Awesome, but Impractical?

The Impact of Novelty and Typicality on the User's Perception of Online Shop Aesthetics and Usability

Master Thesis

Elisa Mekler, B.Sc.

Institute of Psychology

Department for Cognitive Psychology and Methodology

University of Basel

November 2011

Thesis Supervisors:

Alexandre N. Tuch, M.Sc.

Center for Cognitive Psychology and Methodology, University of Basel

Prof. Dr. Klaus Opwis

Center for Cognitive Psychology and Methodology, University of Basel

Abstract

This study investigates the relation between aesthetics and usability by analyzing the effects of novelty and typicality on users' perception of online shop aesthetics and usability and on user performance. In an online study, 93 participants interacted with one of four variations of the same online shop, differing in novelty (high vs. low) and typicality of web object placement (high vs. low). Participants had to find specific products and rate the shop before and after interaction on perceived aesthetics and perceived usability, using several validated instruments. Results show that novelty does affect both perceived usability and aesthetics. However, it does so in different ways: High novelty online shops were considered more aesthetic, while low novelty shops were perceived as more usable. User performance was only affected by typicality. High typicality online shops allowed for more time-efficient navigation and less clicks. Lastly, typicality slightly affected aesthetics over time. This points to a two-way relation between aesthetics and usability. Findings suggest that in the context of online shopping 'what is beautiful' is not usable and what is usable is somewhat beautiful.

Introduction

The field of human-computer interaction (HCI) traditionally focused on usability and usefulness of interactive systems and treated aesthetics as irrelevant (Nielsen, 2009). Yet studies show that the effects of aesthetics are not to be underestimated. It is suggested that aesthetic product design sustains corporate financial performance (Hertenstein, Platt & Veryzer, 2005; Verganti, 2006), promotes corporate credibility (Lowry, Roberts & Higbee, 2006; Robins & Holmes, 2008) and affects consumers' emotions (Porat, Liss & Tractinsky, 2007; Thüring & Malke, 2007).

Several studies have also examined the relation between aesthetics and usability. Some showed that aesthetics affect the users' perception of usability (Kurosu & Kashimura, 1995; Hartmann, Sutcliffe & De Angeli, 2007; Lavie & Tractinsky, 2004; Quinn & Tran, 2010; Sonderegger & Sauer, 2010; Tractinsky, Katz & Ikar, 2000). Tuch, Roth, Hornbaek, Opwis and Bargas-Avila (under review) however, suggest that usability determines whether a product is perceived as aesthetic or not. Although there is evidence for a two-way relation between aesthetics and usability (Ben-Bassat, Meyer & Tractinsky, 2006; Lee & Koubek, 2010). Yet other studies found no relation between these dimensions (Hassenzahl, 2004; Thüring & Mahlke, 2007; van Schaik & Ling, 2009).

However, many studies analyzing the aesthetics-usability relation have one major drawback: It is often unclear how and what aspects of aesthetics and usability have been manipulated. This hampers the understanding of the aesthetics-usability relation and offers only little practical implications, as both aesthetics (Hekkert & Leder, 2008) and usability (ISO/IEC 9241-11, 1998) are multifaceted constructs. And while Hassenzahl (2004) stresses the importance of examining different aesthetic dimensions, he states no reasons why in his study some of the MP3-player skins, which varied greatly in appearance and usability, were perceived as more beautiful than others. Thus, designers receive no practical advice on what aspects of aesthetics or usability they should focus on to create beautiful and usable products.

This study hopes to address this shortcoming by consistently manipulating specific facets of aesthetics and usability of online shops. Hence, novelty was chosen to alter the aesthetics of online shop layouts, as it is an important feature in distinguishing functionally similar products (Hertenstein et al., 2005) and often a key factor to a product's success on the market (Talke, Salomo, Wieringa & Lutz, 2009). Moreover, Hekkert, Snelders and van Wieringen (2003) showed that novel objects are perceived as more aesthetic than less novel objects.

In order to vary usability, typicality of web object placement was manipulated. Roth, Schmutz, Pauwels, Bargas-Avila and Opwis (2010) showed that users have clear expectations on where web objects, such as shopping carts are located and several studies showed that placing web objects according to users' expectations promotes faster user orientation (Oulasvirta, 2004; Roth, Tuch, Mekler, Bargas-Avila & Opwis, under review). Also, compared to how Tractinsky et al. (2000) manipulated aesthetics by rearranging screen objects, which possibly affected the actual usability, typicality can be independently manipulated without influencing novelty.

In short, by understanding the intricacies of the aesthetics-usability relation, user experience can be improved. The goal of this study was to investigate how novelty and typicality affect the users' perception of aesthetics and usability, as well as user performance. Results provide insights on how these facets affect the relation.

Theoretical Background

Aesthetics and Usability Relation

The field of human-computer interaction (HCI) has long followed the principle of 'form follows function'. This phrase was coined by the architect Louis Sullivan (1896) and popularized by the Bauhaus movement, which eschewed ornamental aspects of architecture and product design in favour of function, similar to how HCI often stresses usability and efficiency over aesthetics. Or as Nielsen (2009) put it: 'Users don't care about design'. However, studies (Kurosu & Kashimura, 1995; Tractinsky et al., 2000) found that ATM screens that were deemed aesthetically pleasing, were also perceived as more usable by study participants. Several more studies found evidence that 'what is beautiful is usable' is also applicable to websites (Hartmann et al., 2007; Lavie & Tractinsky, 2004; Sonderegger & Sauer, 2010) and cell phones (Quinn & Tran, 2010). Although in the study of Tuch et al. (under review) it was usability that affected the perception of online shop aesthetics. Further studies (Ben-Bassat et al., 2006; Lee & Koubek, 2010) found that aesthetics and usability affect each other mutually. Other studies (Hassenzahl, 2004; van Schaik & Ling, 2009;) however, found no evidence for a relation between aesthetics and usability.

