usability

The analysis revealed a main effect of aesthetic appeal in the predicted direction, with participants rating the usability of the aesthetically pleasing phone higher than the unpleasing one (see Tab. 1). This difference was statistically significant ($F = 5.39$, $df = 1, 56$, $p < .05$; $\eta^2_{\text{partial}} = .081$). Interestingly, the positive effect of aesthetic appeal was equally large for the work domain as it was for the leisure domain, which was confirmed by the absence of a significant interaction between the two independent factors ($F < 1$). There was also no main effect of context ($F < 1$).

User performance

Task completion rate. The data of this parameter showed that most participants (82.0%) successfully completed the experimental tasks. The means for this parameter may be found in Table 1. Examining the effects of the independent variable revealed no significant effects (aesthetics: $F < 1$; usage context: $F = 1.45$, $df = 1, 56$, $p > .05$; $\eta^2_{\text{partial}} = .025$; interaction: $F < 1$).

Task completion time. The results show that this performance parameter was not affected by any of the independent variables (see Table 1). This was confirmed by the results of the statistical analysis, which all revealed the absence of any significant effects: aesthetics: $F = 1.94$, $df = 1, 55$, $p > .05$; $\eta^2_{\text{partial}} = .034$; usage context: $F < 1$; interaction: $F < 1$.

Efficiency of user-product interaction. This efficiency measure refers to the number of commands entered by the

user to complete the tasks. The results show that task completion was less efficient when an aesthetically pleasing phone was used than an aesthetically displeasing one (see Table 1). This difference was statistically significant ($F = 4.74$, $df = 1, 51$, $p < .05$; $\eta^2_{\text{partial}} = .085$). No difference between usage contexts was observed ($F = 1.31$, $df = 1, 51$, $p > .05$; $\eta^2_{\text{partial}} = .025$) and no interaction emerged ($F < 1$).

Emotion

The data of positive and negative emotion were analyzed separately, using a two-factorial model of analysis of covariance with aesthetic appeal and usage domain as the independent factors and the baseline measure of emotion (i.e. taken prior to task completion) as covariate. A descriptive analysis showed that the score was higher for positive affect ($M = 3.23$) than for negative affect ($M = 1.25$) at both measurement points. The analysis also revealed that operating the cell phone increased positive affect ($M_{\text{pre}} = 3.17$, $M_{\text{post}} = 3.29$) and reduced negative affect ($M_{\text{pre}} = 1.29$, $M_{\text{post}} = 1.20$). Both changes were found to be significant (positive affect: $F = 6.60$, $df = 1, 55$, $p < .05$; $\eta^2_{\text{partial}} = .105$; negative affect: $F = 6.16$, $df = 1, 55$, $p < .05$; $\eta^2_{\text{partial}} = .099$). Aesthetic appeal and usage domain did not show any significant main effects and no interaction, neither for positive nor for negative affect. The covariate 'baseline measure' was highly related to measures of emotion taken after task completion (positive affect: $F = 87.34$, $df = 1, 55$, $p < .001$; $\eta^2_{\text{partial}} = .61$; negative affect: $F = 57.10$, $df = 1, 55$, $p < .001$; $\eta^2_{\text{partial}} = .51$). The data for positive and negative affect are presented in Table 1 (only the post-test data are presented since the pre-test data showed the same pattern). An additional analysis was carried out, examining each of the 20 emotions measured

by PANAS separately. However, none of them was found to be influenced by aesthetic appeal or usage domain.

Correlational analysis of data

In order to facilitate interpretation and comparison of outcomes of usability evaluations, Hornbaek and Law [20] suggest that usability studies should report correlations between measures collected. Analysis of data indicate correlations of perceived usability with attractiveness ($r(54) = .57$, $p < .001$), task completion time ($r(54) = -.47$, $p < .001$) and task efficiency ($r(54) = .53$, $p < .001$). Furthermore, negative emotions decreased with a higher task completion rate ($r(54) = -.29$, $p = .036$) and increased task efficiency ($r(54) = -.31$, $p = .023$). No other correlations reached significance.

DISCUSSION

The present study has revealed three major findings. First, the results have confirmed the well-established effect of aesthetics on perceived usability. Second, with regard to performance-related effects, there was evidence for a negative impact of the aesthetically pleasing design on one performance measure. Third, there was a general absence of domain-specific effects, with little difference emerging between work and leisure domains.

