



Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI

Alexandre N. Tuch^{a,*}, Sandra P. Roth^a, Kasper Hornbæk^b, Klaus Opwis^a, Javier A. Bargas-Avila^a

^a University of Basel, Department of Psychology, Center for Cognitive Psychology and Methodology, Switzerland

^b Department of Computer Science, University of Copenhagen, Njalsgade 128, DK-2300 Copenhagen, Denmark

ARTICLE INFO

Article history:

Available online 21 April 2012

Keywords:

Aesthetics
Beauty
Emotion
Interface design
Usability
User experience

ABSTRACT

This paper analyzes the relation between usability and aesthetics. In a laboratory study, 80 participants used one of four different versions of the same online shop, differing in interface-aesthetics (low vs. high) and interface-usability (low vs. high). Participants had to find specific items and rate the shop before and after usage on perceived aesthetics and perceived usability, which were assessed using four validated instruments. Results show that aesthetics does not affect perceived usability. In contrast, usability has an effect on post-use perceived aesthetics. Our findings show that the “what is beautiful is usable” notion, which assumes that aesthetics enhances the perception of usability can be reversed under certain conditions (here: strong usability manipulation combined with a medium to large aesthetics manipulation). Furthermore, our results indicate that the user's affective experience with the usability of the shop might serve as a mediator variable within the aesthetics–usability relation: The frustration of poor usability lowers ratings on perceived aesthetics. The significance of the results is discussed in context of the existing research on the relation between aesthetics and usability.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Within the field of human–computer interaction (HCI) and especially in the context of user experience (UX) research, the aesthetics of user interfaces has become a topic of major interest. Numerous studies show the influence of aesthetics on for instance trust and credibility (Karvonen et al., 2000; Robins and Holmes, 2007), perception of usability (Ben-Bassat et al., 2006; Thüring and Mahlke, 2007; Tractinsky et al., 2000), usability testing (Sonderregger and Sauer, 2010) and overall impression (Schenkman and Jönsson, 2000; Tractinsky et al., 2000; Tuch et al., 2010). A recent review of empirical studies of UX has shown that aesthetics is one of the most frequently researched dimensions in this field (Bargas-Avila and Hornbæk, 2011).

Interface-aesthetics was mostly neglected until papers by Kurosu and Kashimura (1995) and Tractinsky et al. (2000). They showed that the visual aesthetics of an interface significantly influences users' perceived ease of use of the entire system. Following the publication of Kurosu and Kashimura, numerous studies have investigated the relation between usability and aesthetics (for an overview see Hassenzahl and Monk, 2010). Several of these studies found support for the aesthetics–usability relation (e.g., Hartmann et al., 2008; Lavie and Tractinsky, 2004; Quinn and Tran, 2010), but

other studies could not find this relation (e.g., Hassenzahl, 2004; van Schaik and Ling, 2009).

Based on the empirical findings of previous studies, it is currently unclear under which circumstances the aesthetics of an interface influences perceived usability, or vice versa. This can be attributed to the lack of experimental studies manipulating aesthetics and usability as independent variables. Although most studies are correlative (e.g., De Angeli et al., 2006; Tractinsky, 1997; van Schaik and Ling, 2003), some UX models propose a causal relation between interface-aesthetics and perceived usability (e.g., Hassenzahl, 2004; van Schaik and Ling, 2008). These models are based on correlative data – the causality is solely a matter of theoretical reasoning and cannot be tested by existing data. In a recent article, Hassenzahl and Monk (2010) concluded that there is a lack of studies testing the effects of aesthetics on perceived usability (and vice versa) through experimental manipulation.

In order to create a good user experience, it is important to understand the relation between aesthetics and usability, as well as the processes underlying this relation. We set up a strictly controlled laboratory study within which we systematically manipulated interface-aesthetics and interface-usability. Special attention was paid to the independent manipulation of the experimental factors. Our study provides further insights into the possible causal relations between usability and aesthetics.

* Corresponding author. Tel.: +41 61 2673568.

E-mail address: a.tuch@unibas.ch (A.N. Tuch).

2. Related work

2.1. What is beautiful is usable

In a widely cited study, Tractinsky et al. (2000) showed a high correlation between the perceived beauty of an automatic teller machine (ATM) interface and the users' perceived ease of use of the ATM interface. After participants had experienced the ATM's actual usability, the beautiful-usable correlation became even stronger. Tractinsky was able to show that users' post-use perception of usability was positively affected by the aesthetics of the interface and not by the actual usability of the system. Thus, the aesthetics of the interface affects users' perception of usability. Thinking of the underlying processes of the aesthetics effect, Tractinsky proposed the occurrence of a so-called "halo-effect": The beauty of an interface outshines all other features of the interface and therefore influences users' evaluation of the entire system. Tractinsky also discussed whether there might be an affective response toward the aesthetic interface, which would lead to an improvement of users' mood and hence positively affect the overall evaluation of the system. Nevertheless, Tractinsky called for further research on the contingencies and boundaries of the aesthetics–usability relation. Subsequently, several studies investigated the aesthetics–usability relation with a wide range of different products and different approaches toward measuring of aesthetics and usability. To give an overview of the current state of research, we

have summarized those studies in Table 1 (correlative) and Table 2 (experimental).

2.2. Correlative studies

Most studies investigating the aesthetics–usability relation are correlative. Thus, aesthetics and usability were not systematically manipulated as independent experimental factors. Typically, participants had to rate a product before and/or after usage on perceived usability and perceived aesthetics. Then the two measures were correlated. Most studies found moderate to strong correlations between perceived usability and perceived aesthetics (see Table 1). However, some studies did not find such a correlation (Hassenzahl, 2004; van Schaik and Ling, 2009). In contrast to other studies, Hassenzahl (2004) controlled the correlation between perceived usability (pragmatic quality; PQ) and beauty for hedonic quality (a judgment of a product's potential to support pleasure in use and ownership). Hedonic qualities are strongly correlated to beauty. Therefore, it is not surprising that the beauty-usability correlation disappeared when controlling for hedonic quality. Unlike other studies, van Schaik and Ling (2009) correlated the actual usability (mental workload and task performance) and not the perceived usability with aesthetics, possibly an explanation for the absence of a correlation between aesthetics and usability. In sum, the correlative studies show some convergent evidence for the interrelation between usability and aesthetics. Nevertheless, correlative

Table 1

Summary of studies using correlative analysis to investigate the aesthetics–usability relation.

Source	Product (task)	Correlation (<i>r</i>)	Usability metrics	Aesthetics metrics
Kurosu and Kashimura (1995)	ATM layouts (passive viewing)	Pre-use: .59	Self-made (1 item)	Beautiful (1 item)
Tractinsky (1997)	ATM layouts (passive viewing)	Pre-use: .83 to .92	Self-made (1 item)	Beautiful (1 item)
van Schaik and Ling (2003)	Websites (information retrieval)	Post-use: .49	DES-R (6 items)	Aesthetics (1 item)
Hassenzahl (2004) study 1	MP3 player skins (passive viewing)	Pre-use: .07 ^a	PQ (7 items)	Beauty (1 item)
Hassenzahl (2004) study 2	MP3 player skins (usage scenarios)	Pre-use: .14 ^a ; Post-use: .08 ^a	PQ (7 items)	Beauty (1 item)
Lavie and Tractinsky (2004)	Online shop (shopping task)	Post-use: .68 to .78 (CA); .40 to .46 (EA)	Self-made (4 items)	CA and EA (10 items)
Chawda et al. (2005)	Search tool (search task)	Pre-use: .76; Post-use: .71	SUS	Self-made (items unknown)
De Angeli et al. (2006)	Websites (information retrieval)	Post-use: .38 to .48 ^b (CA); .38 ^b (EA)	Adapted from Lavie and Tractinsky (2004) (5 items)	CA and EA (10 items)
Cyr et al. (2006)	WAP sites (information retrieval)	Post-use: .24 (PLS path coefficient)	PEOU (3 items)	Self-made (4 items)
Hartmann et al. (2007)	Websites (browsing)	Post-use: .43	Self-made (1 item)	Self-made (1 item)
van Schaik and Ling (2008)	Websites (information retrieval)	Pre-use: .12 (beauty); .48 (HQI); −.47 (HQS) Post-use: .41 (beauty); .69 (HQI); .02 (HQS)	PQ (5 items)	Beauty (1 item); HQI (6 items); HQS (3 items)
van Schaik and Ling (2009)	Websites (information retrieval)	Pre-use: .11 to .18 (SMEQ); −.19 to .02 (performance) Post-use: −.02 to −.13 (SMEQ); −.31 to .16 (performance)	SMEQ and task performance	CA and EA (6 items); attractiveness (1 item)
Quinn and Tran (2010)	Cell phones (phone usage)	Post-use: .50 to .53	SUS	Self-made (7 items)

