User evaluation of websites: From first impression to recommendation

Meinald T. Thielsch & Iris Blotenberg
University of Muenster, Germany

Rafael Jaron
Nordlight Research GmbH, Germany

Abstract

Content, usability, and aesthetics are core constructs in users' perception and evaluation of websites, but little is known about their interplay in different use phases. In a first study web users (N=330) stated content as most relevant, followed by usability and aesthetics. In study 2 tests with four websites were performed (N=300), resulting data were modeled in path analyses. In this model aesthetics had the largest influence on first impressions, while all three constructs had an impact on first and overall impressions. However, only content contributed significantly to the intention to revisit or recommend a website. Using data from a third study (N=512, 42 websites), we were able to replicate this model. As before, perceived usability affected first and overall impressions, while content perception was important for all analyzed website use phases. In addition, aesthetics also had a small but significant impact on the participants' intentions to revisit or recommend.

Keywords: World Wide Web, Usability testing, User studies, Empirical studies in HCI

1. Introduction

The World Wide Web plays a major role in the modern society of the 21st century. Its use has grown enormously over the last decade. People spend much of their business and private time online, and the web itself is a complex structure of nearly 50 billion pages (based on a combined analysis of popular web search engines on http://www.worldwidewebsize.com). Furthermore, the web is evolving. Several states of design, closely associated with technological developments, have been identified over the years (Engholm, 2002; Ivory & Megraw, 2005). Because the web and website design are prone to technological innovations, constant facets and effects of website perception and evaluation need to be identified, which has led to the investigation of three core constructs in current research on human perceptions of websites: content, usability, and aesthetics (e.g., Cober et al., 2003; Schenkman & Jönsson, 2000; Tarasewich, Daniel & Griffin, 2001). These general aspects of human perception of websites are broadly defined, and each construct comprises several sub-facets. For example, usability contains interaction-related facets, such as effectiveness of use and fulfillment of the task in question, but also contains facets related to design, such as an easily operable navigation structure and readable typography. There are correlations between the constructs, especially because website designers usually strive to optimize a website with respect to content, usability, and aesthetics. Thus, a successful website is the product of a combination of interesting content, high usability and aesthetic design. Still, these three constructs can be treated separately, and each of them represents a distinct aspect of a user's perception of a website.

The prime importance of the World Wide Web for everyday life of many people leads to questions of how people perceive, evaluate and appraise websites. The aim of this study is to provide a better understanding of key factors that influence user's evaluation of websites, and of the interplay of these factors. Of such key factors content, usability, and aesthetics are named as three of the most influential in several studies (e.g., Hartmann, Sutcliffe, & De Angeli, 2008; Schenkman & Jönsson, 2000; Tarasewich et al., 2001). Unfortunately, website content, usability, and aesthetics are not always interpreted or labeled in the same manner. The following section is therefore devoted to a short overview and a definition of each term. The question arises whether the selection of these three constructs is exhaustive. We argue that the three constructs previously identified are the essence of website perception and, therefore, of a user's evaluation of a website. Content, usability, and aesthetics are directly expressed by objects on a website and thus are very closely attached to users' visual perception and cognitive processing of a website.

1.1. Content, usability, and aesthetics – definitions and overview

The *content* of a website is certainly the main reason why most websites are visited and is therefore of prime importance. Accordingly, researchers emphasize the general importance of website content, such as Palmer (2002), who named content as one main factor for website success, or Agarwal and Venkatesh (2002), who identified content as the most important dimension of websites for specific types of businesses. In empirical models, content factors are stressed and distinguished from design factors (Huizingh, 2000; Robbins & Stylianou, 2003), and cultural differences are found when objective content features are analyzed (Robbins & Stylianou, 2003; Zhao et al., 2003). ISO 9241-151 defines content as "a set of content objects" on a web user interface, describing a content object as "interactive or non-interactive object containing information represented by text, image, video, sound or other types of media" (ISO, 2006, p. 3). Such objective features of website content are focused for

example when examining the World Wide Web as a semantic web (Berners-Lee, 2001; Hendler, 2003), when optimizing content indexing for search engines or for information searches in general (e.g., Fauzi & Belkhatir, 2010; Hendler et al., 2008; Jansen & Spink, 2006), or in research of content quality in terms of accuracy (Sutherland et al., 2005). Furthermore, there are not only objective characteristics of website content (like word count or amount of technical terms) but also subjective content experiences of website users (Huizingh, 2000; Palmer, 2002). Thus, several studies measured content perceptions via subjective user evaluations (e.g. Eysenbach & Köhler, 2002; Hartmann et al., 2008; Kang & Kim, 2006). In the current paper, we focus on the subjective perception of website content in general. This includes how interesting, comprehensible, and useful content is experienced by web users. Recipients' content perception so far has been primarily studied in a specific context, for example such as users' reactions to recruiting websites (e.g., Allen, Mahto & Otondo, 2007; Braddy et al., 2009; Pfieffelmann, Wagner & Libkuman, 2010), or in relation to a specific viewpoint, such as the linking of aspects of credibility to website content (e.g., Dutta-Bergman, 2004; Rains & Karmikel, 2009). But, several studies reveal the impact of website content and information quality on general user reactions like satisfaction, trust, commitment, task performance, website success or website preference (e.g., Aladwani & Palvia, 2002; D'Ambra & Rice, 2001; Kang & Kim, 2006; Kim & Lim, 2001; Liu & Arnett, 2000). Thus, content is treated as a key construct of websites, especially from an organizational perspective (Huizingh, 2000; Ranganathan & Ganapathy, 2002).

The content of a website is of high importance, but *usability* of a website is essential for a quick and successful finding of the contents of interest. ISO 9241-11 defines the usability of a human-system interaction as the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO, 1998, p. 2). Research in human-computer interactions as well as practical work has often focused on usability; Shneiderman and Plaisant (2009) provide a good overview. When examining usability, it is very important to distinguish between objective and subjective measures of usability (Hornbæk, 2006). While a website might appear to be very usable from objective criteria (e.g., a flat link depth, fast loading speed and good search function), it still can be experienced as unusable from a subjective user's point of view. Kurosu and Kashimura (1995) established the terms "inherent" (= objective) and "apparent" (= subjective) usability to describe this difference; these terms have been subsequently used by several authors and will be used in the current paper as well. The term "apparent usability" reveals that user evaluation and perception is central to website usability. To some extent, users might be able to judge the apparent usability of a website without actually using it – but regarding this matter further research is needed. In the current paper, we take a look at the subjective experience of usability by web users. Thus, we ask for and use measures of apparent usability.

In recent years, it has been increasingly recognized that the design of websites and users' needs go beyond pure usability, leading to a shift in focus to a more experiential perspective (e.g., Bargas-Avila & Hornbæk, 2011; Hassenzahl & Tractinsky, 2006; ISO, 2009). This perspective takes the whole user experience into account, including user perception of visual aesthetics. Thus, *aesthetics* has become a core construct as well in website evaluation. However, the term "aesthetics" is not consistently used, there are some other designations for the same construct like for example "beauty" and "pleasure". A current definition is given by Moshagen and Thielsch (2010, p. 690), which describes website aesthetics "as an immediate pleasurable subjective experience that is directed toward an object and not mediated by intervening reasoning". Various studies have demonstrated the importance of website

aesthetics in human-computer interactions and its impact on constructs such as perceived usability, satisfaction or trustworthiness and on users' reactions like preference, urge to buy impulsively or intention to revisit (for an overview see Moshagen & Thielsch, 2010, p. 691). Current research suggests that aesthetic responses occur immediately at first sight (Leder et al., 2004; Thielsch & Hirschfeld, 2012), and thus the importance of first impressions is stressed in the evaluation of website aesthetics (e.g.; Lindgaard et al., 2006 & 2011; Tractinsky et al., 2006; Tuch et al., 2012). In the current paper we follow the definition of Moshagen and Thielsch (2010) and therefore operationalize aesthetics as subjective experience of a user, and measure aesthetics with a questionnaire approach.