Hassenzahl and Monk (2010) came to the conclusion that more studies need to investigate the aesthetics-usability relation through the experimental manipulation of different dimensions of aesthetics and usability. The lack of such studies is all the more surprising, when considering that several validated instruments already exist that measure different dimensions of usability (Brooke, 1996; Lewis, 1991; Yom & Wilhelm, 2004) and aesthetics. Hassenzahl (2004) created the AttrakDiff, an instrument that measures hedonic quality, which describes how engaging and interesting a product is, as well as pragmatic quality, which describes how well a product serves to achieve user goals. Lavie and Tractinsky (2004) defined two clearly distinguishable dimensions of perceived aesthetics. Namely classical aesthetics, which emphasizes clear and simple design, and expressive aesthetics, which

describes how original and creative a product is. Lastly, Moshagen and Thielsch (2010) identified additional facets that pertain to website aesthetics.

In short, a major shortcoming in the research of the aesthetics-usability relation is the lack of studies manipulating aspects of aesthetics and usability as independent variables. Both usability and aesthetics (Hekkert & Leder, 2008) are multifaceted constructs. Hassenzahl (2004) for example stresses the importance of varying different dimensions of aesthetics and usability, thus rendering study results applicable to real product design. Still, he states no reasons why one stimulus in his study was much more favourably received by participants than the other stimuli. This not only contributes fewer insights for future research, but also offers few practical implications, as no concrete information is given to designers on how to optimize the aesthetics and usability of their products. Tractinsky et al. (2000) attempted to manipulate both aesthetics and usability of ATM layouts, but the simple variations offer only little ecological validity. Compare this to Tuch et al.'s (under review) experimental study, which provides insights on how web designers can make an online shop more attractive to users by labeling the information architecture in a way that matches users' expectations. Still, the subtle manipulation of aesthetics (different background colours and textures) in that study limits the practical implications somewhat.

In conclusion, both aesthetics and usability are complex constructs that are more than the sum of their parts. Yet in order to better understand them, it is necessary to scrutinize their components. This study focuses on novelty as an aesthetic property and typicality of web object placement as an aspect of usability.

Novelty

Novelty is an important feature in many industries, as products are often similar in function, but distinguishable by aesthetic design (Hertenstein et al., 2005). Hence, more and more companies successfully focus on product design as a competitive tool (Verganti, 2006). Take Apple's iPod for example, launched in 2001, and compare it to other MP3-players at the

time. According to Talke et al. (2009), it is Apple's focus on novel product design that was key to its success.

Hekkert and Leder (2008) state that people are attracted to novel products or modern art because they derive pleasure from processing new objects. And indeed, Hekkert et al. (2003) demonstrated in several studies that novelty affects the aesthetics of objects as diverse as telephones, teakettles, cars and sanders. Results showed that objects were considered attractive, when they were perceived as novel.

Location Typicality

According to Norman (1983), users form mental models of objects they interact with. These models are further influenced by experiences with similar objects and the users' knowledge. In short, a mental model is the users' knowledge and expectations about how a certain object functions and how it is interacted with. Bernard (2001) showed that users also build mental models of websites and have expectations on where elementary web objects, such as web page logos, navigation or advertisement banners, are located. Roth et al. (2010) expanded these findings to different website types, such as company, online shop and news sites. They showed that users also have expectations of where objects specific to a certain type of website, such as shopping carts, are typically located. Roth et al. also hint at a relation between web object placement and usability. Subsequently, they developed these patterns of user expectations into an index of typicality, an instrument that measures typicality. In this study henceforth, typicality will describe the degree to which web object placement meets user expectations.

Norman (1983) states that by taking the users' expectations into account during the product design, the efficiency and quality of interaction is enhanced. And indeed, studies have shown that web objects placed according to users' expectations, promote more efficient user orientation (Oulasvirta, 2004; Roth et al., under review).

Aim of the study

Goal of this study is to systematically manipulate the degrees of novelty and typicality of online shops, in order to investigate how novelty and typicality affect the perception of online shop aesthetics and usability, as well as user performance. Thus, providing further insights on the aesthetics-usability relation and offering implications for web design.

It is assumed that novelty affects the perception of online shop aesthetics and usability. According to Hekkert et al. (2003), users will perceive the high novelty online shops as more aesthetic than the less novel shops. And because the majority of study findings (Kurosu & Kashimura, 1995; Hartmann et al., 2007; Lavie & Tractinsky, 2004; Quinn & Tran, 2010; Sonderegger & Sauer, 2010; Tractinsky, Katz & Ikar, 2000) conform to the notion of 'what is beautiful is usable', it is expected that online shops that are perceived as aesthetic by study participants, will also be perceived as more usable. User performance should not be affected by novelty. In line with the findings of previous typicality studies (Oulasvirta, 2004; Roth et al., under review), it is to be expected that high typicality will also positively affect the perception of usability, as well as enhance user performance, whereas low typicality will be perceived as less usable and decrease user performance. Finally, pre- and post-use ratings will be compared. According to prior findings (Lee & Koubek, 2010; Tuch et al., under review), novelty will already affect perceived aesthetics and usability prior to shop interaction, whereas both novelty and typicality will impact ratings post-use.

Method

Experimental Design

A three-factor mixed design was used. The between-subject independent variables were novelty (high vs. low) and typicality (high vs. low) of online shop design. The within-subject independent variable was time with two levels (pre-use and post-use). The dependent

variables examined in this study are perceived aesthetics, perceived usability and user performance.

Participants

Participants were recruited from the Department of Psychology's, University of Basel, own database, where people interested in participating in studies may sign up. Six 50.- francs gift coupons for Digitec, an online shop specializing in electronic goods, were raffled among all the study participants. Because of initially low participant numbers, additional participants were recruited from the author's own personal environment.

Participants were contacted via email, which included the link to the study.