The present study has again confirmed the positive effects of an aesthetically appealing design on user perception of product usability. Given the considerable number of previous studies that have reported similar effects [e.g. 51,52], the finding has to be considered very robust. The robustness of this effect is supported by the absence of an influence of usage domain, indicating that even in performance-oriented domains such as work an influence of aesthetics may be observed.

	High aesthetic appeal Means (SD)	Low aesthetic appeal Means (SD)	Overall Means (SD)
Perceived usability (1 – 7)*			
Work	3.71 (1.36)	2.93 (1.00)	3.32 (1.24)
Leisure	3.42 (1.04)	2.96 (0.61)	3.19 (0.87)
Task completion rate (%)			
Work	80.0 (22.7)	77.3 (22.5)	78.7 (22.2)
Leisure	89.3 (12.8)	81.3 (25.6)	85.3 (20.3)
Task completion time (s)			
Work	405.1 (219.9)	343.2 (163.0)	374.1 (192.8)
Leisure	400.1 (221.1)	324.2 (138.7)	363.4 (186.7)
Efficiency index*			
Work	281.5 (106.5)	245.3 (104.5)	264.1 (105.1)
Leisure	272.1 (100.8)	195.5 (63.4)	236.5 (92.6)
Emotion (post test): positive (1 – 5)			
Work	3.38 (.50)	3.33 (.63)	3.36 (.56)
Leisure	3.27 (.56)	3.15 (.55)	3.22 (.55)
Emotion (post test): negative (1 – 5)			
Work	1.14 (.18)	1.19 (.19)	1.17 (.18)
Leisure	1.26 (.35)	1.21 (.21)	1.23 (.28)

Table 1. Perceived usability, performance parameters and emotion as a function of aesthetic appeal and usage domain (significant effects of aesthetics are marked with an asterisk)

This suggests that the occurrence of the “what is beautiful is good” - stereotype [13,16,57] can also be extended to the work domain. This undue influence of aesthetics on outcome variables other than those measuring aesthetic appeal has important repercussions for the evaluation of product usability. It may lead to an overestimate of perceived usability if the product enjoys an aesthetically pleasing design. This may be undesirable for a designer who wishes to distinguish between inherent usability and aesthetic properties even if the latter is clearly a part of the wider concept of user experience [cf. 14].

While the pattern is highly consistent with regard to perceived usability, this is not the case for performance-related effects. The inconsistency of results is not only related to the question of whether aesthetically pleasing designs cause an increase or decrease in performance but also to what kinds of performance measures are affected. The present work adds another study to the small body of research [2,41], which found a decrease in performance for aesthetically pleasing designs compared to displeasing ones, amounting now to three studies [2,41 and the present one]. However, the decrease in performance was only observed for one of the measures while for the other two measures no such effect was recorded. This suggests that this study not only added to the body of research that found a decrease but also provided support in a way for the other body of research, which found no such effect of aesthetics. This indicates that inconsistencies found in the literature could even be illustrated within the present study. Overall, the inconsistencies in the results across studies suggest the presence of several moderating or mediating factors, which have not yet been identified. However, the findings of this study indicate that usage domain does not represent an important moderator or mediator variable for the link between aesthetics and user performance. However, other variables may still be entertained as possible moderators, such as gender [53] and user age [44]. Further empirical research, ideally in the form of field studies conducted in real work and leisure environments, is therefore needed to determine whether these factors actually have a moderating function.

Analyzing the effects of the independent variables on emotion revealed that operating the appliance was considered to be a positive experience by participants, as shown by reported increases in positive affect as well as decreases in negative affect after product usage. However, the interpretation of this result is difficult because no control condition with a different task was used in this study. The change in emotion after task completion was not moderated by usage domain or aesthetic appeal, as demonstrated by the non-significance of all interactions. This raises two points. First, the leisure domain may not amplify the emotional experience more strongly than a work domain, even if there were deliberations by some authors that happiness and satisfaction are more closely associated with leisure [e.g. 38]. Second, no impact of

aesthetic appeal on emotion was observed in the present study, though some previous work on the impact of aesthetics suggested that aesthetically pleasing designs also resulted in more positive emotions [33,41,49]. However, emotions as a consequence of product use may be influenced by several factors, such as inherent usability, ease of use, social prestige and aesthetics [12,39,49,54]. These factors appear to differ in the size of their impact. For example, Thüring and Mahlke [49] showed that both, inherent usability and visual aesthetics, have an influence on user emotions but the effect of aesthetics on user emotion was rather small compared to inherent usability. In contrast, other work showed that user emotions were solely influenced by inherent usability and not by visual aesthetics [54]. Overall, this suggests that the relationship between aesthetics and emotion may be of a rather complex nature, in that sense displaying some similarity to the complex relationship between aesthetics and performance.