1.2. The interplay of content, usability, and aesthetics perception

Beginning with the work of Kurosu and Kashimura (1995) and its replication by Tractinsky (1997), many studies have examined the relationship between usability and aesthetics (for overviews see Hassenzahl & Monk, 2010, and Lee & Koubek, 2012). Lee and Koubek (2010 & 2012) distinguished between pre- and post-use relationships between both constructs and the specific influences on preferences. They found a stronger connection between usability and aesthetics before than after use. Additionally, before use aesthetics influenced preference stronger than usability, after use the preference of web users was equally affected by both constructs. Different explanations have been given for the often reported high correlation between both constructs: halo-effects (e.g., Hartmann et al., 2008), mediation by other constructs (e.g., Hassenzahl & Monk, 2010), mediation by mood (Ilmberger, Schrepp & Held, 2008; Moshagen, Musch & Göritz, 2009), common design features (e.g., Lavie & Tractinsky, 2004; Tarasewich et al., 2001) or, because both constructs can be treated experimentally without affecting each other, common method bias (Thielsch, 2008; for common method bias effects in general see Podsakoff et al., 2003). The connection between usability and aesthetics and the underlying mechanisms are still not completely understood and need to be explored further.

There has been much less research on users' perceived connection between content and usability, content and aesthetics, and the interplay of all three constructs in general. Nevertheless, several studies revealed connections in specific domains. For example, in terms of more objective characteristics of web content there is a clear connection between the accessibility of information and their usability (e.g., Soobrah & Clark, 2011; Watanabe, 2009). Furthermore, there is a great body of research about users' information search and knowledge acquisition online, and how they are affected by users' expertise, content actuality or content comprehensibility (e.g., Eysenbach & Köhler, 2002; Metzger, 2007; Scharrer et al., 2012). Correlations between content and aesthetics have occasionally been reported (e.g., Aladwani & Palvia, 2002; Thielsch, 2008). For example, de Wulf et al. (2006) proposed a model in which, among other factors, website content leads to website success mediated by pleasure. Hartmann et al. (2008) found a positive influence of usability on the perception of content quality. Using screenshots of recruiting websites, Cober et al. (2003) showed that content, usability, and aesthetics are important factors for organizational attraction and recommendation intentions, but in their study, only content and navigational usability contributed independently to the organizational attraction. De Angeli, Sutcliffe and Hartmann (2006) showed that the perception of content, usability, and aesthetics was influenced by different interaction styles: overall preference could be predicted by differences in usability and aesthetics. In a study by Schenkman and Jönnson (2000), aesthetics was the best predictor of overall judgment. In contrast, while analyzing three websites, Tarasewich et al. (2001) found that usability and content were more important, although aesthetics was relevant to the

overall web experience. Sillence et al. (2007) found that sometimes web user rejected even high-quality content due to bad design.
To sum up, there is a great body of research dealing with the connection of usability and aesthetics or effects of web content in specific domains. But research that simultaneously considers users' perceptions of all three constructs while analyzing their general reactions to websites is rare and the resulting evidence is limited. The few existing studies used only small sets of websites, partially non-validated measures, and some evaluated only screenshots – thus it is not very surprising that the results were inconsistent.

1.3. Phases of website use

There are several models describing the different phases of website use (e.g., Lindgaard et al., 2011; Metzger, 2007; Ou & Sia, 2010) as well as general models of technology use, that have been used in website research (e.g., Thüring & Mahlke, 2007: Tractinsky, 2004; van der Heijden, 2003). Most of such models have in common, that there is an exposure and impression formation phase, an evaluation and use phase, as well as intentional outcomes. When visiting a website, information processing and impression building starts immediately within milliseconds (see Thielsch & Hirschfeld, 2012 as well as Tuch et al., 2012). Thus, it is important to distinguish between immediate and deliberate first impressions. While immediate first impressions are highly based on bottom-up processes of our visual perception during website exposure, deliberate first impressions are more top-down driven and based on reflective cognitive processes and reasoning during a website evaluation phase. Thus, the current research proposed two different stages of first impression; both mostly build within the first seconds of use (Lindgaard et al., 2006: Robins & Holmes, 2008; Tuch et al., 2012). In the current paper we will investigate the impact of deliberate first impressions.
While exploring and using a website overall impressions are build. Such impressions could be influenced by task characteristics (van Schaik & Ling, 2009). Overall impressions of websites in our paper are operationalized as post-use measurements of participants' impressions. Moreover, even when looking at a single website visit, there are several possible outcomes of website use, like for example buying intentions, costumer loyalty or intentions to revisit and recommend a website. In the current paper we take a look at the two latter ones, as these can easily be assessed for any type of website.
To sum up, we distinguish between three phases of website use: A pre-use first impression phase, a post-use overall impression phase, and a phase where intentions towards a website are build. The latter one will result in future usage behavior of a web user, which can only be assessed in longitudinal research approaches. A web user might experience these phases sequently, but it is also possible that he or she builds for example intentions only based on first impressions. Leaving a website just after a few seconds (and skipping a deep evaluation or use phase) is very common among web users (Robins & Holmes, 2008).

1.4. Aim of our research

The aim of our research is to further explore and analyze the interplay of content, usability, and aesthetics in evaluations of websites given by website users. Previous studies have seldom considered these three constructs together to predict users' reactions to websites, and the results with respect to the impact of each construct have been mixed. Therefore, we observe all three constructs at the same time and analyze different phases of the users' reactions. The definition of the constructs as well as prior research suggests that the three constructs should have different impacts depending of the time of the evaluation (first versus

overall impression) and the aim of the evaluation (focus on the website itself versus focus on revisiting or recommending the website). For example, the importance of aesthetics in first impressions has been stressed in prior research, while the general importance of content could suggest that content is relevant to all phases of website use.

In a first step, our aim is to explore the general perspective of web users, and to analyze their beliefs about the constructs in question. Thus, in study 1 we directly ask web users which aspect of website perception is important at which phase of use and if there are any other relevant aspects besides content, usability, and aesthetics. Furthermore, we test if such general beliefs are influenced by actual visiting a website or by familiarity with a website.
Asking users about their beliefs and evaluations of a website is a common, but limited approach. Therefore, our aim in the second study is to build a path model describing the connections between content, usability, and aesthetics perceptions, and the different use phases based on website tests.
Subsequently, the goal of study 3 is to replicate this path model using different stimuli, different measures, and a different sample. In doing so, a validation of the found model is provided.
In total, we analyzed data from three different studies. We combined a direct evaluation of user beliefs with more indirect predictions based on large website evaluations working towards a model of website evaluation based on users' reactions in different use phases.

2. Study 1: General website evaluation – user beliefs

The aim of the first study was to explore the general perspective of web users regarding the different aspects of website evaluation. We asked web users to rate the given constructs without evaluating a specific website in detail. We examined if other aspects besides content, usability, and aesthetics were rated as important. Furthermore, we determined if these general importance ratings were influenced by actually visiting a website (known or unknown). Therefore, participants were divided into three groups: One group was asked to directly answer the given questions; two groups were asked to freely choose a known or unknown website before answering the questions.

2.1. Methods

2.1.1. Participants

A total of 330 participants (61.5 % female) anonymously took part in this web-based study on a voluntary basis. Ages ranged from 14 to 68 years, with a mean age of 31.64 years (*SD* = 10.05). The education level of 83.3 % of the participants was high school diploma or higher. All subjects had used the Internet before. On average, they had been using the Internet for 12.50 years (Min = 4, Max = 25, SD = 3.34) and 17.02 hours a week (Min = 1, Max = 75, SD = 13.36).

2.1.2. Materials

Participants in two of three conditions were allowed to freely choose a website. In the condition "known website", only a few websites were visited by more than one respondent; in sum 88 different websites were retrieved. In the condition "unknown website," each site was chosen only once (in sum 104 different websites).