Note: CA = classical aesthetics (Lavie & Tractinsky, 2004); DES-R = display evaluation scale (Spenkelink et al., 1993); EA = expressive aesthetics (Lavie & Tractinsky, 2004); HQI = hedonic quality identification (Hassenzahl, 2004); HQS = hedonic quality stimulation (Hassenzahl, 2004); PEOU = perceived ease of use (Koufaris, 2003); PLS = partial least square; PQ = pragmatic quality (Hassenzahl, 2004); SMEQ = subjective mental effort questionnaire (Zijlstra, 1993); SUS = system usability scale (Brooke, 1996).

^a Correlation is controlled for HQI and HQS.

^b Original work reports only significance levels.

studies only allow for limited inference on the directions of any causality between aesthetics and usability.

2.3. Experimental studies

Only a few experimental studies investigate the directions of the aesthetics–usability relation. These studies manipulate aesthetics and usability as experimental factors and examine the impact of that manipulation on perceived usability and aesthetics. This allows for conclusions about the directions of the aesthetics–usability relation (see Table 2). Three out of five studies showed a significant effect of interface-aesthetics on perceived usability (Ben-Bassat et al., 2006; Lee and Koubek, 2010; Tractinsky et al., 2000), whereas two report a trend in this direction (Mahlke and Thüning, 2007; Thüning and Mahlke, 2007). More beautiful interfaces were evaluated higher on usability. In addition to the aesthetics effect, Ben-Bassat et al. (2006) and Lee and Koubek (2010) found a significant effect of interface-usability on perceived aesthetics, meaning that a more usable interface led to higher aesthetics ratings. Nevertheless, the effects of aesthetics were more pronounced than the effects of usability (see main effects (η_p^2) in Table 2). In conclusion, it seems that a pure “what is beautiful is usable” notion is only partially supported. There is also some evidence that in certain cases the relation is best described as “what is usable is beautiful”.

2.4. Shortcomings of previous research

Prior research reveals that several aspects of the aesthetics–usability relation remain unclear. In the following, we summarize the shortcomings of existing research.

2.4.1. Manipulation of aesthetics and usability

As mentioned earlier, the directions of the aesthetics–usability relation are currently unclear. To our knowledge, only five experimental studies provide reliable information on the directions of the relation. Because their findings are mixed, it is not possible to draw

a clear conclusion. An explanation for the mixed result pattern might be the different ways of experimental manipulation and assessment of the dependent measures between the studies. A key difficulty for experimental studies is the systematic and independent manipulation of aesthetics and usability. This is crucial for drawing causative conclusions about the aesthetics–usability relation. In some studies, aesthetics and usability appear to be confounded. For instance, manipulating the aesthetics factor through changing or moving certain interface-elements involves the risk of unintentionally changing the interface’s usability (Tractinsky et al., 2000). Manipulating usability may have a similar confounding effect on aesthetics.

2.4.2. Boundary conditions

Previous research put rather limited attention to possible boundary conditions of the aesthetics–usability relation, even though Tractinsky et al. already emphasized the importance of boundary conditions: “Obviously, more research is needed to assess the contingencies and boundaries of the aesthetics–usability relationships.” (Tractinsky et al., 2000, p. 142). With boundary conditions we describe the possibility that different degrees of usability and aesthetics manipulation may affect the aesthetics–usability relation differently. For instance, a large effect size of the usability manipulation may attenuate the effect of interface-aesthetics on perceived usability. No previous experimental studies took these conditions into account and indicated the strength of their manipulation by means of effect sizes. For all experimental studies we therefore calculated the effect size of their manipulation checks as a proxy for the strength of the experimental manipulation. Table 2 shows that most studies manipulate aesthetics as well as usability quite strong and with equal strengths (large effect sizes of similar strengths). Two exceptions are the studies of Ben-Bassat et al. (2006), where the strength of the usability manipulation is extremely high, and Thüning and Mahlke (2007), where the strength of the usability as well as the aesthetics manipulation is medium.

Table 2
Summary of experimental studies of the aesthetics–usability relation.

Source	Product (tasks)	Main effects (η_p^2) ^a	UB metrics	AE metrics	UB factor	AE factor	MC (Cohen's f) ^a
Tractinsky et al. (2000)	ATM layouts (ATM usage)	AE on perceived UB (.037)	Self-made (1 item)	Self-made (1 item)	2 Levels (system delays)	3 Levels (placing of buttons)	UB: large (1.87) AE: large (2.37)
Ben-Bassat et al. (2006)	Simulated phone book (data entry)	AE on perceived UB (.189) ^b UB on perceived AE (.056) ^b	Adapted ^c (4 items)	Self made (3 items)	2 Levels (numbers of key strokes)	2 levels (visual design)	UB: large (3.50) AE: large (.79) ^e
Thüning and Mahlke, 2007, study 2	Simulated audio players (player usage)	Trend AE on perceived UB (.034)	SUMI (sub dimensions)	CA or EA ^d	2 Levels (navigation elements)	2 Levels (different skins)	UB: medium (.30–.32) AE: medium (.37)
Thüning and Mahlke, 2007, study 3	Simulated audio players (player usage)	Trend AE on perceived UB (.035)	SUMI (sub dimensions)	CA or EA ^d	2 Levels (navigation elements)	2 Levels (different skins)	UB: large (.73–1.00) AE: large (.81)
Mahlke and Thüning (2007)	Simulated audio players (player usage)	Trend AE on perceived UB (.035)	SUMI (sub dimensions)	CA or EA ^d	2 Levels (navigation elements)	2 Levels (different skins)	UB: large (.83) AE: large (.81)
Lee and Koubek (2010)	Websites (information retrieval)	AE on perceived UB (.167) UB on perceived AE (.141)	PSSUQ (8 items)	CA and EA (10 items) ^f	2 Levels (content organization)	2 Levels (color, layout, font)	UB: large (.77) AE: large (.81)

Note: AE = aesthetics; CA = classical aesthetics (Lavie & Tractinsky, 2004); EA = expressive aesthetics (Lavie & Tractinsky, 2004); MC = manipulation check; PSSUQ = post-study system usability questionnaire (Lewis, 2002); SUMI = software usability measurement inventory (Kirakowski & Corbett, 1993), UB = usability.

^a Effect sizes were not reported in the original papers, we calculated them from the indicated F -values and df.

^b From the reported data it was unclear which F -value goes with which main effect.

^c from Lavie & Tractinsky (2004).

^d It is not indicated which dimension was used.

^e Data from a pilot-study.

^f CA and EA were averaged and analyzed as a single scale.

2.4.3. Assessment of aesthetics and usability

As illustrated in Tables 1 and 2, many different instruments are used to assess aesthetics and usability. Some of them are self-made and not validated, others are mature and validated scales. On the one hand, it is important to assess usability and aesthetics by means of different instruments, because they are multi-faceted constructs. On the other hand, the use of different instruments makes it difficult to compare the results among the different studies. This is also a reason for the difficulty of drawing a strong overall conclusion from existing studies. To counter this issue there is a need to assess aesthetics and usability within the same experiment by means of different validated scales.