In total, 173 participants took part in the study, 114 of whom did actually complete the questionnaire. Participants that did not complete the entire study, were excluded from analysis. Another 19 participants were discarded from the sample, as they did not properly interact with the online shops according to tracking data (no clicks on product pages). This is either due to participants not following task instructions or technical problems, which were mentioned by three participants and which might have influenced participants' ratings of the online shop designs. Another two participants were excluded from analysis, because they did not properly fill in the questionnaires. The final sample of 93, consisted of 41 male and 51 female participants. One participant chose not to disclose his or her gender. Participants' experience with computers, online shops or designs did not differ significantly over all experimental conditions (for descriptives, refer to Table 1).

Table 1

Descriptive statistics over all experimental conditions

	Novelty high		**Novelty low**	
	Typicality high	**Typicality low**	**Typicality high**	**Typicality low**
	M (SD)	M (SD)	M (SD)	M (SD)
N	17	22	26	28
Age	32.2 (12.4)	33.2 (15.6)	31.1 (11.5)	31.7 (15.3)
Computer experience	5.6 (1.1)	5.1 (1.2)	5.4 (1.0)	5.0 (0.8)
Design experience	3.3 (1.4)	2.8 (1.6)	3.1 (1.3)	2.8 (1.6)
Internet experience	5.7 (0.7)	5.4 (0.9)	5.4 (1.0)	5.3 (0.8)
Online Shopping experience	5.2 (1.1)	4.2 (1.7)	4.6 (1.4)	4.3 (1.3)
Web Design experience	3 (1.8)	2.2 (1.4)	2.6 (1.4)	2.3 (1.6)

Note. Experience was rated on a 7-point Likert scale (1 = no experience; 7 = expert)

Materials

Online Shop: In order to conduct the experiment in a realistic web environment, a fully functional online shop specializing in beach apparel for men and women was programmed. One navigation bar was used to switch between the four top-level categories women, men, girls and boys. Each of these main categories had four subcategories, represented in another navigation bar: Swimwear, beach fashion, footwear and accessories. Each of those subcategories was further divided into two to four subcategories (e.g. Women's swimwear -> One-pieces, bikinis and tankinis). By clicking on a category of the lowest level, products were displayed in the center of the page in a 2 x 3 matrix. Each product item was represented by a picture, a label and a price tag. Clicking on a product led to a detail page, containing name, price and a bigger image of the product, as well as a product description. The detail page also contained a button to add the product to the shopping cart. The shopping cart, listing all added

products, could be accessed at any time. Altogether, the shop contained well over 500 unique products and product descriptions.

Layouts: In order to operationalize the independent variables, four different versions of the online shop had to be created. To manipulate typicality, the placement of three web objects was varied: The logo, two navigation bars and the shopping cart. Data from the study of Roth et al. (2010) provided information on where users expect or do not expect those web objects to be located. For the high typicality conditions, the logo was placed in the upper left corner, the shopping cart in the upper right corner, the top-level navigation bar beneath them and the lower-level navigation to the left of the page. In the low typicality conditions the logo was placed in the lower left corner, the shopping cart in the lower right corner, top-level navigation above those and the lower-level navigation to the right of the web page.

To manipulate novelty, existing websites were taken and their background images and web objects altered. In the case of shopping carts for example, very simple shopping cart icons were used in the low novelty conditions, whereas more colorful and complex images were implemented in the high novelty conditions (Figure 1).

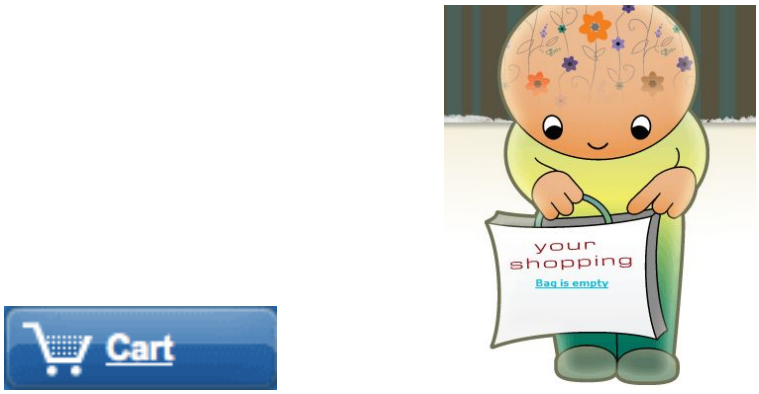

Figure 1. Examples of a low novelty shopping cart (Amazon.com) and a high novelty shopping cart (Lennartz, 2008).

In total, three sets of mock-ups were created in Photoshop. Each set contained four design variations: 1) High novelty, high typicality; 2) high novelty, low typicality; 3) low novelty,

high typicality and 4) low novelty, low typicality. Web object placement was identical in all three mock-up sets, but they varied in appearance and products offered. In order to evaluate which one of the three sets was best suited for implementation in the main study, a preliminary online study was conducted. Participants were asked to rate the 12 mock-ups and another six screenshots of existing online shops on novelty and typicality using a previously validated instrument[1]. All of the 54 participants rated all 18 screenshots.

After data collection was completed, all items describing novelty and all items describing typicality were aggregated into the dependent variables novelty resp. typicality. An analysis of variance (ANOVA) for the independent variables novelty and typicality and the dependent variables novelty and typicality showed that the set depicted in Figures 2 and 3 was best suited for the main study. For the dependent variable novelty a main effect for novelty ($F = 10.769$, $p = .001$, $\eta^2_p = .058$) and no main or interaction effect for typicality was found. The ANOVA for the dependent variable typicality resulted in a significant main effect for typicality ($F = 15.096$, $p < .001$, $\eta^2_p = .066$) and no main or interaction effect for novelty.

[1] Refer to Appendix (p. 33) for more information on instrument creation and validation.

High Typicality

Low Typicality

Figure 2. The Start pages of the high novelty online shop designs

High Typicality

Low Typicality

Figure 3. The Start pages of the low novelty online shop designs

Measurements

Several validated instruments were used for the study and overall internal consistency for all study conditions was acceptable to excellent. Alpha values are displayed in brackets.