Although there was some evidence in previous work that suggested differences in perception and behavior between the two usage domains work and leisure, no such differences were found in the present study. The absence of such a difference may be due to several factors. Generally, the distinction between work and leisure has been considered a difficult one [1,17]. This is because different users perceive the same domain in different ways. Cautious evidence in support of this interpretation was found in the present study, indicated by the considerable variance in user ratings of the scale measuring the user’s perception of the lab environment as work-like or leisure-like. This suggests that not all users perceived the work domain as work-related and the leisure domain as leisure-related. Furthermore, it emerged that participants rated the work domain just above the middle of the 100-point scale ($M = 52.0$) while the leisure domain received a rating much closer to the expected scale end ($M = 26.6$). This suggests that the leisure domain was experienced more intensely than the work domain. Participants might have found it easier to immerse into the leisure context because of the closer experience with the situations described in the task instructions (all of them had already sent a text message to a friend to cancel a meeting but few had already sent out such a message to a university professor). Overall, it shows that the impact of usage domain may be smaller than expected on the basis of the small body of literature. With the current study making the first direct empirical comparison between work and leisure domains, we believe that there is a need for more empirical work, making comparisons between the two domains in similar ways, in order to determine with more confidence the influence of the usage domain in usability evaluation.

CONCLUSION

The present study has revealed three major implications for designers. First, the influence of aesthetics on perceived usability is highly robust, extending to several usage domains. It implies that even for systems mainly used in

work settings, it is worth trying hard to make interactive products aesthetically pleasing if one wishes to obtain high usability ratings. Conversely, designers need to be aware that for work systems too, there may be a risk to overestimate the usability of a product if relying only on subjective measures of a highly appealing product. Second, it may be sufficient to test dual-domain products only in one domain since there were no main effects of domain and no interactions with aesthetics. However, designers need to ensure that the tasks chosen are not only meaningful to the domain to be tested but also relevant to the domain not selected for testing (i.e. the tasks need to be representative for both domains). Third, findings of this study indicate that efficiency of task completion may decrease when a highly aesthetical product is used. Although this effect is in contrast to findings of other studies, it indicates that design aesthetics may not only influence subjective measures of usability tests but may also have a significant impact on (objective) measures of user behavior.

ACKNOWLEDGEMENTS

The authors are very grateful to the Swiss National Science Foundation for their financial support of the study (research grant No. 100014/122490).