2.5. Hypotheses

The aim of the present study is to address some of the reported shortcomings of existing research. To give further insights into the relation between aesthetics and usability, we set up an experimental study to independently manipulate the factors usability and aesthetics. Further, we assessed perceived aesthetics and usability by means of several different validated scales, allowing us to compare the scales among each other within the experiment and to compare our results with those of previous studies. As the experimental setting (product), we used a realistic online shop with a shopping task.

Based on the previous findings of the summarized studies in Tables 1 and 2, we used the following hypotheses: Interface-aesthetics affects perceived usability before usage (H_1). Interface-aesthetics affects perceived usability after usage (H_2). Interface-usability affects perceived aesthetics after usage (H_3).

3. Method

3.1. Design

The experiment used a three-factor, mixed design. The between-subject independent variables were *interface-usability* with two

levels (low vs. high) and *interface-aesthetics* with two levels (low vs. high). The within-subject independent variable was *time* with two levels (pre-use vs. post-use). The primary dependent variables were *perceived usability* and *perceived aesthetics* (see Fig. 1).

3.2. Participants

A total of 80 participants took part in the experiment (42 females), with a mean age of 25.7 years ($SD = 7.7$). Participants' mean experience in using the web was 10.8 years ($SD = 2.9$) and all had previously shopped online. As compensation for taking part in the experiment, participants received an equivalent of 50 US\$. In order to motivate the participants during the experiment, they were told that the four best performing participants would win an iPod shuffle. Participants were randomly assigned to one of the four experimental conditions.

3.3. Independent variables

In order to operationalize the independent variables, four different versions of an online shop were fully implemented (see Section 3.5). We manipulated interface-usability and interface-aesthetics as independently as possible by changing only those parts of the user interface that affect either usability or aesthetics.

3.3.1. Interface-usability

To manipulate usability, two different information architectures (IAs) were built: one good (*high interface-usability*) and one bad (*low interface-usability*). The relation between quality of IA and usability is well researched (e.g., Blackmon et al., 2005; Chi et al., 2001; Chi et al., 2000; Ivory et al., 2001; Leuthold et al., 2008; Leuthold et al., 2011). Thus, interface-usability was manipulated by solely changing the labels of the online shop's IA. The usable and unusable versions of the shop looked identical, contained exactly the same product items and had the same menu depth and breadth. The versions only differed in menu labels and assignment of the items to menu categories.

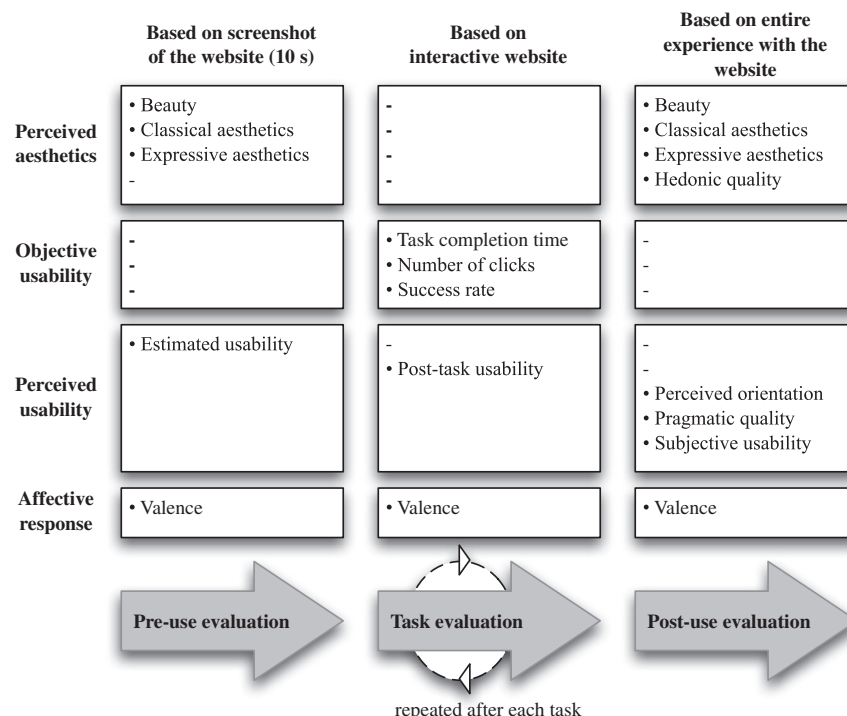


Fig. 1. Overview of experimental procedure (see Table 4 for detailed description of the dependent variables).

Table 3

Latent semantic analysis: mean information scent for navigation-paths leading to the target item compared with mean paths not leading to a target item in the high and low usability condition.

High usability		Low usability	
NPT	NPnT	NPT	NPnT
<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
.51 (.09)	.30 (.08)	.22 (.09)	.20 (.06)

Note: NPT = navigation-path leading to target; NPnT = navigation-path not leading to target.

The information scent of the two IAs was compared using latent semantic analysis (LSA; e.g., Katz and Byrne, 2003). Table 3 illustrates that the LSA showed high scent for the navigation-paths leading to the target item (NPT) and low scent for the navigation-paths not leading to the target item (NPnT) in the good IA condition, whereas in the bad IA condition both path types showed a low scent. For examples of navigation-paths in the good and the bad IA, see Fig. 2.

3.3.2. Interface-aesthetics

To create a beautiful (*high interface-aesthetics*) and an ugly (*low interface-aesthetics*) online shop, we selected 30 professionally designed website templates from the royalty-free website <http://freecsstemplates.org>. These templates were screened by a panel of experts ($n = 4$), who chose the 10 most beautiful ones. All templates used the same content. For each of those 10 “beautiful” versions an “ugly” counterpart was created. To avoid confounding with usability, we manipulated the aesthetics of the websites without changing any interaction elements (e.g., navigation elements, shopping basket) regarding its position, text background color or layout. The only graphical factors that were manipulated were background color, background texture and decorative graphic elements. Thus, the aesthetic manipulation would not affect usability factors such as readability or interfere with users’ expectations about placement of interaction elements (see Roth et al., 2010).

Finally, this procedure led to 20 design versions of online shops (10 “ugly-beautiful” pairs).

In a preliminary online study, 178 users rated each of the 20 design versions regarding beauty. For our main study, we selected the “ugly-beautiful” pair with the largest difference in the beauty rating, $t(177) = 12.63$, $p < .01$. $d = .95$ (see Fig. 3). The beautiful version achieved a mean beauty appraisal of 4.30 ($SD = 1.61$) on a 7-point scale, whereas the ugly version reached 2.21 ($SD = 1.41$).

3.4. Dependent measures

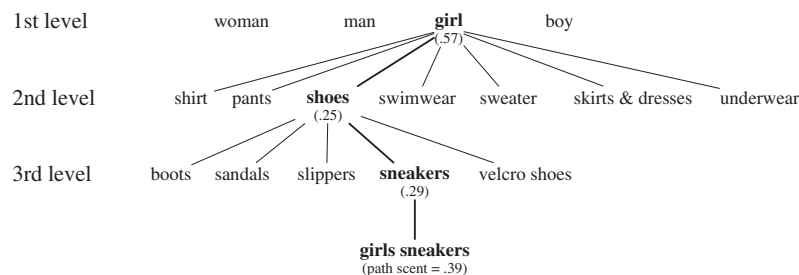
The main dependent variables were *perceived usability* and *perceived aesthetics*. To ensure the validity and reliability of the measurements, we assessed usability and aesthetics by means of several different scales and items (see Table 4). Besides the subjective appraisals, we also assessed task completion time, number of clicks, and success rate as objective usability metrics (Hornbæk, 2006), and valence as emotional measurement using the self-assessment manikin (SAM) scale (Lang, 1980; Lang et al., 2005).