Perceived aesthetics: To assess perceived aesthetics of the online shop designs, the dimensions of classical aesthetics (α = .77 - .93) and expressive aesthetics (α = .83 - .98) (Lavie & Tractinsky, 2004) were used, in addition to the items on hedonic quality for stimulation (α = .82 - .94) (Hassenzahl, 2004), as well as one item on beauty (Jacobsen, Buchta, Kohler & Schroger, 2004). In order to keep the study around 10 minutes long and because it was believed that users do not identify with an unknown online shop, the items of hedonic quality for identification were dropped.

Perceived usability: Different usability scales were used to operationalize perceived usability post-use, namely the System Usability Scale (SUS) (Brooke, 1996) (α = .81 - .93) and the questionnaire for perceived orientation in online shops (WOOS) (Yom & Wilhelm, 2004) (α = .87 - .96), as well as the items of pragmatic quality (PQ) (Hassenzahl, 2004) (α = .94 - .98), One item was used to estimate usability prior to shop use (Tuch et al., under review).

User performance: Objective user performance was measured through task completion time and number of clicks, which were tracked during the two online shopping tasks. Unlike the scales (SUS, WOOS, PQ) used for perceived usability, the after scenario questionnaire (ASQ) (Lewis, 1991) (α = .90 - .95) does not measure overall system usability. Instead, participants are asked to rate their own user performance on task level. Therefore, the ASQ was deemed most suitable to assess subjective user performance.

Tasks: The first task consisted of navigating the online shop in order to find a specific pair of men's swimming trunks and then adding it to the shopping cart, after which participants would return to the questionnaire. The second task was practically identical to the first, but instead another product, a beach dress, had to be found and added to the shopping cart.

Procedure

Study participants were randomly assigned to one of the four online shop layouts. Participants first answered a few general questions on demographics and on their experience with the internet, online shops and design. Participants were then presented with a screenshot of the online shop, with whom they would later interact with, and could look at it for as long as they wished to do so. After that they were asked to rate the aesthetics and estimate the usability of the online shop (see Figure 4 for details). Afterwards, participants were introduced to the first shopping task and subsequently referred to the online shop. Instructions were always available during the task in the upper left corner of the page. While interacting with the online shop, task completion time and clicks of study participants were tracked. After solving the first shopping task, participants were then transferred back to the questionnaire, where they had to fill in the ASQ items, before being introduced to the second shopping task. Task completion time and clicks were again tracked during the shopping task. After solving the second shopping task, participants again filled in the ASQ. Lastly, participants rated the aesthetics and the usability of the shop, they had just used. After which they could leave their email address, if they wished to participate in the raffle.

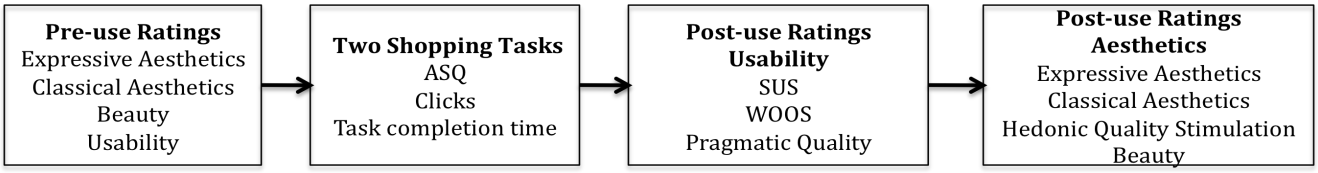

Figure 4. Overview of experimental procedure and measurements.

Results

For all statistical tests an alpha level of .05 was used. In order to investigate the effects of novelty and typicality on perceived aesthetics, perceived usability and user performance, an analysis of variance (ANOVA) with novelty and typicality as independent variables was calculated. In order to meet the criteria required for ANOVA, data was square root

transformed, log-transformed or inverse transformed, thus assuring normal distribution and homogeneity of variance.

Effects on perceived aesthetics

Pre-use: No interaction effect for novelty and typicality was found for pre-use expressive aesthetics (Refer to Table 2 for statistical values). But there was a significant main effect for novelty, yet no main effect for typicality. For pre-use classical aesthetics, results show a tendency for an interaction effect for novelty and typicality, but no significant main effects for either novelty or typicality could be found. Lastly, no interaction effect for pre-task beauty was found, but there was a significant main effect for novelty on pre-task beauty. No main effect for typicality was found.

Table 2

ANOVA for Pre-Use Perceived Aesthetics

	M (SD)	M (SD)	F	η^2_p	p
Novelty	high	low			
Beauty	4.0 (0.2)	3.3 (0.2)	5.830	.061	.018
Classical Aesthetics	3.9 (0.2)	4.3 (0.2)	2.019	.023	.159
Expressive Aesthetics	3.8 (0.2)	2.5 (0.2)	19.311	.180	<.001
Typicality	high	low			
Beauty	3.6 (0.2)	3.6 (0.2)	.000	.000	.998
Classical Aesthetics	4.1 (0.2)	4.0 (0.2)	.049	.001	.825
Expressive Aesthetics	3.2 (0.2)	3.2 (0.2)	.008	.000	.928
Novelty x Typicality					
Beauty			.041	.000	.839
Classical Aesthetics			3.435	.039	.067
Expressive Aesthetics			.219	.002	.641

Note. The displayed values are not transformed; statistical tests are based on the transformed data.

Post-use: For post-use expressive aesthetics no interaction effect was found (Table 3). There was a significant main effect for novelty and no main effect for typicality. The high novelty shops had higher ratings on the dimensions of expressive aesthetics. For classical aesthetics, no interaction or main effects were found whatsoever. Nor were interaction effects for hedonic quality stimulation found. Instead there was a significant main effect for novelty and again no main effect for typicality. The high novelty shops had much higher HQ-S ratings than the low novelty shops. There were no significant interaction or main effects for novelty and typicality on post-use beauty though. Still, there was a tendency that novelty influences

the post-task beauty rating. The high novelty shop was considered more beautiful when it was also low on typicality. Apart from the dimensions of classical aesthetics, findings overall are in line with the hypothesis that high novelty shops are perceived as more aesthetic than low novelty shops.