2.1.3. Measures

Participants were asked to rate the importance of certain aspects of website perception. Each participant was given four questions in a completely random order to reflect four web users' reactions during website use:

(1) Regarding the first impression of a website: How important are the following aspects from your point of view?
(2) Regarding the overall impression of a website: How important are the following aspects from your point of view?
(3) For revisiting a website repeatedly: How important are the following aspects from your point of view?
(4) For recommending a website to friends and acquaintances: How important are the following aspects from your point of view?

The participants rated each item with respect to content, usability (labeled with the common German term "Benutzerfreundlichkeit"), aesthetics (labeled as "design/aesthetics"), and a fourth category for other aspects (labeled "other"). Participants valued the importance of the four given aspects by assigning each a portion of 100 percent. As with the questions, the anchors were presented in random order to avoid position effects. If "other" was rated with more than zero percent, an open text field was shown in which participants could indicate the meaning of "other".

2.1.4. Procedures

After a welcome screen, participants were randomly assigned to one of three conditions. In the first condition (n = 119), participants were only asked to answer the four given questions. In the second condition (n = 107), participants were asked to visit a known website of their choice and answer the four questions with respect to that website. The third condition (n = 104) was similar to the second condition, except that participants were asked to select an unknown website. Participants could choose the unknown website, for example, via a spontaneous search word typed into a search engine. There were no significant demographic differences between the three groups.

If participants were instructed to retrieve a website we asked them to indicate which website they visited, followed by a control question asking if the website was known before this study. Then, each of the four questions described above was shown on a single page; those were presented to the participants in a random order. At the end of the study, the participants were asked to provide some demographic background information and to indicate their agreement with the use of their data for scientific purposes. Participation in this study took approximately five to six minutes.

2.2. Results

For each of the four given questions, we determined if the condition (not visiting a website or visiting a known or unknown website) had an effect. We calculated MANOVAs with the condition as independent and the four aspects of each question as dependent variables. There were no significant differences in the participants' ratings whether a website was visited or not or if a website was known or unknown. Thus, we combined the answers given for the three conditions for analysis (see table 1).

From the participants' perspective, content was the most relevant aspect for all four use aspects, with a rising tendency from first impression (34.87 %) to recommendation (54.97 %). Usability was rated as second in importance, except for first impressions, for which aesthetics was rated as second in importance (32.57 %). While usability was constantly rated around 25

to 30 percent, the perceived importance of aesthetics decreased from first impression (32.57 %) to overall impression (21.75 %) to revisit and recommendation (14.99 % and 16.25 %, respectively). Other aspects were of minor importance (≤ 3.24 %), and the aspects named in this category could mostly be assigned to content or usability (e.g., aspects like structure, overview, functionality, loading speed, safety, actuality, readability, seriousness, and comprehensibility).

Table 1. Importance ratings of core website aspects given by web users in percent

Aspect	Content		Usability		Aesthetics		Other	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
First impression	34.87	18.63	29.40	13.04	32.57	19.16	3.16	5.78
Overall impression	45.85	16.89	29.16	11.80	21.75	12.22	3.24	5.19
Intention to revisit	53.70	17.71	28.33	13.10	14.99	10.21	2.98	5.67
Recommendation	54.97	18.80	25.87	13.24	16.25	11.26	2.91	5.66

Note. Participants (N = 330) were asked to give percentage ratings.

2.3. Discussion

This short study shows that web users assign high importance to website content for all analyzed phases of website use. While this result may seem trivial, it is interesting that usability and aesthetic aspects garnered as well high importance ratings. Furthermore, the assumptions made by the participants are in line with empirical research stressing the importance of aesthetics in first impressions (e.g., Lindgaard et al., 2006 & 2011; Tractinsky et al., 2006; Thielsch & Hirschfeld, 2012). While the participants assumed little difference in the importance of usability, the relevance of content increased and was most meaningful for the recommendation of a website.

This short study had some limitations. First, the single item measurements used are prone to measurement errors. If a construct is multidimensional (as has been shown for example for aesthetics by Lavie & Tractinsky, 2004, and by Moshagen & Thielsch, 2010), single item measurement could lead to an imperfect measure of a construct. As we gave no specific definition of the constructs in forehand, the understanding of the labels might differ between participants. Thus, it is important to analyze the given constructs with standardized scales consisting of more than a single item to validate the results found.
Second, in this study, participants could (in two of the conditions) decide with complete freedom which website to visit, and there was no direct and detailed evaluation of the chosen website. This very general style of assessment was used intentionally in our study but may have led to superficial answers. Thus, it is important to analyze the actual use and differentiated evaluation of a given website with respect to all of the constructs in question. To perform such an evaluation was the aim of our second study.

3. Study 2: The influence of content, usability, and aesthetics on different user evaluation phases

The aim of the second study was to explore users' evaluations of websites by analyzing their rating behavior. Participants were asked to use one of four different websites and to rate the site with respect to not only content, usability, and aesthetics but also first and overall impressions, intention to revisit and intention to recommend the website. We used the ratings of these three constructs to predict users' reactions during website use.

3.1. Methods

3.1.1. Participants

The websites given were evaluated online by 300 participants aged between 18 and 80 years (M = 42.10, SD = 14.17). Of the participants, 168 were male (56 %), 131 were female, and one participant gave no information with respect to gender. Of the participants, 58.7 % hold a high school diploma or higher degree. They had been using the Internet on average for 9.25 years (Min = 2, Max = 20, SD = 3.53). Participants were selected via a German online panel and received credit points within the panel for completing the study (value was about 2.50 €).

3.1.2. Materials

The websites of four large electricity suppliers in Germany served as the stimulus materials (see appendix). The reasons for choosing websites of electricity suppliers were based on the idea to find a commercial surf scenario which most households can easily relate to and which makes it possible for us to ask for information search as well as for more complicated interaction surf tasks. To measure first impression, only a screenshot of the initial web page was presented at the beginning of the online study. Afterwards, participants were linked to the full functional website to explore it in a realistic scenario. The websites themselves were quite comparable (see appendix for means and standard deviations), as determined by examining the differences between the best- and worst-evaluated website in this study with respect to each construct and by calculating Cohen's *d.* According to the guidelines provided by Cohen (1988), standardized mean differences of 0.2, 0.5 and 0.8 are considered small, medium, and large effects, respectively. The four tested websites were relatively similar, as the differences between the best- and worst-rated websites with respect to content (d = 0.40), usability (d = 0.42), and aesthetics (d = 0.37) were rather moderate.

3.1.3. Measures

Participants evaluated the website in question with respect to content, usability, and aesthetics. As measure a standardized market research questionnaire called "NLR web scan" (Nordlight Research, 2008; see appendix) was used. The aim of the NLR web scan is to provide key performance indicators for the quality of Internet and e-media applications. It measures users perceptions of content (nine items), usability (seven items), and aesthetics (seven items) as well as additional items not pertinent for the present study. Participants answered the questions on a four-point Likert scale ranging from 1 ("disagree") to 4 ("agree"). An even numbered answer scale was used to grasp even small tendencies in participants' evaluations. If a participant was not able to answer an item, he or she could indicate this via a "don't know" option (which was coded as missing). The scale construction of the NLR web scan was based on factor analysis; Cronbach's α of the three scales ranged between .68 and .93 (Nordlight Research, 2008; see appendix).

First impression was measured on a seven-point scale (Cronbach's α was .91) with respect to six different aspects such as "To what extent did the starting page look: 'professional', 'interesting' and 'appropriate for an energy supplier'?"
Overall impression of the website in question was rated on a five-point scale ranging from "bad" to "excellent".
Intentions to revisit or to recommend a website to friends and acquaintances were asked on a four-point Likert scale (ranging from "disagree" to "agree").