3.5. Materials

3.5.1. Online shop

To conduct the experiment in a realistic web environment, we fully implemented an online shop for clothing. The shop had four categories at the top-level represented with tabs at the top of the page. Each main category had seven subcategories represented in form of a menu on the left side of the page and each subcategory had again five to seven subcategories also represented in form of a menu on the left side of the page. By clicking on a category of the lowest level, product items were displayed in the center of the page. Product items were organized in a 2×3 matrix (Schmutz et al., 2010), displaying a picture, the name and the price for each item. Clicking on a product item led to a detail page displaying a picture of the product item, the name, the price, and a detailed product description. In total, the shop contained more than 1300 different product items. Moreover, a shopping cart listing all added product items was available.

High Interface Usability



Low Interface Usability

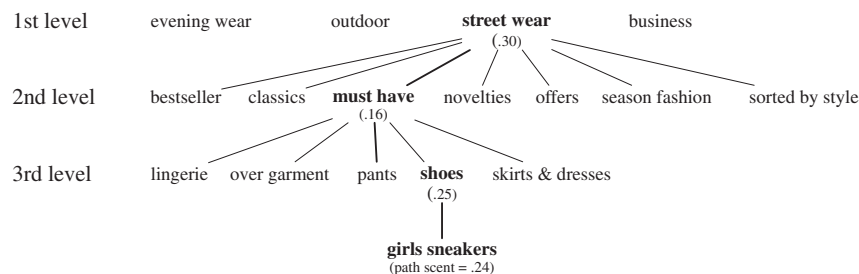


Fig. 2. Example of navigation-paths and corresponding information scents in the high and low usability condition for a specific target item.

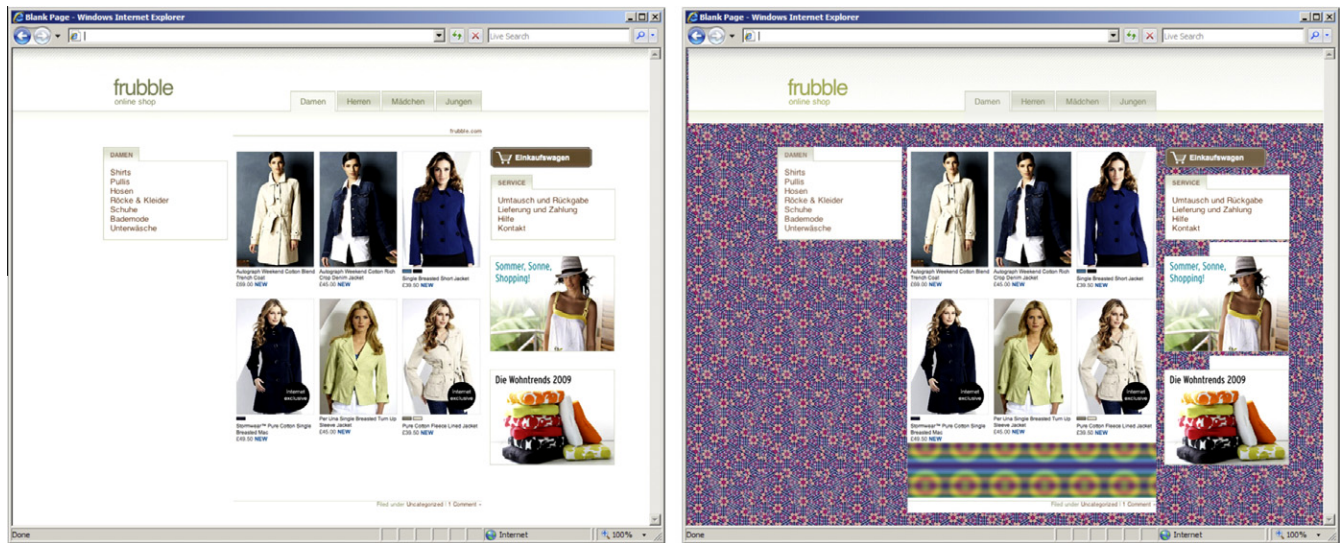


Fig. 3. The high-aesthetics (left) and low-aesthetics (right) version.

Table 4
Measurements used for this study.

	Questionnaire; reference	Item(s)	Anchors	Reliability	Range	Used in phase
<i>Perceived aesthetics</i>						
Beauty	Jacobsen et al. (2004) and Hassenzahl (2004)	1	Ugly/beautiful	–	1–7	Pre, post
Classical aesthetics	Lavie and Tractinsky (2004)	5	Disagree/agree	.78/.80	1–7	Pre, post
Expressive aesthetics	Lavie and Tractinsky (2004)	5	Disagree/agree	.73/.83	1–7	Pre, post
Hedonic quality (stimulation) ^a	AttracDiff 2; Hassenzahl (2004)	7	Semantic differential	.80	1–7	Post
Hedonic quality (identification) ^a	AttracDiff 2; Hassenzahl (2004)	7	Semantic differential	.84	1–7	Post
<i>Perceived usability</i>						
Estimated usability	None ^b	1	Unusable/usable	–	1–7	Pre
Post task usability	ASQ; Lewis (1991)	3	Strongly disagree/strongly agree	.92–.94	1–7	During
Perceived orientation	WOOS; Yom and Wilhelm (2004)	7	Disagree/agree	.95	1–5	Post
Pragmatic quality	AttracDiff 2; Hassenzahl (2004)	7	Semantic differential	.92	1–7	Post
Subjective usability	SUS; Brooke (1996)	10	Disagree/agree	.94	1–5	Post
<i>Affective response</i>						
Valence	SAM; Lang (1980) and Lang et al. (2005)	1	Negative/positive	–	1–7	Pre, during, post

^a Hedonic qualities are not actual metrics for aesthetics, nevertheless they are strongly related to perceived aesthetics.

^b Item developed by the authors for this study: “How do you rate the usability of this online shop?”.

3.5.2. Tasks

Participants had to solve four tasks in the online shop. Each task consisted of browsing for a specific product item (target item) and adding it to the shopping cart. Participants were shown a picture of the target item and asked to add an identical item to the shopping cart. Iconic task instructions were used to avoid words that may induce specific search strategies. The four target items are displayed in Fig. 4. If participants were not able to find a target item within 5 min, they were asked by the experimenter to move onto the next task.

3.6. Procedure

The experiment took place in the University's HCI lab. Instructions on the procedure of the experiment were displayed on the computer screen. Task descriptions and questionnaires were provided in paper form. It was clearly indicated when participants had to switch from the computer screen to paper and back.

Initially, participants were presented a screenshot of the online shop for 10 s. They then rated the screenshot using scales



Fig. 4. Target items for browse tasks: a tie (task 1), a spaghetti strap shirt (task 2), girls' sneakers (task 3), and a man's suit (task 4).

regarding perceived aesthetics and usability. Next, participants solved four tasks in the online shop. After each task, they rated their use-experience by means of several items. Finally, after solving all tasks in the shop, they were told to evaluate the entire interaction with the shop, especially in terms of perceived usability and aesthetics. For an overview of the experimental procedure and the measurements obtained, see Fig. 1 in conjunction with Table 4.

Table 5
Manipulation check for the experimental factors.

	<i>M</i> (<i>SD</i>) High	<i>M</i> (<i>SD</i>) Low	<i>F</i> (1,76)	η_p^2	Cohen's <i>f</i>	<i>p</i>
<i>Interface-aesthetics</i>						
Classical aesthetics	4.8 (1.0)	3.7 (1.0)	22.461	.228	.543 ^b	<.001
Beauty	3.8 (1.4)	3.0 (1.1)	7.801	.093	.320 ^a	.007
ASQ (Lewis, 1991)	3.6 (1.4)	3.5 (1.5)	.063	.001	.032	.803
Completion time (s) ^c	477 (362)	443 (354)	.471	.006	.078	.495
Number of clicks ^c	155 (130)	144 (122)	.538	.007	.084	.466
Success rate (%)	74 (24)	74 (28)	.000	.000	.000	.995
<i>Interface-usability</i>						
Classical aesthetics	4.3 (1.1)	4.2 (1.3)	.023	.000	.000	.879
Beauty	3.5 (1.5)	3.3 (1.2)	.488	.006	.078	.487
ASQ (Lewis, 1991)	2.4 (0.8)	4.7 (1.0)	130.951	.633	1.313 ^b	<.001
Completion time (s) ^c	155 (94)	781 (230)	225.507	.748	1.723 ^b	<.001
Number of clicks ^c	50 (43)	269 (95)	195.134	.720	1.604 ^b	<.001
Success rate (%)	92 (12)	55 (23)	76.859	.503	1.006 ^b	<.001

Note: No significant interactions found for all dependent variables.