Table 3

ANOVA for Post-Use Perceived Aesthetics

	M (SD)	M (SD)	F	η^2_p	p
Novelty	high	low			
Beauty	4.2 (0.2)	3.5 (0.2)	3.444	.039	.067
Classical Aesthetics	4.0 (0.2)	4.3 (0.2)	.997	.011	.321
Expressive Aesthetics	3.8 (0.2)	2.5 (0.2)	16.628	.157	<.001
Hedonic Quality Stimulation	4.3 (0.2)	2.8 (0.1)	27.958	.248	<.001
Typicality	high	low			
Beauty	3.8 (0.2)	3.9 (0.2)	.038	.000	.847
Classical Aesthetics	4.3 (0.2)	4.1 (0.2)	.621	.007	.433
Expressive Aesthetics	3.4 (0.2)	3.0 (0.2)	1.634	.018	.205
Hedonic Quality Stimulation	3.6 (0.2)	3.5 (0.2)	.276	.003	.601
Novelty x Typicality					
Beauty			.296	.003	.588
Classical Aesthetics			2.126	.024	.148
Expressive Aesthetics			.027	.000	.869
Hedonic Quality Stimulation			.010	.000	.919

Note. The displayed values are not transformed; statistical tests are based on the transformed data.

Effects on perceived usability

Pre-use: Homogeneity of variance could not be achieved for pre-use usability, but the ANOVA indicates that there are no interaction or main effects for novelty or typicality.

Post-use: First of all, an ANOVA with the independent variables novelty and typicality was conducted using the different usability scales as dependent variables. There were no interaction effect for the perceived usability measurements, but significant main effects for novelty could be found for all of them (Table 4). Low novelty shops were perceived as more usable than high novelty shops. On all three scales, the low novelty, high typicality online shop scored highest. These findings refute the hypothesis that shops that are perceived as novel and aesthetic, would also be perceived as usable. Instead, low novelty shops, while perceived as less aesthetic, are perceived as more usable.

Table 4

ANOVA for Post-Use Perceived Usability

	M (SD)	M (SD)	F	η^2_p	p
Novelty	high	low			
Perceived orientation (WOOS)	3.3 (0.1)	3.8 (0.1)	6.240	.070	.014
Pragmatic Quality (PQ)	4.2 (0.2)	5.6 (0.2)	14.680	.150	<.001
Subjective usability (SUS)	64.4 (3.1)	77.0 (2.7)	9.126	.099	.003
Typicality	high	low			
Perceived orientation (WOOS)	3.6 (0.1)	3.5 (0.1)	.425	.005	.516
Pragmatic Quality (PQ)	5.1 (0.2)	4.7 (0.2)	1.110	.013	.295
Subjective usability (SUS)	72.5 (3.1)	68.9 (2.7)	.722	.009	.398
Novelty x Typicality					
Perceived orientation (WOOS)			1.035	.012	.312
Pragmatic Quality (PQ)			.008	<.001	.931
Subjective usability (SUS)			1.790	.021	.185

Note. The displayed values are not transformed; statistical tests are based on the transformed data.

Effects on user performance

Over both tasks, no interaction effects and no main effects for novelty were found for task completion time, number of clicks or the ASQ. Yet results show a main effect for typicality on task completion time and number of clicks (see Table 5 for statistical values). For the ASQ a marginally not significant main effect for typicality was found. Still, participants who used the high typicality shops were faster and used less clicks than participants who worked with the low typicality shops. And the former also had the tendency to rate their own user performance more favorably than the latter.

Table 5

ANOVA for User Performance

	M (SD)	M (SD)	F	η^2_p	p
Novelty	high	low			
ASQ	5.0 (0.2)	5.5 (0.2)	2.148	.024	<.146
Number of clicks	6.8 (0.4)	7.0 (0.3)	.115	.002	.735
Task completion time	46.0 (4.7)	48.5 (3.8)	.163	.003	.688
Typicality	high	low			
ASQ	5.6 (0.2)	5.0 (0.2)	3.897	.042	.051
Number of clicks	6.2 (0.3)	7.5 (0.3)	6.674	.086	.012
Task completion time	40.0 (4.4)	54.6 (4.1)	5.880	.084	.018
Novelty x Typicality					
ASQ			.670	.007	.415
Number of clicks			.006	<.001	.937
Task completion time			.162	.003	.688

Note. The displayed values are not transformed; statistical tests are based on the transformed data.

Effects over time

Perceived aesthetics: To compare pre- and post-use perceived aesthetics over time, a repeated measures ANOVA with time (pre-use and post-use) as within-subject factor and novelty and typicality as between-subject factors was conducted. The dependent variables were the expressive aesthetics, classical aesthetics and beauty scales used before and after shop usage.

For expressive aesthetics an interaction effect for time and typicality was found ($F = 4.367$, $p = .040$, $\eta^2_p = .047$), but no other interaction or main effect for time. There was also no

interaction effect for novelty and typicality. A significant main effect for novelty ($F = 19.193$, $p < .001$, $\eta^2_p = .179$) was found though, yet no main effect for typicality. Novelty determined whether shops scored high on expressive aesthetics, but typicality apparently changed the expressive aesthetics ratings over time, as high typicality shops scored higher on expressive aesthetics after usage, whereas low typicality shop scored lower after usage (Figures 5 and 6). Still, there was no typicality main effect for post-use expressive aesthetics (Table 3). No interaction of main effects whatsoever were found for classical aesthetics.

No interaction effects for time were found for beauty. But there was a main effect for time ($F = 4.442$, $p = .038$, $\eta^2_p = .049$). There was no interaction effect for novelty and typicality, but a main effect for novelty ($F = 5.056$, $p = .027$, $\eta^2_p = .056$). No main effect for typicality was found. This means that high novelty shops were perceived as more beautiful than low novelty shops, but after usage all shop variations were considered more beautiful.