REFERENCES

1. Beatty JE, Torbert WR (2003) The false duality of work and leisure. *J Manage Inq* 12 (3): 239–252
2. Ben-Bassat T, Meyer J, Tractinsky N (2006) Economic and subjective measures of the perceived value of aesthetics and usability. *ACM T Comput-Hum Int* 13 (2): 210–234
3. Bevan N (2006) Practical issues in usability measurement. *Interact* 13 (6): 42–43
4. Brady L, Phillips C (2003) Aesthetics and usability: a look at color and balance. *Usability News* 5(1).
5. Brave S, Nass C (2003) Emotion in human-computer interaction. In: Sears A., Jacko JA (eds). *The human computer interaction handbook: fundamentals, evolving technologies and emerging applications*. Lawrence Erlbaum Associates, New Jersey, pp. 81–96
6. Chang HC, Lai HH, Chang YM (2007) A measurement scale for evaluating the attractiveness of a passenger car form aimed at young consumers. *Int J Ind Ergonom* 37 (1): 21–30
7. Chawda B, Craft B, Cairns P, Rüger S, Heesch D (2005). Do “ Attractive Things Work Better ”? An Exploration of Search Tool Visualisations. In N. McKinnon, L, Bertelsen, O B (eds), *Proceedings of 19th British HCI Group Annual Conference (HCI2005)*. Edinburgh, UK, pp. 2–5
8. Connell IW, Blandford AE, Green TRG (2004) CASSM and cognitive walkthrough: usability issues with ticket vending machines. *Behav Inform Technol* 23 (5): 307–320
9. Crilly N, Moultrie J, Clarkson PJ (2004) Seeing things: consumer response to the visual domain in product design. *Des Stud* 25 (6): 547–577
10. Csikszentmihályi M (1997) *Finding flow: The psychology of engagement with everyday life*. Basic Books, New York
11. Dedeoglu AO (2004) The symbolic use of mobile telephone among Turkish consumers. *J Euro-Mark*, 13: 143–162
12. Desmet PMA, Hekkert P (2002) The basis of product emotions. In: Green WS, Jordan PW (eds) *Pleasure with Products: Beyond Usability*. Taylor & Francis, London, pp. 60–68
13. Dion K, Berscheid E, Walster E (1972) What is beautiful is good. *J Exp Soc Psychol*, 24 (3): 285–290
14. Garrett JJ (2002) *Elements of user experience*. New Riders Press, Indianapolis
15. Hartmann J, Sutcliffe A, de Angeli A (2007). Investigating Attractiveness in Web User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, San Jose, April 28–May 3. San Jose, USA: ACM New York, USA, pp. 387–396
16. Hassenzahl M, Monk A (2010) The Inference of Perceived Usability From Beauty. *Hum-Comput Interact* 25(3): 235 - 260
17. Haworth JT, Veal, AJ (2004) *Work and leisure*. Routledge, London
18. Hekkert P, Snelders D, van Wieringen PCW (2003) Most advanced, yet acceptable: typicality and novelty as joint predictors of aesthetic preference. *Brit J Psychol*, 94: 111–124
19. Helander MG, Khalid HM (2006) Affective and pleasurable design. In: Salvendy G (ed) *Handbook of Human Factors and Ergonomics*. Wiley Interscience, New York, pp. 543–572
20. Hornbæk K, Law EL-C (2007) Meta-analysis of correlations among usability measures. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*, San Jose, USA. ACM: New York, pp. 617–626
21. Jordan PW (2000) *Designing Pleasurable Products*. Taylor & Francis, London
22. Krohne HW, Egloff B, Kohlmann, CW, Tausch A (1996) Untersuchungen mit einer deutschen Version der „Positive and Negative Affective Schedule“ (PANAS). *Diagnostica* 42 (2): 139–156
23. Kurosu M, Kashimura K (1995) Apparent usability vs. inherent usability: experimental analysis on the determinants of the apparent usability. *Conference companion on human factors in computing systems*. ACM, New York, pp. 292–293