^a Medium effect.

^b Large effect.

^c The displayed values are not log-transformed; statistical tests are based on the log-transformed data.

3.7. Manipulation check

Before the main analysis we checked whether the factors interface-aesthetics and interface-usability were successfully manipulated. Therefore, we ran a two-way ANOVA with perceived aesthetics (pre-use) and performance as dependent variables. Table 5 summarizes results of the manipulation check, which show that all manipulations were successful. Note that interfaces did not differ regarding expressive aesthetics ($F(1,76) = .02$, $p = .940$). This was not expected because expressive aesthetics implies attributes such as creativity, fascination or use of special effects. Given that our aesthetic manipulation aimed only at the visual surface, expressive attributes remained unchanged. No interactions between interface-aesthetics and interface-usability were found.

Regarding boundary conditions, the data show that in this study the aesthetics manipulation ($f = .32$ – $.54$; medium to large) was weaker than the usability manipulation ($f = 1.01$ – 1.72 ; large). We will return to this difference in the discussion of the results.

4. Results

All data were checked for normal distribution and linearity. For all statistical tests an alpha level of .05 was used. To investigate the effect of the experimental manipulations a set of 2×2 ANOVA with interface-aesthetics and interface-usability as independent variables was calculated. Fig. 5 provides a summary of the effects found in the pre-use and post-use phases. In the following sections, the effects will be presented in more detail.

4.1. Effects of interface-aesthetics on perceived usability

We find no effect of interface-aesthetics ($F(1,76) = .582$, $p = .448$, $\eta_p^2 = .008$) nor of interface-usability ($F(1,76) = .093$, $p = .761$, $\eta_p^2 = .001$) on perceived usability at the pre-use phase. No interaction occurred ($F(1,76) = .000$, $p = 1.000$, $\eta_p^2 = .000$). Against our expectations (H_1), the participants did not use the interface's aesthetics as a proxy for pre-use perceived usability.

In the post-use phase, interface-aesthetics also did not affect perceived usability (SUS, WOOS, PQ), contradicting H_2 rather than supporting the “what is beautiful is usable” notion. Unsurprisingly, we found a significant main effect for interface-usability. No interface-aesthetics \times interface-usability interaction occurred. Hence, the interface-aesthetics did not affect users' perception of usability

after interacting with the online shop. Perceived usability was solely affected by the actual usability of the interface (see Table 6 and Fig. 5).

4.2. Effect of interface-usability on perceived aesthetics

Interface-usability had a significant impact on perceived aesthetics during the post-use phase. We found a significant main effect for interface-usability on classical aesthetics and on hedonic quality identification (HQI). Expressive aesthetics and hedonic quality stimulation (HQS) were not affected by interface-usability. After using the online shop, participants' perceived aesthetics and HQI were influenced by the experienced usability (supporting H_3). Thus, perceived aesthetics and HQI were not solely affected by interface-aesthetics but also by interface-usability (see Table 7 and Fig. 5). For a more detailed analysis of this effect, see Section 4.2.3 in combination with Table 8.

4.2.1. A pre- and post-use comparison for perceived usability

In order to investigate the changes in perceived usability from the pre-use to the post-use phase we perform a ANCOVA with interface-usability and interface-aesthetics as independent variables. As dependent variables we use the usability scales from the post-use phase (SUS, WOOS, PQ) and as co-variable we enter estimated usability from the pre-use phase. This analysis shows no effects for the co-variable, meaning that the pre-use usability ratings are not related to post-use usability ratings. Further, the ANCOVA shows a strong main effect for interface-usability. We find neither an interaction nor a main effect for interface-aesthetics (data pattern is equivalent to the ANOVA results in Table 7). Hence, participants adapted their usability ratings after using the shop depending on its interface-usability.

4.2.2. A pre- and post-use comparison for classical and expressive aesthetics

To investigate the changes in perceived aesthetics from the pre-use to the post-use phase, a repeated measures ANOVA with time (pre-use vs. post-use) as within-subject factor and interface-usability and interface-aesthetics as between-subject factors was performed. As dependent variables we used the classical and expressive aesthetics (these scales were used in both pre and post-use phases).

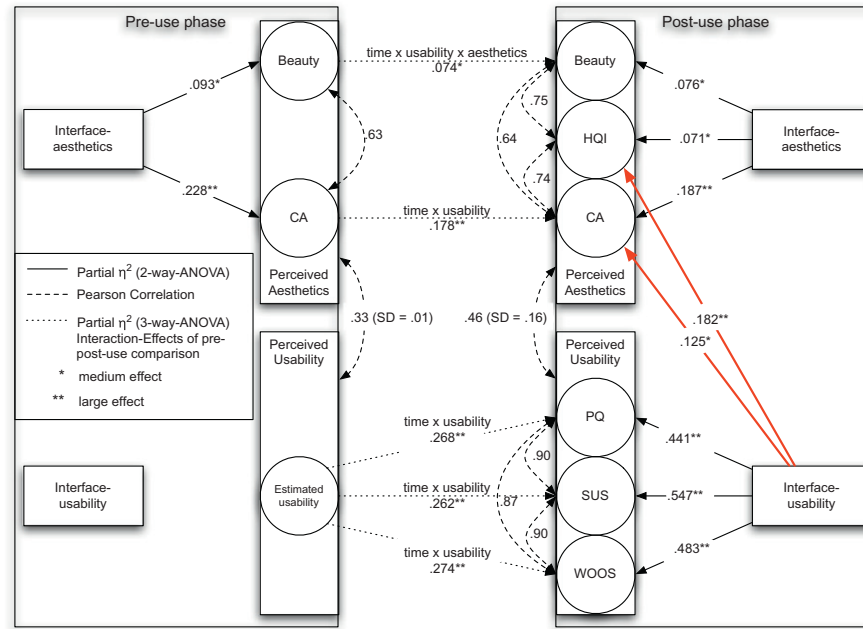


Fig. 5. Summary of results.

Table 6
ANOVA for perceived usability at the post-use phase.

	M (SD) High	M (SD) Low	F (1,76)	η_p^2	p
<i>Interface-aesthetics</i>					
Subjective usability (SUS)	3.0 (1.0)	3.1 (1.1)	.046	.001	.830
Perceived orientation (WOOS)	2.7 (1.1)	2.6 (1.3)	.154	.002	.696
Pragmatic quality (PQ)	3.8 (1.8)	3.7 (1.8)	.216	.003	.643
<i>Interface-usability</i>					
Subjective usability (SUS)	3.8 (0.6)	2.3 (0.8)	91.854	.547 ^a	<.001
Perceived orientation (WOOS)	3.4 (0.9)	1.8 (0.8)	70.896	.483 ^a	<.001
Pragmatic quality (PQ)	4.9 (1.2)	2.6 (1.4)	60.002	.441 ^a	<.001
<i>Interface-aesthetics × interface-usability</i>					
Subjective usability (SUS)	–	–	.247	.003	.621
Perceived orientation (WOOS)	–	–	.006	<.001	.926
Pragmatic quality (PQ)	–	–	.072	.001	.790

^a Large effect.