3.1.4. Procedures

After a welcome screen and some demographic questions, participants were presented with the website of the local electricity supplier (at the time of the study, each of the four energy suppliers was active in a different region). In addition to simply viewing the website, the aim was to simulate the natural perception of the website during the survey. During the first stage, a screenshot of the website in question was presented, followed by the first impression items. During the second stage, participants performed three tasks on the given website: First, they were asked to freely explore the website. Then, they were instructed to search for information about electricity rates, and at last to calculate an individual tariff for their household. This allowed the participants to examine the website more closely and become more familiar with it. The tasks were followed by the item regarding the overall impression of the website.
In step three, participants were asked whether they would visit the website again and whether they would recommend the website to friends and acquaintances. After answering these questions the three NRL web scan scales measuring content, usability, and aesthetics perception were presented in randomized order.

3.2. Results

To identify the impact of the users' evaluations of content, usability, and aesthetics on their opinions and behaviors toward the website, a path analysis was performed with the statistics software MPlus (Muthén & Muthén, 2010) using the scale sum scores. In this analysis we allowed the predictors to correlate with each other, all three predictors were highly correlated (see table 2) as well as dependent variables (see table 3). The model (see figure 1) showed good fit indices: $\chi^2 = 6.39$; df = 4; p = .17; CFI = 1.00; TLI = .99; RMSEA = .05; SRMR = .01. For the CFI (Comparative Fit Index) and TLI (Tucker Lewis Index), values > .95 constitute good fit (Hu & Bentler, 1999); RMSEA (Root Mean Square Error of Approximation) and SRMR (Standardized Root Mean Square Residual) should be close to zero, whereas an RMSEA ≤ .05 and an SRMR ≤ .08 indicate close fit (Hu & Bentler, 1999).

Table 2. Intercorrelations of content, usability, and aesthetics in study 2 (values on the left of the slash), and study 3 (values on the right of the slash).

	Usability	Aesthetics
Content	.76** / .61**	.64** / .62**
Usability	-	.56**/ .64**

Note. * p < 0.05, **p < 0.001

Table 3. Intercorrelations of dependent variables in study 2 (values on the left of the slash), and study 3 (values on the right of the slash).

	Overall impression	Intention to revisit	Intention to recommend
First impression	.71** / .59**	.50* / .18**	.58** / .12**
Overall impression	.-	.57** / .21**	.69** / .17**
Intention to revisit		-	.68** / .64**

Note. * $p < 0.05$, **$p < 0.001$

The results confirm the importance of each of the three constructs, but the particular importance of each construct was dependent on the time and the focus of the evaluation. Figure 1 shows the impact of the three constructs on the first and overall impressions and the intention to revisit or recommend a website, arrows indicate regression coefficients. The path analysis showed that the decisive factor for the first impression of a website was aesthetics (.48, $p < 0.01$), with a considerably weaker influence of usability (.17, $p < 0.01$) and content (.15, $p < 0.05$). For the overall impression of a website, content (.40, $p < 0.01$) and aesthetics (.30, $p < 0.01$) were the most important factors, but usability (.16, $p < 0.01$) also had an impact.
Furthermore, Content played the decisive role in the intention to revisit a website (.65, $p < 0.01$) and the willingness to recommend a website (.69, $p < 0.01$), while neither usability nor aesthetics had a significant impact.

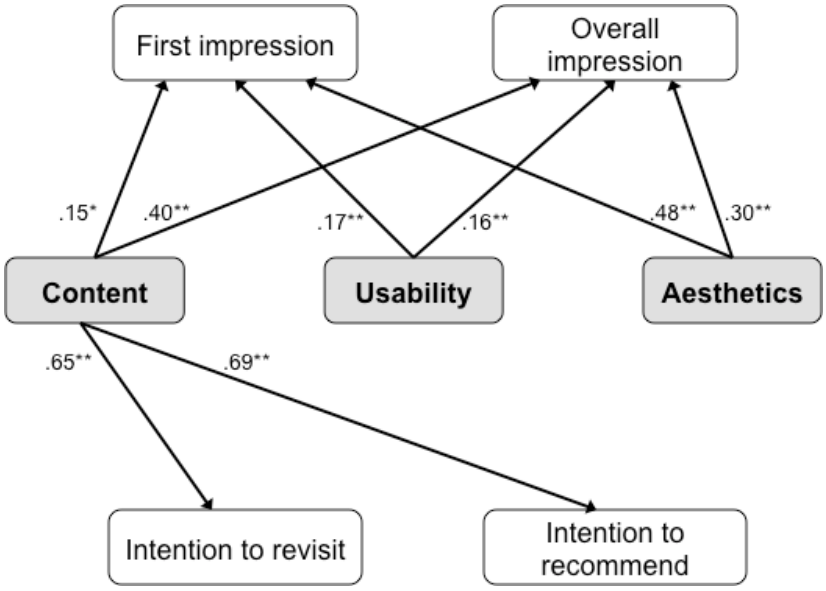

Figure 1. Path model of study 2
Note. Intercorrelations of independent variables are shown in table 2, of dependent variables in table 3.

3.3. Discussion

The results of our second study confirm the importance of the key website constructs content, usability, and aesthetics in users' website evaluation. Furthermore, we again found that the impact of each construct depended on the time and the focus of the evaluation. While all three constructs contribute to first and overall impressions of a website, aesthetics is especially

crucial for the first impression (as previously stressed empirically by Lindgaard et al., 2006 & 2011; Tractinsky et al., 2006; Thielsch & Hirschfeld, 2012; Tuch et al., 2012). However, users decide whether they are willing to visit a website again and recommend it based on content. This generally corresponds with the user estimates in study 1, but there are some differences (especially with respect to usability) that will be examined in the general discussion.
Because the builder of a website normally wants visitors to inspect the website more closely and remain on it, a convincing first impression and thus an aesthetic website are desirable. Furthermore, the activating impact of the website is of special importance because it causes visitors to visit the website again and, ideally, to recommend the website. To achieve the latter, the results of study 2 suggest that only content perception is critical. However, a high overall impression of a website depends on all three constructs, although the impact of usability perception is lower than the one of content and aesthetics.

While the path model met nearly all criteria for good fit (Chi-square test non-significant, CFI and TLI > .95, SRMR < .08) as well as these results are of high practical value, there are some drawbacks:
First, it is noteworthy that perceived usability had no influence on users' intentions to revisit or recommend a website. It seems obvious that poor usability should result in no recommendations and that few users would be willing to visit a totally unusable website again. Does this assumption reflect reality? In the given study, the tested websites were relatively similar; the difference between the best and worst rated websites with respect to usability was relatively small ($d = 0.42$), and the websites were quite usable. But, comparable absolute values and differences were also found for content ($d = 0.40$) and aesthetics ($d = 0.37$) – and content has a large influence on the intention to revisit or recommend. Thus, as the effect sizes for the differences between the tested websites were very comparable, the results found are highly valid for typical well-designed company websites like the ones used in our study. How users react when poorly designed websites are also tested is analyzed below in the next study.
Second, because this survey is a field study and not an experimental one, it has all the advantages and disadvantages associated with a field study. Potential problems with this study include the non-randomized assignment of the stimuli to the participants and the use of only a small set of different websites from one line of business. However, our aim was to get a realistic impression about the extent to which each construct influences website evaluation. The advantage of field studies is their closeness to reality. The procedure in this study is very close to natural usage and therefore very close to reality. This survey is also satisfactorily representative with regard to the age range of the participants. The results found need to be validated, however.
Third, there are significant correlations between the dependent as well as the independent variables in this study. Even as we controlled for such intercorrelations within the path analysis, one should not forget that neither analyzed constructs nor dependent variables are completely independent from each other. This reflects the fact, that web designers strive to optimize all three aspects in the same manner and that a prior use phase has influence on following ones.
Fourth, path analyses are like regressions analyses capable of identifying causal relationships between variables. But one has to keep in mind, that these relations are based on correlations and the theoretical approach of the build model. The fact that the found model is well fitting is a good hint for its validity, but there is still a need of a further validation. This could be done by a replication of this model or further causal analysis via experimental research designs.

In summary, this study provides a realistic impression of the decisive factors in a user's evaluation of a typical company website. The evaluation of each construct was dependent on the time and the focus of website use. But, a greater number of more diverse stimuli should be tested (ideally in randomized order), and the results should be replicated in a second sample with established measures from academic research. We fulfill these demands in study 3.