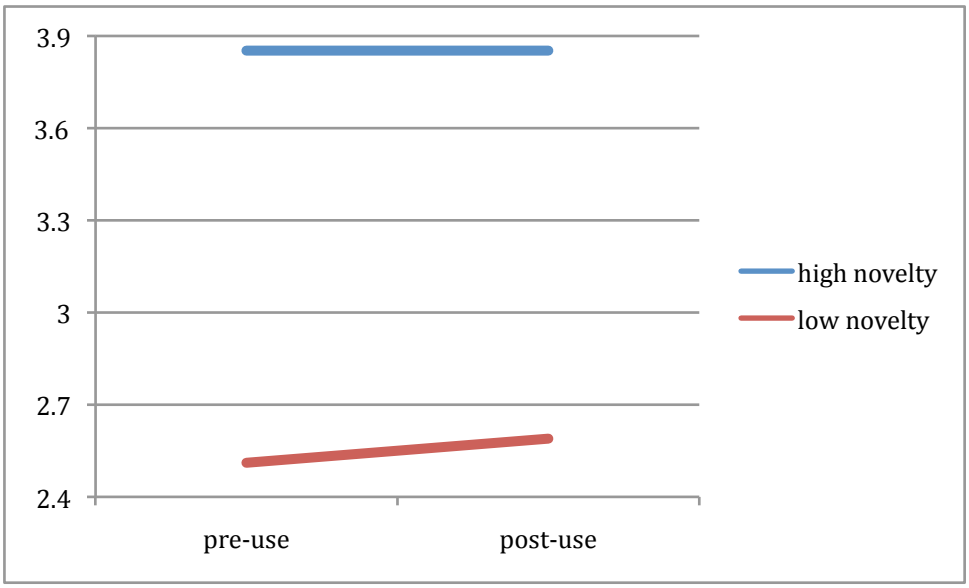

Figure 5. Pre- and post-use score for expressive aesthetics and novelty. Note that the displayed values are not transformed; statistical tests are based on the transformed data.

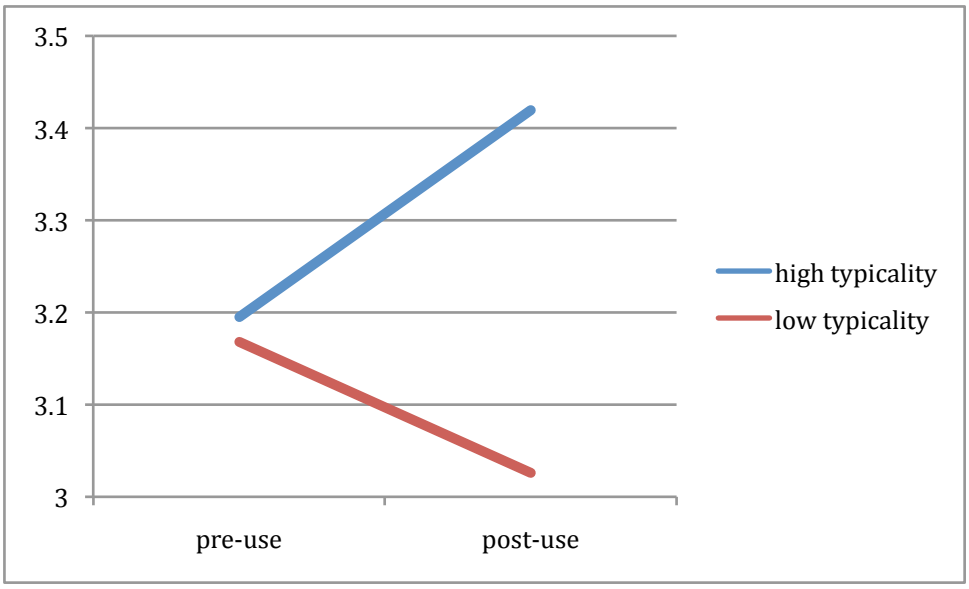

Figure 6. Pre- and post-use score for expressive aesthetics and typicality. Note that the displayed values are not transformed; statistical tests are based on the transformed data.

Perceived usability: To examine how perceived usability changed before and after the shopping tasks, an ANCOVA with novelty and typicality as the independent variables was performed. The dependent variables were the post-use perceived usability measures (SUS, WOOS, PQ) and the co-variable was pre-use estimated usability. No effect for pre-use usability was shown, and no interaction or main effect for typicality either. A significant main effect for novelty (refer to Table 2) was found though. This suggests that novelty made participants adapt their usability ratings.

Discussion

Results show that users perceive novel online shops as more aesthetic than less novel online shops. However, novel online shops are also perceived as less usable than low novelty online shops. User performance is unaffected by novelty, but typical web object placement improves user performance. Furthermore, all shop variations were considered more beautiful

after usage, but typicality determined whether a shop scored higher or lower on expressive aesthetics after use.

These findings further indicate that aesthetics affects usability. But instead of enhancing the perception of usability, novelty acts to its detriment. And while there was only one indication of typicality affecting expressive aesthetics, this still hints at a two-way relation between aesthetics and usability. It is also important to note that perceived usability does not match actual user performance.

Although these findings provide further evidence for the relation between aesthetics and usability, they differ from previous research in several ways. Apparently, Tractinsky et al.'s statement 'what is beautiful is usable' (2000) does not apply to online shopping web sites. Users rather seem to think that 'what is novel is beautiful, but not usable'. Instead, the fact that typicality affected expressive aesthetics ratings after use conforms to studies that found evidence for both aesthetics affecting usability and to a lesser extent, usability affecting aesthetics (Ben-Bassat et al., 2006; Lee & Koubek, 2010). Furthermore, results confirm Hekkert et al.'s (2003) findings that novelty is perceived as aesthetic by users. However, typicality did not affect the perception of usability, but its effect on speeding up user performance is in line with the findings of Bernard (2001), Oulasvirta (2004) and Roth et al. (under review). Finally, the fact that participants perceived shops as more beautiful after use, adheres to the mere exposure effect (Zayonc, 1968), which states that objects are more favorably perceived by people after repeated exposure.