For classical aesthetics the ANOVA shows a significant *time × interface-usability* interaction ($F(1,76) = 16.443, p < .001, \eta_p^2 = .236$) and a main effect for *time* ($F(1,76) = 23.467, p < .001, \eta_p^2 = .178$). There was no three-way interaction nor a *time × interface-aesthetics* interaction. Further, the ANOVA showed a significant main effect for *interface-aesthetics* ($F(1,76) = 23.597, p < .001, \eta_p^2 = .237$), whereas the factor *interface-usability* failed marginally to reach a significant level ($F(1,76) = 3.740, p = .057, \eta_p^2 = .047$). There was no *interface-aesthetics × interface-usability* interaction. As illustrated in Figs. 5 and 6, interface-usability only affects the classical aesthetics appraisals after online shop usage: Participants changed their ratings from the pre-use to the post-use phase depending on interface-usability. Participants in the low-usability condition lowered their ratings after using the online shop, whereas participants' ratings in the high-usability

condition remained stable. Regarding expressive aesthetics, there is no interaction nor any main effects. In the following we present further analyses that provide some explanations for this pattern.

4.2.3. Effects on single items within aesthetics scales

To see which items of classical aesthetics and HQI were affected by the interface's usability, we conducted a series of independent *t*-tests with *interface-usability* as the independent variable and all items of classical aesthetics and HQI as dependent variables. The analysis showed that for classical aesthetics only the items "clean" and "organized" differed significantly, whereas for HQI, all items, except "gaudy" and "takes me distant from people", differed. It seems that classical aesthetics is only partially affected by usability. The items that were sensitive to the usability manipulation (i.e., "clean" and "organized") may be relevant not only for perceived aesthetics, but also for usability. In contrast, most items of HQI were sensitive to the usability manipulation. Although HQI is expected to be associated with aesthetics, in our study it was affected more strongly by interface-usability than interface-aesthetics (see Table 8).

4.2.4. The user's affective experience as a mediator variable

A more general explanation for the interface-usability effect on perceived aesthetics might be found in the user's affective reaction towards the usability of the online shop: A frustrating usability experience may lead to a negative affective experience for the user, which in turn reduces the ratings of perceived aesthetics. In other words, users get frustrated by poor usability and therefore they rate the interface more negatively in general, including its aesthetics. To investigate this hypothesis, we set up two path-models (see Fig. 7): (1) one with a direct effect of *objective usability*¹ on classical aesthetics and (2) the other with the user's affective response as mediator-variable. The following variables were entered in the models: Objective usability as independent variable, classical aesthetics

¹ The variable objective usability is formed using the factor score resulting from a exploratory factor analysis including the performance measures from the tasks (task completion time, number of clicks, and success rate). The extracted factor explains 83% of the variance.

Table 7
ANOVA for perceived aesthetics at the post-use phase.

	M (SD) High	M (SD) Low	F (1,76)	η_p^2	p
<i>Interface-aesthetics</i>					
Classical aesthetics	4.3 (1.1)	3.3 (1.2)	17.524	.187 ^b	<.001
Expressive aesthetics	2.0 (0.9)	2.0 (0.9)	.016	<.001	.901
Hedonic quality identification	3.6 (1.0)	3.1 (1.1)	5.804	.071 ^a	.018
Hedonic quality stimulation	2.9 (1.0)	2.8 (0.9)	.006	<.001	.940
Beauty	3.9 (1.6)	3.1 (1.3)	6.220	.076 ^a	.015
<i>Interface-usability</i>					
Classical aesthetics	4.2 (1.1)	3.4 (1.1)	10.809	.125 ^a	.002
Expressive aesthetics	2.1 (0.9)	1.9 (0.8)	.582	.008	.448
Hedonic quality identification	3.8 (1.0)	2.9 (0.9)	16.919	.182 ^b	<.001
Hedonic quality stimulation	2.7 (0.8)	3.0 (0.9)	.053	.001	.819
Beauty	3.7 (1.5)	3.3 (1.5)	.919	.012	.341
<i>Interface-aesthetics × interface-usability</i>					
Classical aesthetics	–	–	.700	.009	.405
Expressive aesthetics	–	–	1.818	.023	.182
Hedonic quality identification	–	–	2.271	.029	.136
Hedonic quality stimulation	–	–	4.843	.060 ^a	.031
Beauty	–	–	1.602	.021	.209

^a Medium effect.

^b Large effect.

Table 8
Effects of interface-aesthetics and interface-usability on classical aesthetics and hedonic quality identification broken down to single items ($N = 80$).

	Interface-aesthetics		Interface-usability	
	t	p	t	p
<i>Classical aesthetics</i>				
The design of the online-shop was ...				
Aesthetic	1.986	.051	0.760	.450
Symmetrical	2.187	.032	1.130	.262
Pleasant	3.635	<.001	1.285	.203
Organized	2.852	.006	2.714	.008
Clean	3.589	.001	5.084	<.001
<i>Hedonic quality identification</i>				
Takes me distant from people—brings me closer to people	1.678	.097	−0.269	.788
Gaudy—classy	2.278	.025	1.624	.108
Cheap—valuable	2.642	.010	1.965	.053
Noninclusive—inclusive	0.303	.762	2.694	.009
Isolating—integrating	0.653	.516	2.914	.005
Amateurish—professional	1.752	.084	4.438	<.001
Unpresentable—presentable	1.56	.123	5.345	<.001

Note: P-values in bold are significant at the 5% level (two-tailed test).

and perceived usability (SUS, WOOS, PQ) as dependent variables, and the user's affective response (SAM valence) as mediator-variable. Path coefficients suggest that the effect of objective usability on classical aesthetics is fully mediated through the user's affective response, whereas the effect on usability is only partially mediated (see Fig. 7). Similar results were obtained in the model when classical aesthetics was replaced with HQI. For more information on path analysis in general, see Kline (2005).

To get further support for this mediation effect, we reanalyzed the effects of interface-usability on perceived aesthetics (Table 7),

by entering the user's affective response (SAM valence) from the post-use phase as co-variable into the ANOVA. The consequence was that the interface-usability effect on perceived aesthetics disappeared ($F(1,75) = .084$, $p = .772$ for classical aesthetics and $F(1,75) = .215$, $p = .644$ for HQI) and only the interface-aesthetics effect remained ($F(1,75) = 15.867$, $p < .001$ for classical aesthetics and $F(1,75) = 4.325$, $p = .041$ for HQI). This finding supports the idea of mediation and suggests that the effect of usability on perceived aesthetics depends on the user's affective experience with the usability of the online shop.

5. Discussion

The results of our study show that interface-aesthetics did not affect perceived usability, whereas a low interface-usability significantly lowered classical aesthetics and HQI after usage of the online shop. Before usage, interface-aesthetics did not affect perceived usability. Further, our results suggest that the effect of interface-usability on classical aesthetics and HQI was mediated through the user's affective experience with the usability of the online shop. Users were frustrated by the low usability and lowered their aesthetics ratings.

5.1. Relation between aesthetics and usability

Our findings differ from previous research in at least two ways. First, our study is the first one able to show an interface-usability effect on perceived aesthetics and at the same time finding no interface-aesthetics effect on perceived usability. Most existing studies found an effect (or at least a trend) of interface-aesthetics on perceived usability. Two studies (Ben-Bassat et al., 2006; Lee and Koubek, 2010) also found an interface-usability effect on perceived aesthetics alongside the interface-aesthetics effect, but the effect was less pronounced. Our results show that Tractinsky's notion ("what is beautiful is usable") can be reversed to a "what is usable is beautiful" effect under certain circumstances. This may also have an implication for existing UX models. For instance, Hassenzahl and Monk (2010) propose several causal models for the aesthetics–usability relation. Within these models, the general direction of the effect is the same: aesthetics affects usability. They conclude that the implied causality in the models is solely a matter of theoretical reasoning and that experimental studies testing the effect of beauty on perceived usability (and vice versa) are needed. Our data suggest that additional factors are needed to understand the directions of the effects in these models.