4. Study 3: Replication and validation

In our third study, we replicated the findings of study 2 in another sample with a large set of stimuli and validated research questionnaires. Therefore, we reanalyzed a data set of Moshagen and Thielsch (2010), which was partly used for scale validation. Those data had not been analyzed with respect to this research question before.

4.1. Methods

4.1.1. Participants

A total of 512 participants took part in this study; 347 were female (67.8 %). Ages ranged from 15 to 82 years (M = 30.50, SD = 10.61). The education level of 83.2 % of the participants was high school diploma or higher. On average, the subjects had been using the Internet for 10.25 years (Min = 2, Max = 25, SD = 3.29) and 15.72 hours a week (Min = 1, Max = 90, SD = 13.29). Participants took part voluntarily on an anonymous basis and received no compensation for completing the study.

4.1.2. Materials

A set of 42 websites from nine different content domains was used (see appendix). The websites were selected to represent a broad range of corporate and institutional websites in Germany, including, for example, corporate websites, e-commerce, e-recruiting, entertainment, and information sites. Readers are referred to Thielsch and Hirschfeld (2010) for a more detailed description of this categorization scheme. Due to technical difficulties, only three search engines could be included in this set; all of the other website categories each consisted of four to five websites. In contrast to the stimulus set used in study 2, this set contains a broad range of websites in terms of content, usability, and aesthetic appraisal (see appendix for means and standard deviations). The standardized mean differences between the best and worst evaluated websites were very large for each construct ($d_{content}$ = 3.67, $d_{usability}$ = 2.58, $d_{aesthetics}$ = 1.85).

4.1.3. Measures

We used several established measures for the evaluation of the website in question:
Perceived quality of content (PQC): This questionnaire (Thielsch, 2008; see appendix) consists of nine items on three subscales ("liking", "intelligibility", and "quality and use") representing a general factor, "quality of content". Cronbach's α of the three scales ranged between .71 and .90, and is .88 in the present study for the general factor. Factor and content validity of the PQC is demonstrated by Thielsch (2008).
Perceived website usability (PWU): The one-dimensional scale measuring perceived website usability (adapted based on Flavián, Guinalíu & Gurrea, 2006; see appendix) is a seven-item measure assessing perceived ease of use, ease of understanding and speed of information retrieval. Flavián and colleagues (2006) reported a Cronbach's α of .90 for this scale as well

as evidence for construct validity. Thielsch (2008) found a Cronbach's α of .95 for the adapted German version and demonstrated factor and convergent validity.
Visual aesthetics of websites inventory (VisAWI): This questionnaire (Moshagen & Thielsch, 2010; see appendix) uses 18 items to measure a general aesthetic factor consisting of four facets ("simplicity", "diversity", "color", and "craftsmanship"). Moshagen and Thielsch (2010) report Cronbach's α between .85 and .94, and provided evidence for convergent, divergent, discriminative, concurrent and experimental validity. Additional analyses as well as the items in German and English language can be found at Moshagen and Thielsch (in press).

First and overall impression: Participants were asked to rate their first impression of the website ("My first impression: I would mark the website with…") as well as their overall impression ("Altogether: I would mark the website with…") on a six-point grading scale ranging from "insufficient" to "very good", such a scale is very common in Germany.

Intention to revisit and to recommend: Three items were used to assess participants' intention to revisit the website being evaluated: (1) "I will visit the website again"; (2) "I will visit the website on a regular basis"; (3) "If I had interest in the content of the website in the future, I would consider visiting the website". The responses to these three items were averaged to form an index of the participants' intentions to revisit the website. Cronbach's α of this scale was .89 in the current study. Additionally, website recommendation was measured with one question: "I would recommend the website to my friends".

If not indicated differently above, participants were asked to indicate their level of agreement to each item of these questionnaires on seven-point Likert scales ranging from 1 ("strongly disagree") to 7 ("strongly agree").

4.1.4. Procedures

The study was announced as a survey for evaluating websites. After providing some demographic information, participants were randomly assigned to one website from the stimulus set and presented with a split screen. The questions regarding the website were presented in the smaller upper panel. At the beginning, participants were asked to rate their first impression of the website. Next, they answered the PQC, PWU, and VisAWI (and one other measure not pertinent to this study); these questionnaires used in the middle part of the study were given in random order, and the items within the questionnaires were randomized as well. Afterwards, the overall impression was rated on the same scale used at the beginning. At the end, the intention to revisit and to recommend was measured. On average, participants needed approximately 15 minutes to complete the study.

4.2. Results

As in study 2, a path analysis using the sum scores of the scales was performed with MPlus (Muthén & Muthén, 2010). Again predictors were allowed to correlate with each other, and all three predictors were highly correlated (see table 2). As in study 2, independent variables were correlated, but showed mostly small intercorrelations (see table 3). The path analysis widely confirmed the results of study 2 (see figure 2). The model showed very good fit indices: $\chi^2 = 3.08$; df = 2; p = .21; CFI = 1.00; TLI = .99; RMSEA = .03; SRMR = < .01.
As before, the crucial factor for the evaluation of first impression was aesthetics (.47, $p < .01$), significantly ahead of content (.26, $p < .01$), and usability (.17, $p < .01$). The importance of aesthetics is in agreement with the results of study 2. While usability had the same influence as before, content had a slightly higher influence on first impressions than in study 2.

For overall impression, the results were nearly the same, with a small decrease with respect to content (.30, $p < .01$) from study 2. Aesthetics again had a large influence on overall impression (.49, $p < .01$); indeed, the standardized path coefficient in this study is noticeable higher than in study 2.
When predicting the index regarding the willingness to visit the website again, content was clearly the most important construct (.64, $p < .01$). While usability and aesthetics had no significant influence at this point in study 2, a small but significant influence of aesthetics on the intention to revisit was found in study 3 (.11, $p < .01$).
The same pattern occurred when asking participants about their intention to recommend the website in question. As in study 2, content was by far the most important aspect (.61, $p < .01$), while usability had no influence and aesthetics had a small influence (.10, $p < .01$).
Thus, the results were mostly comparable to those found in study 2, especially with respect to usability, the importance of aesthetics for first impressions, and the importance of content for the intention to revisit or recommend a website. In study 3, aesthetics had a slightly higher influence on the latter two aspects as well as on overall impression.

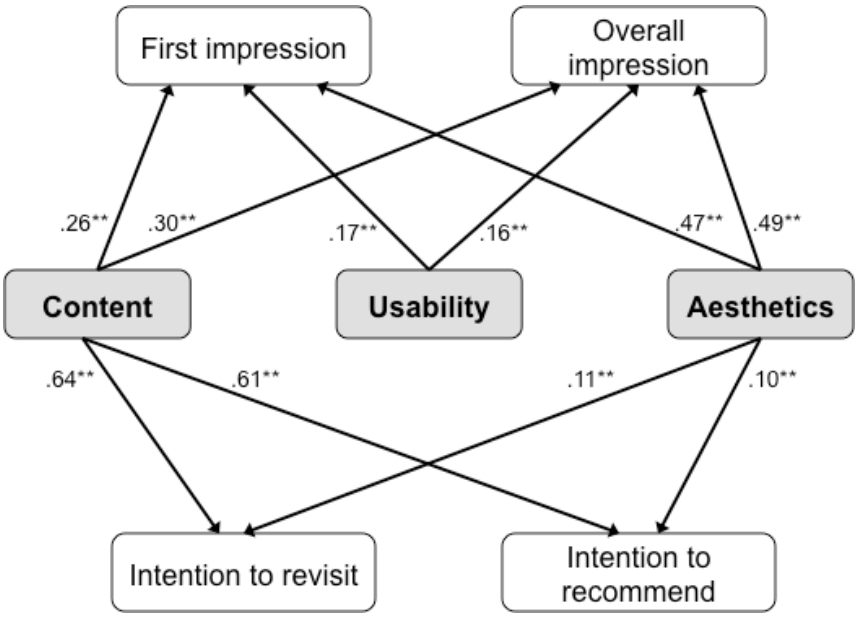

Figure 2. Path model of study 3
Note. Intercorrelations of independent variables are shown in table 2, of dependent variables in table 3.