Despite being perceived as aesthetic, high novelty shops scored lower on usability ratings than low novelty shops. This may be due the nature of the shopping tasks. According to van Schaik and Ling (2009) users' judgment and perception of interactive products depend on the mode of use. Participants were given a clear task, before interacting with the online shops. Hence, participants were in goal mode, which made them strive for efficiency over arousal and perceive novelty as a hindrance to usability. Afterwards, when participants were asked to

rate aesthetics, they switched to the action mode. High arousal is preferred in the action mode, so aesthetic ratings were maybe favorably influenced by novelty. Or perhaps aesthetic ratings would have been extremer if participants' action modes had been triggered by different tasks.

According to the results, users were well able to judge a website's aesthetic dimensions by looking at a screenshot. Yet it seems that users were not able to judge a website's usability by simply looking at it. This is contrary to the findings of Tractinsky et al. (2000), where participants used aesthetics as an indicator for judging usability, even before interaction. Again, this may be due to the context (van Schaik & Ling, 2009) in which participants had to rate usability. Firstly, participants were not sensitized to usability issues and therefore, may have been unable to anticipate any potential usability problems they would encounter during interaction. Secondly, there was only one item measuring usability pre-use, amidst several items describing aesthetic dimensions. Thus, participants were primarily focused on rating aesthetics and not usability.

Finally, this study offers some implications for web design. In order to distinguish their online shop from other online shops, designers are encouraged to experiment with novel online shop design, provided that the design does not jeopardize user performance, by compromising users' expectation of a typical online shop (see also Tuch et al., under review). Or as Hekkert and Leder (2008) put it: 'Attractive designs comprise a thoughtful balance between novelty and typicality'. The fact that users perceive novel online shops as less usable, poses less of an issue, as long as user performance is unaffected by these novel designs. However, further research is needed to examine whether novelty impairs online shop credibility or users' willingness to buy.

Limitations and further research

Next, several shortcomings of this study have to be addressed, as they limit the validity of the results: First of all, results of this study apply to online shops only. It still has to be seen whether these findings can be replicated for other website types or even other interactive

systems. Second, the final sample in the main study was rather small, especially for the high novelty conditions, as many participants had to be excluded from analysis due to technical problems. Lastly, the online shop variations were not created by professional web designers. It would be interesting to examine whether the results presented in this paper would persist in professionally designed online shop variations.

Also, participants were never asked to interact with any web element other than navigation and product pages. The identical shopping tasks might offer a plausible explanation for the few main effects and the lack of interaction effects for typicality. Therefore, future studies concerned with typicality of web object placement should require participants to interact with all web objects on a website. Besides, the identical structure of the tasks raises another issue. Participants might have given different ratings if they could have freely explored the online shops, instead of browsing them for specific products. According to van Schaik and Ling (2009), the context in which participants have to interact with websites, does impact the users' aesthetic perceptions. Therefore, future studies should vary the nature of tasks or compare the differences in ratings between participants required to follow a given task and participants asked to freely navigate a website.

Further research is required to observe the effects of aesthetic properties (Hekkert & Leder, 2008) other than novelty on users' perception of aesthetics and usability. The relation between aesthetics and usability has already been examined through the dimensions of complexity (Tuch, Bargas-Avila, Opwis & Wilhelm, 2009) and symmetry (Tuch, Bargas-Avila & Opwis, 2010). And while there is evidence that cultural sensitivities influence aesthetic preference (Zhang, Feick & Price, 2006), little is known of its influence on usability. It would also be interesting to observe how aesthetics and usability ratings change over time. Karapanos, Hassenzahl and Martens (2008) showed that aesthetics ratings mostly persist, while novelty wears off after several weeks. Future studies could examine how those fluctuations affect perceived usability and user performance. In short, for a better understanding of how product

aesthetics are related to usability, future studies need to vary the properties and the degree of manipulation of aesthetics and usability.

Finally, further research is needed to find out how novelty or other properties of aesthetics affect dimensions other than usability. Willingness to buy is crucial in the context of online shopping, as is overall user satisfaction. Several studies demonstrated that aesthetics affect user satisfaction (Schenkman, & Jönsson, 2000; Tractinsky et al., 2000; Tuch et al., 2010) and credibility (Lowry et al., 2006; Robins & Holmes, 2008). It would be interesting to observe how novelty affects user satisfaction or whether users would perceive a novel corporate website as credible or untrustworthy. To conclude, further insights are gained by manipulating aspects of aesthetics and usability.

References

Amazon. (n. d.) Retrieved October 26, 2011, from http://www.amazon.com/

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.

Bernard, M. L. (2001). User Expectations For the Location of Web Objects. In: *Proceedings of the ACM CHI 2001*, 171–172.

Brooke, J. (1996). SUS - A quick and dirty usability scale. Usability evaluation in industry, 189–194.

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*, ACM, 387-396.

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.

Hekkert, P., Snelders, D. & van Wieringen, P. C. W. (2003). ‚Most advanced, yet acceptable': Typicality and novelty as joint predictors of aesthetic preference in industrial design. *British Journal of Psychology*, *94*, 111-124.

Hekkert, P. & Leder, H. (2008). Product Aesthetics. In Schifferstein, N. J. & Hekkert, P. (Eds.), *Product Experience* (259-285). San Diego, CA: Elsevier.

Hertenstein, J.H., Platt, M.B. & Veryzer, R.W. (2005). The Impact of Industrial Design Effectiveness on Corporate Financial Performance. *Journal of Product Innovation Management, 22*, 3–21.

ISO/IEC 9241-11 (1998). *Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability*. International Organization for Standardization.

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.

Karapanos, E., Hassenzahl, M. & Martens, J.-B. (2008). User experience over time. In: *CHI '08 extended abstracts on Human factors in computing systems*. ACM, 729-738.