Second, aesthetic perception of the interface changes significantly over time: both classical aesthetics and HQI were affected by the experienced usability. In contrast, prior work sees perceived aesthetics and hedonic attributes as stable constructs, not affected by the user's interaction experience (Hassenzahl, 2004). This raises the question as to whether there are certain conditions where perceived aesthetics and hedonic qualities are affected by experienced usability; and if there are such conditions, what do they look like? In our study, it might be that the strong usability manipulation and the performance-oriented tasks were such conditions.

There are several explanations for this divergence. One might be the strength of the manipulation of the experimental factors. Table 2 shows that aesthetics and usability were mostly manipulated to a similar degree (e.g., medium vs. medium, large vs. large effects). In contrast, our aesthetics manipulation ($f = .32$ –.54; medium to large) was weaker than the usability manipulation ($f = 1.01$ –1.72; large). It might be that in this study the strong usability manipulation outshone the interface-aesthetics and eliminated the "what is beautiful is usable" effect. Our results show the importance to research and discuss the conditions that

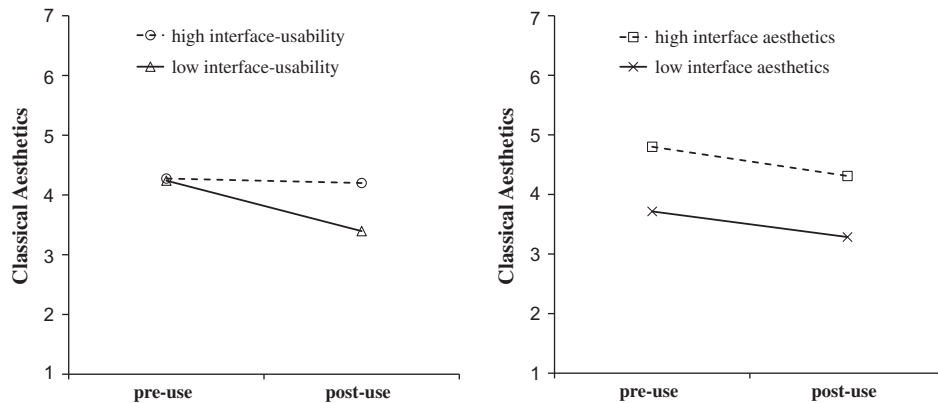


Fig. 6. Pre- and post-use scores for classical aesthetics for the usability (left) and aesthetics (right) conditions.

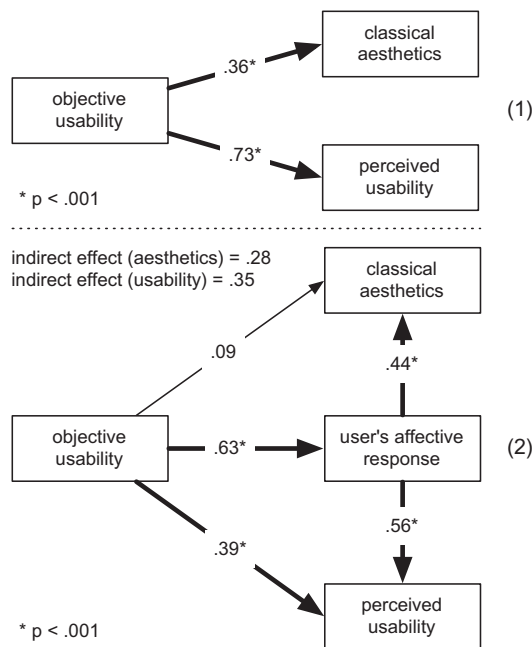


Fig. 7. Path-analytic models: (1) the effect of objective usability on classical aesthetics and usability and (2) the effect of objective usability on classical aesthetics and usability including user's affective response as mediator variable.

shape the aesthetics–usability relation (e.g., the relative strength of the manipulations) in the future more carefully.

Further reasons for the usability effect could be the applied task type. Participants had to solve rather performance-orientated tasks and they were certainly very involved in the tasks by knowing that depending on their performance, they could be rewarded with an iPod shuffle. Thus, interface-usability probably was an important factor for the users, trying to solve the task as quickly and as correctly as possible.

5.2. How do we measure: a closer look at scales

The use of multiple measurement concepts of aesthetics and usability allows a comparison between these concepts. Whereas all usability scales were highly interrelated and equally affected by the manipulations, the aesthetics measures were comparatively less interrelated and also behaved differently to the manipulations: in contrast to beauty, classical aesthetics and HQI were affected by the usability manipulation. We analyzed the items of the two affected scales separately (Table 8). Looking at classical aesthetics,

the usability manipulation affected the items “clean” and “organized” of the scale, whereas the aesthetics manipulation had an impact on nearly all items. “Clean” and “organized” are attributes that may also be seen as usability features, this would explain why classical aesthetics was affected by usability. Hassenzahl and Monk suggested that classical aesthetics differs from beauty. They even put forward that “classic aesthetics can be understood as a form of “visual” usability [...] complementing the usability of interaction” (Hassenzahl and Monk, 2010, p. 255). Regarding HQI, the usability manipulation had an impact on different items than the aesthetics manipulation. It seems that aesthetics and usability tap different aspects on the HQI scale. This is interesting, because it is assumed that hedonic qualities are system attributes that are influenced by users’ aesthetic impression of the interface.

5.3. Mediation through the user's affective experience

In regard to the underlying processes of the aesthetics–usability relation, we suggest a possible mediation through the user's affective response triggered by the usability of an interface. Tractinsky has already suggested that the user's affective experience might play an important role regarding the aesthetics–usability relation: “An affective response to the design's aesthetics may improve users' mood and their overall evaluations of the system” (Tractinsky et al., 2000, p. 130). Further, Thüring and Mahlke (2007) were able to show that users' subjective emotional appraisals were affected by interface-usability and interface-aesthetics, but psychophysiological measures were solely affected by interface-usability. Our analyses suggest that the effect of interface-usability on perceived aesthetics is indeed mediated through the user's affective response. This was not the case for perceived usability (Fig. 7). It seems that the user's perception of usability is at least partially based on the actual usability of the interface, whereas the evaluation of aesthetics is influenced by the user's affective response caused by the interaction experience. It might be important to take the user's affective experience into account as a relevant variable that significantly influences the aesthetics–usability relation.

5.4. Limitations and further research

The findings of our study are limited in several ways. First, the results are based on the evaluation of a single product (online shop). It would be interesting to know if the same findings also occur when using different products. To understand the relation between aesthetics and usability, it is crucial to show the same effects by testing several products and several user populations.

Second, we used a performance-oriented task with clearly defined instrumental goals (buying specific shopping items). It is

known that the type of task (presence or absence of a specific goal) can have an impact on users' evaluation of a system (Hassenzahl and Ullrich, 2007; van Schaik and Ling, 2009). In our case, users might have focused on usability aspects when evaluating the system. For instance, it is possible that a more "hedonic" task (e.g., buying whatever they want) or the total absence of a task would have led to a different system evaluation, focusing more on aesthetic aspects of the system. For further research, it would be worth including the task type as an additional experimental factor.

Third, the suggested influence of the user's affective response on the aesthetics–usability relation is just based on a correlation. A causal interpretation should therefore be treated with caution. To clarify the role of affective experience in aesthetic evaluation of a product, experiments which actively manipulate the user's affective experience (e.g., mood induction) are needed.

Fourth, to avoid confounding between the experimental factors, we manipulated interface-aesthetics rather subtly (i.e., background color and texture). This reduced perspective on aesthetics might lower the ecological validity of our study.

Finally, we highlighted that findings of experimental studies are mixed regarding causal relations between aesthetics and usability. For a better understanding of the causality of this relation it is important to find the boundary conditions of the specific effects in the future. Therefore, we suggest that future studies in this field should provide the effect sizes of their manipulations. This may help us to better understand the boundary conditions of the aesthetics–usability relation. In addition, further studies should manipulate aesthetics and usability to different degrees in order to observe which effects occur under which conditions.

Acknowledgments

The authors wish to thank Erik Frøkjær and Jakob Simonsen for their helpful comments on an earlier draft of this paper and Basil Keller for his help in data collection.