4.3. Discussion

In this study, we largely replicated the results from study 2 with a large set of websites, a different sample, and different measures. The path model met all criteria for good fit (Chi-square test non-significant, CFI and TLI > .95; RMSEA < .05, SRMR < .08). Again, all three constructs contributed to the first and overall impressions of a website. The main results were replicated: aesthetics had a high influence on deliberate first impression, and content had a very high impact on the intention to revisit or recommend a website. Furthermore, in this study, content had a slightly higher impact on first impressions, while aesthetics had a small but significant impact on intention to revisit and intention to recommend. As before, usability had a small to mediocre influence on first and overall impressions but was not relevant for intention to revisit or recommend.

The latter result is very interesting with respect to the question that arose in study 2: Should a really unusable website lead to a decreased intention to revisit or recommend the website? While this may be true if a very unusable website is tested directly, there is no such effect in our data. The tested websites showed a wide range of apparent usability, and the correlation between the evaluation of usability and the intention to revisit or recommend was even lower when we solely analyzed the evaluated websites that were rated as unusable. In the future, it may be interesting to examine websites classified in terms of inherent usability. Based on the given data, we can assume that perceived usability has a significant impact on first and overall impressions, but the influence of content and aesthetic perception is persistently higher at later phases of website use.

Furthermore, there are some small but interesting differences in the results when compared to study 2. Using a broad set of stimuli, content had slightly more influence on first impressions, while aesthetics had the same effect on the intention to revisit or recommend. The latter result is in agreement with prior research investigating the impact of aesthetics on the intention to revisit a website (Mahlke, 2002; Yoo & Donthu, 2001). Because results in the literature differ when using small sets of stimuli and static screenshots (e.g., Cober et al., 2003), it seems important to stress that this effect may not occur with every single website and may be masked by the unique content of a specific website. This would also explain why no impact of aesthetics on intention to revisit or recommend was evident in our second study, in which we used a small and thematically narrow set of websites. But, an alternative explanation could be given on the difference between users' tasks between study 2 and 3. Van Schaik and Ling (2009) found influences of different task modes (action versus goal mode) on aesthetics perceptions. Thus, further research is needed on such influences on the found model. Therefore, a possibly substantial limitation of study 3 is that participants did not fulfill any specific tasks like those given in study 2. Participants in study 3 were instructed to explore the website in question, saw a fully functional version of the website itself, and showed typical usage duration times for such a website test. But, future research should control for different use scenarios, for example like done by van Schaik and Ling (2009).
Furthermore, as in study 2, there are significant correlations between the dependent as well as the independent variables in this study one should keep in mind while interpreting the results (even as they are statistically controlled in the path analysis). And again it is to mention, that path analyses are only capable of identifying causal relationships based on theoretical reasoning. The supposed theoretical model was replicated and showed excellent fit to the empirical data in study 3, but further experimental validation will be useful.

In summary, we confirmed the results of the second study and extended them. We used a large set of websites and established measures and questioned a large sample. Although the weaknesses of study 2 were addressed, we obtained a mostly similar pattern of results. Thus, we expect the effects found will be stable in general when business and institutional websites are tested.

5. General Discussion

While content is the most important aspect of a website from a user's perspective (as stated in study 1), we showed in a series of two studies that this depends on the time and the focus of the website evaluation. In doing so, we may shed some light on the previously partially mixed

results regarding the impact of content, usability, and aesthetics by showing how the impact of each construct is dependent on users' evaluations.
In study 1, all three constructs – content, usability, and aesthetics – were stated as important. Furthermore, when asked directly users named content as most important. Usability was rated a distant second, while aesthetics was third, except for first impressions, for which aesthetics was rated second in importance. In addition, it is noteworthy that only few additional aspects were named in this study, and most of those were assignable to content and usability.
In study 2, we used a survey design that was very similar to a user's natural use of a website and asked for several ratings on a set of four websites. With path analysis, we showed that aesthetics had the highest impact on deliberate first impressions, while all three constructs contributed to first and overall impressions. Furthermore, only content was crucial for the activating impact of a website, such as the intention to revisit or recommend the site.
We replicated those results in study 3 with a different sample, different measures and a large set of 42 websites. Our proposed model showed excellent fit to the empirical data in study 3.

In general, studies 2 and 3 showed that aesthetics has a greater impact on the evaluation of a website than one might expect from directly asking users (as we did in study 1). Thus, participants may underestimate the importance of aesthetic factors on first and overall impressions. Nevertheless, aesthetics likely play an important role in users' perceptions and evaluations, as well as in online product presentations and branding. The users' assumptions in study 1 regarding the impact of content and aesthetics on intentions to revisit or recommend were confirmed in study 3, where content was obviously the absolutely most important stimulus for returning to or recommending a website. The importance of aesthetics for all analyzed phases of website use was still underestimated in a direct evaluation. From current research on website aesthetics we know, that this construct is processed very early in visual perception (Leder et al., 2004; Thielsch & Hirschfeld, 2010 & 2012; Tuch et al., 2012). Moreover, one may suggest that the processing of content and especially usability occurs slightly later in perception (see also Lee & Koubek, 2012); this may explain the high impact of aesthetics on immediate and deliberate first impressions. Several authors have discussed what determines first impressions, especially with respect to aesthetics. From a general design perspective, visual complexity in particular appears to have an impact on users' reactions towards a website (e.g., Geissler, Zinkhan & Watson, 2006; Nadkarni & Gupta, 2007; Tuch et al., 2009). In addition to visual complexity, aspects of order and prototypicality seem to be very important for users' first impressions (Deng & Poole, 2010; Tuch et al., 2012). From a more cognitive perspective, such design elements are reflected in the different spatial frequencies perceived by a user, and recent research has shown the importance of low spatial frequencies for immediate first impressions (Thielsch & Hirschfeld, 2010 & 2012).

Thus, while we showed that content is the most important aspect of a website in general and that content has an impact from deliberate first impressions to the point of recommendation, one should not forget that a website is more than a collection of content pieces. Usability is relevant for first and overall impressions, and aesthetic factors influence the whole perception of a website from first impression until recommendation. However, in contrast to content perception, aesthetics is most important at first sight and less important for more complex decisions like the willingness to revisit or recommend the website. As this activating consequence of aesthetics does not occur in our second study or in some of the literature (e.g., Cober et al., 2003) when small sets of stimuli are used, it is important to stress that this effect may not occur on every single website, may be masked by content aspects of a specific website or be influenced by specific tasks users have to fulfill (see van Schaik & Ling, 2009).

Perceived usability is of less importance than estimated by web users in study 1. Users usability perception has some impact on first and overall impressions but not on the intentions to revisit and to recommend in studies 2 and 3, respectively. We showed that this was not due to restricted variance in the usability evaluations (see discussion of study 3). Thus, perceived usability may not have an impact in this context, and more inherent measures may be needed to evaluate usability. We will discuss this in the next section.