24. Lavie T, Tractinsky N (2004) Assessing dimensions of perceived visual aesthetics of web sites. *Int J Hum-Comput St* 60 (3): 269–298
25. Leder H, Belke B, Oeberst A, Augustin D (2004) A model of aesthetic appreciation and aesthetic judgments. *British journal of psychology*, 95(4) 489–508
26. Lewis JR (2002) Psychometric evaluation of the PSSUQ: Using data from five years of usability studies. *J Hum-Comput Interact* 14 (3-4): 463–488
27. Liu Y (2003) Engineering aesthetics and aesthetic ergonomics: theoretical foundations and a dual-process research methodology. *Ergonomics* 46 (13-14): 1273–1292.
28. Maguire M (2001) Methods to support human-centred design. *Int J Comput St* 55: 587–634
29. Moshagen M, Musch J, Göritz AS (2009) A blessing, not a curse: Experimental evidence for beneficial effects of visual aesthetics on performance. *Ergonomics* 52 (10): 1311–1320
30. Nakarada-Kordic I, Lobb B (2005) Effect of perceived attractiveness of web interface design on visual search of web sites. In *Proceedings of the 6th ACM SIGCHI - CHINZ '05*. ACM: New York, pp. 25–27
31. Ngo DCL, Teo LS, Byrne JG (2003) Modelling interface aesthetics. *Inf Sci* 152: 25–46
32. Norman DA (2004) *Emotional design: why we love (or hate) everyday things*. Basic Books, New York
33. Porat T, Tractinsky N (2012) It's a pleasure buying here: The effects of web-store design on consumer's emotions and attitudes. *Hum-Comput Interact* 27(3), 235-276
34. Postrel V (2003) *The substance of style: how the rise of aesthetic value is remaking a commerce, culture and consciousness*. Harper Collins Publishing, New York
35. Prentice DA (1987) Psychological correspondence of possessions, attitudes, and values. *J Pers Soc Psychol* 53: 993–1003
36. Prümper J, Heinbokel T, & Küting HJ (1993) Virtuelle Prototypen als Werkzeuge zur benutzerzentrierten Produktentwicklung. Anwendung einer handlungstheoretischen Fehlertaxonomie auf reale und simulierte Oberflächen von Waschmaschinen. *Z Arbwiss* 4:160–167
37. Reinecke K, Yeh T, Miratrix L, Mardiko R, Zhao Y, Liu J, Gajos KZ (2013) Predicting users' first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI'13*. ACM: New York, pp. 2049–2058
38. Rheinberg F, Manig Y, Kliegl R, Engeser S, Vollmeyer R (2007) Flow bei der Arbeit, doch Glück in der Freizeit. *Z Arb Organ Psychol* 51 (3): 105–115
39. Rafaeli A, Vilnai-Yavetz I (2004) Instrumentality, aesthetics and symbolism of physical artefacts as triggers of emotion. *Theor Issue in Ergonomics Sci* 5 (1): 91–112
40. Sauer J, Wastell D, Schmeink C (2009) Designing for the home: a comparative study of support aids for central heating systems. *Appl Ergonomics* 40: 165–174
41. Sauer J, Sonderegger A (2009) The influence of prototype fidelity and aesthetics of design in usability tests: effects on user behaviour, subjective evaluation and emotion. *Appl Ergonomics* 40: 670–677
42. Sauer J, Sonderegger A (2011) The influence of product aesthetics and user state in usability testing. *Behav Inform Technol* 30: 787-796
43. Schenkman BN, Jönsson FU (2000) Aesthetics and preferences of web pages. *Behav Inform Technol* 19 (5): 367–377
44. Sonderegger A, Sauer J (2010) The influence of design aesthetics in usability testing: effects on user performance and perceived usability. *Appl Ergonomics* 41: 403–410
45. Sonderegger A, Sauer J (2013) The influence of socio-cultural background and product value in usability testing. *Appl Ergonomics* 44: 341-349
46. Sonderegger A, Sauer J, Eichenberger J (in press) Expressive and classical aesthetics: two distinct concepts with highly similar effect patterns in user-artefact interaction. *Behav Inform Technol*.
47. Sonderegger A, Zbinden G, Uebelbacher A, Sauer J (2012) The influence of product aesthetics and usability over the course of time: a longitudinal field experiment. *Ergonomics* 55(7), 713-730
48. Thielsch MT, Blotenberg I, Jaron R (2013) User Evaluation of Websites: From First Impression to Recommendation. *Interact. Comput.*, 26(1): 89–102
49. Thüring M, Mahlke S (2007) Usability, aesthetics and emotions in human-technology interaction. *Int J Psychol* 42: 253–264
50. Tinsley HEA, Hinson JA, Tinsley DJ, Holt MS (1993) Attributes of leisure and work experiences. *J Couns Psychol* 40 (4): 447–455
51. Tractinsky N (1997) Aesthetics and apparent usability: empirically assessing cultural and methodological issues. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM: New York, pp. 115-122
52. Tractinsky N, Katz AS, Ikar D (2000) What is beautiful is usable. *Inter Comput* 13 (2): 127–145

53. Tuch AN, Bargas-Avila JA, Opwis K (2010) Symmetry and aesthetics in website design: it's a man's business. *Comput Hum Behav* 26 (6): 1831–1837
54. Tuch AN, Roth S P, Hornbæk K, Opwis K, Bargas-Avila JA (2012) Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. *Comput Human Behavior* 28(5), 1596–1607.
55. Tzou RC, Lu HP (2009) Exploring the emotional, aesthetic, and ergonomic facets of innovative product on fashion technology acceptance model. *Behav Inform Technol* 28: 311–322
56. Van der Heijden, H (2003) Factors influencing the usage of websites: the case of a generic portal in The Netherlands. *Information & Management* 40(6): 541–549
57. van Schaik P, Hassenzahl M, Ling J (2012). User-experience from an inference perspective. *ACM Trans. Comput.-Hum. Interact.* 19(2): 1-25
58. van Schaik P, Ling J (2009). The role of context in perceptions of the aesthetics of web pages over time. *Int J Hum Comput St* 67(1): 79 - 89
59. Watson D, Clark LA, Tellegen A (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. *J Pers Soc Psychol* 54: 1063–1070
60. Yamamoto M, Lambert DR (1994) The impact of product aesthetics on the evaluation of industrial products. *J Prod Innov Manage* 11: 309–324
61. Zeng L, Proctor RW, Salvendy G (2012) User-based assessment of website creativity: a review and appraisal. *Behav Inform Technol*, 31(4), 383-400