References

- Bargas-Avila, J., & Hornbæk, K. (2011). Old wine in new bottles or novel challenges: A critical analysis of empirical studies of user experience. In *Proceedings of the 2011 annual conference on human factors in computing systems* (pp. 2689–2698). ACM.
- Ben-Bassat, T., Meyer, J., & Tractinsky, N. (2006). Economic and subjective measures of the perceived value of aesthetics and usability. *ACM Transactions on Computer–Human Interaction*, 13(2), 210–234.
- Blackmon, M., Kitajima, M., & Polson, P. (2005). Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 31–40). ACM.
- Brooke, J. (1996). SUS – A quick and dirty usability scale. *Usability Evaluation in Industry*, 189–194.
- Chawda, B., Craft, B., Cairns, P., Rüger, S., & Heesch, D. (2005). Do "Attractive Things Work Better"? An exploration of search tool visualisations. In *Proceedings of 19th British HCI group annual conference*.
- Chi, E., Pirolli, P., Chen, K., & Pitkow, J. (2001). Using information scent to model user information needs and actions and the web. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 490–497). ACM.
- Chi, E., Pirolli, P., & Pitkow, J. (2000). The scent of a site: A system for analyzing and predicting information scent, usage, and usability of a web site. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 161–168).
- Cyr, D., Head, M., & Ivanov, A. (2006). Design aesthetics leading to m-loyalty in mobile commerce. *Information and Management*, 43(8), 950–963.
- De Angeli, A., Sutcliffe, A., & Hartmann, J. (2006). Interaction, usability and aesthetics: What influences users' preferences? In *Proceedings of the 6th ACM conference on designing interactive systems* (pp. 271–280).
- 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* (pp. 396). ACM.
- Hartmann, J., Sutcliffe, A., & De Angeli, A. (2008). Towards a theory of user judgment of aesthetics and user interface quality. *Transactions on Computer–Human Interaction*, 15(4).
- Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human–Computer Interaction*, 19(4), 319–349.
- Hassenzahl, M., & Monk, A. (2010). The inference of perceived usability from beauty. *Human–Computer Interaction*, 25(3), 235–260.
- Hassenzahl, M., & Ullrich, D. (2007). To do or not to do: Differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with Computers*, 19(4), 429–437.
- Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human–Computer Studies*, 64(2), 79–102.
- Ivory, M., Sinha, R., & Hearst, M. (2001). Empirically validated web page design metrics. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 53–60). ACM.
- Jacobsen, T., Buchta, K., Kohler, M., & Schroger, E. (2004). The primacy of beauty in judging the aesthetics of objects. *Psychological Reports*, 94(2), 1253–1260.
- Karvonen, K., Cardholm, L., & Karlsson, S. (2000). Cultures of trust: A cross-cultural study on the formation of trust in an electronic environment. In *Proceedings of the nordic workshop on secure IT systems, Reykjavik, Iceland*.
- Katz, M., & Byrne, M. (2003). Effects of scent and breadth on use of site-specific search on e-commerce web sites. *ACM Transactions on Computer–Human Interaction*, 10(3), 198–220.
- Kirakowski, J., & Corbett, M. (1993). SUMI: The software usability measurement inventory. *British Journal of Educational Technology*, 24(3), 210–212.
- Kline, R. (2005). *Principles and practice of structural equation modeling*. The Guilford Press.
- Koufaris, M. (2003). Applying the technology acceptance model and flow theory to online consumer behavior. *Information systems research*, 13(2), 205–223.
- Kurosu, M., & Kashimura, K. (1995). Apparent usability vs. inherent usability: Experimental analysis on the determinants of the apparent usability. In *CHI '95: Conference companion on human factors in computing systems* (pp. 292–293). ACM.
- Lang, P., Bradley, M., & Cuthbert, B. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical report a-6, University of Florida.
- Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 119–137). Norwood, NJ: Ablex Publishing.
- Lavie, T., & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human–Computer Studies*, 60(3), 269–298.
- Lee, S., & Koubek, R. J. (2010). Understanding user preferences based on usability and aesthetics before and after actual use. *Interacting with Computers*, 22(6), 530–543.
- Leuthold, S., Bargas-Avila, J., & Opwis, K. (2008). Beyond web content accessibility guidelines: Design of enhanced text user interfaces for blind internet users. *International Journal of Human–Computer Studies*, 66(4), 257–270.
- Leuthold, S., Schmutz, P., Bargas-Avila, J., Tuch, A., & Opwis, K. (2011). Vertical versus dynamic menus on the world wide web: Eye tracking study measuring the influence of menu design and task complexity on user performance and subjective preference. *Computers in Human Behavior*, 27(1), 459–472.
- Lewis, J. (1991). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: The ASQ. *ACM SIGCHI Bulletin*, 23(1), 78–81.
- Lewis, J. (2002). Psychometric evaluation of the PSSUQ using data from five years of usability studies. *International Journal of Human–Computer Interaction*, 14(3), 463–488.
- Mahlke, S., & Thüring, M. (2007). Studying antecedents of emotional experiences in interactive contexts. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 915–918). ACM.
- Quinn, J. M., & Tran, T. Q. (2010). Attractive phones don't have to work better: Independent effects of attractiveness, effectiveness, and efficiency on perceived usability. In *CHI '10: Proceedings of the 28th international conference on human factors in computing systems* (pp. 353–362). New York, NY, USA: ACM.
- Robins, D., & Holmes, J. (2007). Aesthetics and credibility in web site design. *Information Processing and Management*, 44(1), 386–399.
- Roth, S., Schmutz, P., Pauwels, S., Bargas-Avila, J., & Opwis, K. (2010). Mental models for web objects: Where do users expect to find the most frequent objects in online shops, news portals, and company web pages? *Interacting with Computers*, 22(2), 140–152.
- Schenkman, B., & Jönsson, F. (2000). Aesthetics and preferences of web pages. *Behaviour and Information Technology*, 19(5), 367–377.
- Schmutz, P., Roth, S., Seckler, M., & Opwis, K. (2010). Designing product listing pages – Effects on sales and users' cognitive workload. *International Journal of Human–Computer Studies*, 68(7), 423–431.
- Sonderegger, A., & Sauer, J. (2010). The influence of design aesthetics in usability testing: Effects on user performance and perceived usability. *Applied Ergonomics*, 41(3), 403–410.
- Spengelink, G., Besuijen, K., & Brok, J. (1993). An instrument for the measurement of the visual quality of displays. *Behaviour and Information Technology*, 12(4), 249–260.
- Thüring, M., & Mahlke, S. (2007). Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology*, 42(4), 253–264.
- Tractinsky, N. (1997). Aesthetics and apparent usability: Empirically assessing cultural and methodological issues. In *CHI '97: Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 115–122). ACM.
- Tractinsky, N., Katz, A. S., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, 13(2), 127–145.

- Tuch, A., Bargas-Avila, J., & Opwis, K. (2010). Symmetry and aesthetics in website design: It's a man's business. *Computers in Human Behavior*, 26(6), 1831–1837.
- van Schaik, P., & Ling, J. (2003). The effect of link colour on information retrieval in educational intranet use. *Computers in Human Behavior*, 19(5), 553–564.
- van Schaik, P., & Ling, J. (2008). Modelling user experience with web sites: Usability, hedonic value, beauty and goodness. *Interacting with Computers*, 20(3), 419–432.
- van Schaik, P., & Ling, J. (2009). The role of context in perceptions of the aesthetics of web pages over time. *International Journal of Human–Computer Studies*, 67(1), 79–89.
- Yom, M., & Wilhelm, T. (2004). WOOS-Ein Messinstrument für die wahrgenommene Orientierung in online-shops. In *Mensch and Computer* (pp. 43–53).
- Zijlstra, R. (1993). *Efficiency in work behaviour. A design approach for modern tools*. Ph.D. thesis, Delft University of Technology, Delft, The Netherlands: Delft University Press.