5.1. Limitations and further research

Some limitations should be considered when interpreting the results of our present research. First, we only examined business and institutional websites and excluded private websites from our analyses in studies 2 and 3. Given the myriad of existing websites, we tested only a limited sample of stimuli. However, we successfully replicated the results with different stimuli, different measures and a different sample in study 3, and thus we expect our results to be applicable to other types and kinds of websites. Additional replication with different stimuli would be ideal to validate our results.
Second, all of the tested participants and the stimuli used shared the same cultural background. There is some evidence for the existence of cultural aspects of website content (Robbins & Stylianou, 2003), and several authors also stress the importance of cultural factors in website usability and design perception (e.g., Marcus & Gould, 2000; Simon, 2001; Sun, 2001; Tractinsky, 1997), especially for web design aspects like color and images (Cyr et al., 2009; Cyr, Head & Larios, 2010). The extent to which our findings are prone to cultural differences should be analyzed by a cross-cultural approach.
Third, due to the large amount of websites tested, we decided to analyze only the perceived/apparent usability. Although there are correlations between apparent and inherent variables of usability (Tractinsky, 1997), inherent usability seems to reflect some other aspects of graphical user interfaces (Fu & Salvendy, 2002; Kurosu & Kashimura, 1995; Tractinsky, 1997). Thus, future replication of these results should include an evaluation of websites in terms of inherent usability, which may yield interesting insights beyond the rather visual impression of apparent usability. Additionally, as mentioned in the discussion of study 3, usability perception might be influenced by task characteristics. This could be by tested by giving participants different use scenarios (like done by van Schaik and Ling, 2009).
Fourth, it will be interesting to test the causal relationships and extend our results using experimental designs and – as far as applicable for the constructs in question – more objective measures like reaction times, task duration times or objective attributes of stimuli like word count, link depth and so on. Our studies examined direct website evaluations and users' beliefs and used website evaluations to predict website use. The next step should be an experimental validation and extension of our results. Such a validation should take different use scenarios and use tasks into account to clear their influence on the found model. The analyzed use phases could be extended by longitudinal approaches, taking repeated use of a website into account. In doing so, it is possible to take further use phases into account, like revisiting a website or specific behavior like buying, downloads etc..

5.2. Conclusion

When building a successful website, designers and content creators have to work hand in hand. Aesthetics is most important for the first impression, while content is decisive for the entire duration of use. Content perception is of special significance when the users are asked to get active themselves – like for example in recommending a website or revisiting it. Our

results for website usage and evaluation bring to mind the elaboration likelihood model (Petty & Cacioppo, 1986): Aesthetics, which may be processed primarily on a peripheral route, is especially important for spontaneous reactions and overall impressions, while the user is reading and navigating the website. Content becomes more and more important during use and is probably centrally processed. When higher cognitive processes are necessary, such as contemplating the applicability of a website for recommending to friends, the influence of aesthetics weakens but is not eliminated. This is in agreement with the current definition of aesthetics, which is assumed to not be mediated by intervening reasoning (Moshagen & Thielsch, 2010, p. 690).

Let us think of a website user as a customer: he or she will be attracted by good aesthetics and bound by interesting and appealing content presented in a usable manner. When contemplating revisiting the website or recommending it, content is crucial, but the user will still be influenced by aesthetic factors. Aesthetics serves as a door opener as well as a reinforcement to the users. Content builds the vital base of a website, usability is essential, and aesthetics is much more than "just" decoration.

Acknowledgments

The authors would like to thank Angela Buchholz for her help in analyzing data, Morten Moshagen for his permission to reanalyze the data used in study 3 and Gerrit Hirschfeld for very helpful comments on prior versions of the manuscript.

References

Agarwal, R. & Venkatesh, V., 2002. Assessing a firm's web presence: A heuristic evaluation procedure for the measurement of usability. Information Systems Research, 13, 168-186.

Aladwani, A.M. & Palvia, P.C., 2002. Developing and validating an instrument for measuring user-perceived web quality. Information & Management, 39, 467–476.

Allen, D.G., Mahto, R.V. & Otondo R.F., 2007. Web-based recruitment: Effects of information, organizational brand and attitudes toward a web site on applicant attraction. Journal of Applied Psychology, 92, 1696 -1708.

Bargas-Avila, J.A. & Hornbæk, K., 2011. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. *Proceedings of the 2011 annual conference on Human factors in computing systems* (pp. 2689–2698). ACM.

Berners-Lee, T., 2001. Publishing on the semantic web. Nature, 410, 1023–1024.

Braddy, P.W., Meade, A.W., Michael, J.J. & Fleenor, J.W., 2009. Internet recruiting: Effects of website content features on viewers' perceptions of organizational culture. International Journal of Selection and Assessment, 17, 19-34.

Cober, R.T., Brown, D.A., Levy, P.E., Cober, A.B. & Keeping, L.M., 2003. Organizational web sites: Web site content and style as determinants of organizational attraction. International Journal of Selection and Assessment, 11(2/3), 158-169.

Cohen, J., 1988. Statistical power analysis for the behavioural sciences. Hillsdale: Erlbaum.

Cyr, D., Head, M. & Larios, H., 2010. Colour appeal in website design within and across cultures: A multi-method evaluation. International Journal of Human-Computer Studies, 68(1-2), 1-21.

Cyr, D., Head, M., Larios, H. & Pan, B., 2009. Exploring human images in website design: A multi-method approach. MIS Quarterly, 33, 539–566.

D'Ambra, J. & Rice, R.E., 2001. Emerging factors in user evaluation of the World Wide Web. Information & Management, 38, 373–384.

De Angeli, A., Sutcliffe, A. & Hartmann, J., 2006. Interaction, usability and aesthetics: What influences users' preferences? Proceedings of the 6th conference on Designing Interactive systems, p. 271–280. New York: ACM.

De Wulf, K., Schillewaert, N., Muylle, S., & Rangarajan, D., 2006. The role of pleasure in web site success. Information & Management, 43, 434-446.

Deng, L. & Poole, M.S., 2010. Affect in web interfaces: A study of the impacts of web page visual complexity and order. MIS Quarterly, 34, 711–730.

Dutta-Bergman, M.J., 2004. The impact of completeness and web use motivation on the credibility of e-Health Information. Journal of Communication, 54, 253-269.

Engholm, I., 2002. Digital style history: The development of graphic design on the Internet. Digital Creativity, 13, 193-211.

Eysenbach, G. & Köhler, C., 2002. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *British Medical Journal*, *324*, 573–577.

Fauzi, F., & Belkhatir, M., 2010. A user study to investigate semantically relevant contextual information of WWW images. International Journal of Human-Computer Studies, 68, 270-287.

Flavián, C., Guinalíu, M. & Gurrea, R., 2006. The role played by perceived usability, satisfaction and consumer trust on website loyalty. Information & Management, 43, 1–14.

Fu, L.M. & Salvendy, G., 2002. The contribution of apparent and inherent usability to a user's satisfaction in a searching and browsing task on the Web. Ergonomics, 45, 415-424.

Geissler, G.L., Zinkhan, G.M. & Watson, R.T., 2006. The influence of home page complexity on consumer attention, attitudes, and purchase intent. Journal of Advertising, 35, 69–80.

Hartmann, J., Sutcliffe, A. & De Angeli, A., 2008. Towards a theory of user judgment of aesthetics and user interface quality. ACM Transactions on Computer-Human Interaction, TOCHI, 15, Article 15.

Hassenzahl, M. & Monk, A., 2010. The inference of perceived usability from beauty. Human-Computer Interaction, 25, 235-260.

Hassenzahl, M. & Tractinsky, N., 2006. User experience - a research agenda. Behaviour & Information Technology, 25, 91-97.

Hendler, J., 2003. Science and the Semantic Web. Science, 299(5606), 520-521.

Hendler, J., Shadbolt, N., Hall, W., Berners-Lee, T. & Weitzner, D., 2008. Web science: An interdisciplinary approach to understanding the web. Communications of the ACM, 51, 60–69. ACM.

Hornbæk, K., 2006. Current practice in measuring usability: Challenges to usability studies and research. International Journal of Human-Computer Studies, 64, 79-102.

Hu, L., & Bentler, P. M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling, 6, 1-55.

Huizingh, E., 2000. The content and design of web sites: An empirical study. Information & Management, 37, 123–134.

Ilmberger, W., Schrepp, M. & Held, T., 2008. What kind of cognitive process causes the relationship between aesthetics and usability. In: Holzinger, A. (Ed.): *USAB 2008, LNCS 5298*, p. 43-54.

ISO, 2009. ISO FDIS 9241-210: Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems. Geneva: International Organization for Standardisation.

ISO, 2006. ISO 9241: Ergonomics of Human-System Interaction – Part 151: Guidance on World Wide Web Interfaces. Geneva: International Organization for Standardisation.