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*. ACM, 292–293.

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*, 530-543.

Lennartz, S. (2008). Shopping Carts Gallery: Examples and Good Practices. In *Smashing Magazine*. Retrieved October 25, 2011, from http://www.smashingmagazine.com/2008/02/07/shopping-carts-gallery-examples-and-good-practices/.

Lewis, J. (1991). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *ACM SIGCHI Bulletin, 23* (1), 78–81.

Lowry, P. B., Roberts, T. L. & Higbee, T. (2007). First Impressions with Websites: The Effects of the Familiarity and Credibility of Corporate Logos on Perceived Consumer Swift Trust of Websites. In: Jacko, J. (Ed.), *Human Computer Interaction, Part IV,* (77-85). Berlin / Heidelberg: Springer-Verlag.

Moshagen, M. & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies, 68* (10), 689-709.

Nielsen, J. (2009). Fresh vs. Familiar: How Aggressively to Redesign. In *Alertbox*. Retrieved November 2, 2011, from http://www.useit.com/alertbox/familiar-design.html

Norman, D. (1983). *Mental Models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Oulasvirta, A. (2004). Task-processing demands and memory in web interaction: A levels-of-processing approach. *Interacting with Computers, 16* (2), 217–241.

Porat, T., Liss, R. & Tractinsky, N. (2007). E-Stores Design: The Influence of E-Store Design and Product Type on Cuonsumers' Emotions and Attitudes. In: Jacko, J. (Ed.), *Human Computer Interaction, Part IV,* (712-721). Berlin / Heidelberg: Springer-Verlag.

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*, ACM, 353-362.

Robins, D. & Holmes, J. (2008). Aesthetics and credibility in web site design. *Information Processing and Management, 44* (1), 386–399.

Roth, S. P., Schmutz, P., Pauwels, S. L., Bargas-Avila, J. A., 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.

Roth, S. P., Tuch, A. N., Mekler, E. D., Bargas-Avila, J. A. & Opwis, K. (under review). Meet the users' expectations – Web objects are found faster on web pages when they are placed according to users' mental models. [submitted to *International Journal of Human-Computer Studies*]

Schenkman, B. & Jönsson, F. (2000). Aesthetics and preferences of web pages. *Behaviour & Information Technology, 19* (5), 367–377.

Sonderegger, A. & Sauer, J. (2010). The influence of design aesthetics in usability testing: Effects on user performance and perceived usability. *Applied Ergonomics, 41*, 403-410.

Sullivan, L. (1896). The Tall Office Building Artistically Considered. In *Lippincott's Magazine*. Retrieved October 29, 2011, from http://academics.triton.edu/faculty/fheitzman/tallofficebuilding.html

Talke, K., Salomo, S., Wieringa, J. E. & Lutz, A. (2009). What about Design Newness? Investigating the Relevance of a Neglected Dimension of Product Innovativeness. *Journal of Product Innovation Management*, *26*, 601-615.

Thüring, M. & Mahlke, S. (2007). Usability, aesthetics, and emotions in human-technology-interactiion. *International Journal of Psychology*, *42*, 253-264.

Tractinsky, N., Katz, A. S. & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, *13* (2), 127–145.

Tuch, A. N., Bargas-Avila, J. A., Opwis, K. & Wilhelm, F. H. (2009). Visual complexity of websites: Effects on users' experience, physiology, performance, and memory. *International Journal of Human-Computer Studies*, *67* (9), 703-715.

Tuch, A. N., Bargas-Avila, J. A. & Opwis, K. (2010). Symmetry and aesthetics in website design: It's a man's business. *Computers in Human Behavior*, *26* (6), 1831-1837.

Tuch, A.N., Roth, S.P., Hornbæk , K., Opwis, K. & Bargas-Avila, J.A. (under review). Toward Understanding the Relation Between Usability, Aesthetics, and Affect in HCI. [submitted to *Computers in Human Behavior*]

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*, 79-89.

Verganti, R. (2006). Innovating through Design. *Harvard Business Review, 84* (12), 114–22.

Yom, M. & Wilhelm, T. (2004). WOOS-Ein Messinstrument fü¨r die wahrgenommene Orientierung in Online-Shops. In: Mensch & Computer. 43–53.

Zayonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9,* 1-27.

Zhang, Y., Feick, L. & Price, L. J. (2006). The Impact of Self-Construal on Aesthetic Preference for Angular Versus Rounded Shapes. *Personality and Social Psychology Bulletin, 32,* 794-805.

Appendix

Instrument creation and validation

First, an expert panel consisting of HCI master students generated a list of items describing novelty and typicality. 8 items were chosen to delineate novelty: Original, novel, creative, unusual, innovative, fresh, exciting and peculiar. Another 8 items were chosen to characterize typicality: Placement of web objects meets expectations, typical, classical, clear, conventional, open (übersichtlich in German), meets expectations as an online shop.

To check whether this instrument would adequately cover novelty and typicality of online shop screenshots, a preliminary online study was conducted. Participants were asked to rate 18 screenshots of online shops using the 16 items generated by the expert panel. The stimuli were made up of 3 x 4 mock-up sets of online shops, all selling the same t-shirts. Each set consisted of four different novelty/typicality variations. Another 6 existing online shops were added for comparison. All 113 participants had to rate all 18 screenshots.

Internal consistency was high for items describing novelty ($\alpha_{Nov} = .93$) and for items describing typicality ($\alpha_{Typ} = .90$). A factor analysis showed that novelty explained 65.1%-74-5% of variance, whereas typicality explained 51.1%-63.4% of variance. After eliminating four novelty items and four typicality items that correlated less strongly than the other items, another factor analysis showed that novelty explained 79.2%-87.7% of variance, while typicality explained 69.6%-78.0% of variance. Thus, the final instrument consisted of four items describing novelty (novel, original, creative and innovative) and four items describing typicality (classical, typical, conventional, web object placement meets expectations).