ISO, 1998. ISO 9241: Ergonomic requirements for office work with visual display terminals, VDTS) – part 11: Guidance on usability. Geneva: International Organization for Standardisation.

Ivory, M.Y. & Megraw, R., 2005. Evolution of web site design patterns. ACM Transactions on Information Systems, 23, 463-497.

Jansen, B. & Spink, A., 2006. How are we searching the World Wide Web? A comparison of nine search engine transaction logs. Information Processing & Management, 42, 248-263.

Kang, Y. & Kim, Y., 2006. Do visitors' interest level and perceived quantity of web page content matter in shaping the attitude toward a web site? Decision Support Systems, 42, 1187-1202.

Kim, S.Y. & Lim, Y.J., 2001. Customers' perceived importance of and satisfaction with Internet shopping. Electronic Markets, 11, 148–154.

Kurosu, M. & Kashimura, K., 1995. Apparent usability vs. inherent usability: experimental analysis on the determinants of the apparent usability. In: Conference companion on Human factors in computing systems, p. 292-293. New York: ACM Press.

Lavie, T. & Tractinsky, N., 2004. Assessing dimensions of perceived visual aesthetics of web sites. International Journal of Human - Computer Studies, 60, 269–298.

Leder, H., Belke, B., Oeberst, A. & Augustin, D., 2004. A model of aesthetic appreciation and aesthetic judgements. British Journal of Psychology 95, 489-508.

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.

Lee, S. & Koubek, R.J., 2012. Users' perceptions of usability and aesthetics as criteria of pre- and post-use preferences. *European J. of Industrial Engineering*, *6*, 87.

Lindgaard, G., Fernandes, G., Dudek, C. & Brown, J., 2006. Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & Information Technology 25,* 115-126.

Lindgaard, G., Dudek, C., Sen, D., Sumegi, L., and Noonan, P., 2011. An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. *ACM Transactions on Computer-Human Interaction, 18*, 1-30.

Liu, C. & Arnett, K.P., 2000. Exploring the factors associated with Web site success in the context of electronic commerce. *Information & Management, 38*, 23-33.

Mahlke, S., 2002. Factors influencing the experience of website usage. In: Proceedings of the CHI'02 Extended Abstracts on Human Factors in Computing Systems. ACM, Minneapolis, pp. 846-847.

Marcus, A. & Gould, E.W., 2000. Crosscurrents: Cultural dimensions and global web user-interface design. *Interactions, 7*, 32–46.

Metzger, M., 2007. Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, *58*, 2078–2091.

Moshagen, M., Musch, J. & Göritz, A.S., 2009. A blessing, not a curse: Experimental evidence for beneficial effects of visual aesthetics on performance. *Ergonomics, 52*, 1311-1320.

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

Moshagen, M. & Thielsch, M.T., in press. A short version of the visual aesthetics of websites inventory. *Behaviour & Information Technology*.

Muthén, L.K., & Muthén, B.O., 2010. Mplus. User's Guide. (6th ed.). Los Angeles, CA: Muthén & Muthén.

Nadkarni, S. & Gupta, R., 2007. A task-based model of perceived website complexity. *MIS Quarterly, 31*, 501-524.

Nordlight Research, 2008. NLR web scan Instrumentenentwicklung [NRL web scan development]. Hilden, Germany: Rafael Jaron.

Ou, C.X. & Sia, C.L., 2010. Consumer trust and distrust: An issue of website design. *International Journal of Human-Computer Studies*, *68*, 913–934.

Palmer, J.W., 2002. Web site usability, design, and performance metrics. Information Systems Research, 13, 151-167.

Petty, R.E. & Cacioppo, J.T., 1986. The elaboration likelihood model of persuasion. Advances in experimental social psychology, 19, 123-205.

Pfieffelmann, B., Wagner, S.H. & Libkuman, T., 2010. Recruiting on corporate web sites: Perceptions of fit and attraction. International Journal of Selection and Assessment, 18, 40-47.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., & Podsakoff, N.P., 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903.

Rains, S.A. & Karmikel, C.D., 2009. Health information-seeking and perceptions of website credibility: Examining Web-use orientation, message characteristics, and structural features of websites. Computers in Human Behavior, 25, 544-553.

Ranganathan, C. & Ganapathy, S., 2002. Key dimensions of business-to-consumer web sites. Information & Management, 39, 457-465.

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

Robbins, S.S. & Stylianou, A.C., 2003. Global corporate web sites: An empirical investigation of content and design. *Information & Management, 40*, 205–212.

Scharrer, L., Bromme, R., Britt, M.A., & Stadtler, M., 2012. The seduction of easiness: How science depictions influence laypeople's reliance on their own evaluation of scientific information. *Learning and Instruction, 22*, 231–243.

Schenkman, B.N. & Jönsson, F.U., 2000. Aesthetics and preferences of web pages. Behaviour & Information Technology, 19, 367-377.

Shneiderman, B. & Plaisant, C., 2009. Designing the user interface: Strategies for Effective Human-Computer Interaction, 5th ed.. Boston: Addison-Wesley.

Soobrah, R. & Clark, S.K., 2011. Your patient information website: how good is it? *Colorectal disease*, 14, 90–94.

Sillence, E., Briggs, P., Harris, P.R., & Fishwick, L., 2007. How do patients evaluate and make use of online health information? *Social Science & Medicine*, *64*, 1853–62.

Simon, S.J., 2001. The impact of culture and gender on web sites: An empirical study. ACM SIGMIS Database, 32, 18–37.

Sun, H., 2001. Building a culturally-competent corporate web site: An exploratory study of

cultural markers in multilingual web design. In Proceedings of the 19th annual international conference on computer documentation, pp. 95–102. New York: ACM Press.

Sutherland, L.A., Wildemuth, B., Campbell, M.K., & Haines, P.S., 2005. Unraveling the web: an evaluation of the content quality, usability, and readability of nutrition web sites. *Journal of nutrition education and behavior*, *37*, 300–305.

Tarasewich, P., Daniel, H.Z., & Griffin, H.E., 2001. Aesthetics and web site design. *Quarterly Journal of Electronic Commerce, 2*, 67-81.

Thielsch, M.T., 2008. *Ästhetik von Websites [Aesthetics of websites]*. Münster: MV Wissenschaft.

Thielsch, M.T. & Hirschfeld, G., 2010. High and low spatial frequencies in website evaluations. Ergonomics, 53, 972-978.

Thielsch, M.T. & Hirschfeld, G., 2012. Spatial frequencies in aesthetic website evaluations – explaining how ultra-rapid evaluations are formed. Ergonomics, 55 (7), 731-742.

Thüring, M., & Mahlke, S., 2007. Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology*, *42*, 253 – 264.

Tractinsky, N., 1997. Aesthetics and apparent usability: Empirically assessing cultural and methodological issues. CHI 97 proceedings, 115–123.

Tractinsky, N., 2004. Toward the study of aesthetics in information technology. In J. I. DeGross (Ed.), *Proceedings of the 25th ICIS* (pp. 771–780). Washington, DC: ICIS Press.

Tractinsky, N., Cokhavi, A., Kirschenbaum, M. & Sharfi, T., 2006. Evaluating the consistency of immediate aesthetic perceptions of web pages. International Journal of Human-Computer Studies 64, 1071-1083.

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, 703-715.

Tuch, A.N., Presslaber, E.E., Stöcklin, M., Opwis, K. & Bargas-Avila, J.A., 2012. The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. International Journal of Human-Computer Studies, 70(11), 794-811.

Van der Heijden, H., 2003. Factors influencing the usage of websites: the case of a generic portal in The Netherlands. *Information & Management*, *40*, 541–549.

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.

Watanabe, T., 2009. Experimental evaluation of usability and accessibility of heading elements. *Disability and Rehabilitation: Assistive technology*, *4*, 236–47.

Yoo, B. & Donthu, N., 2001. Developing a scale to measure the perceived quality of an Internet shopping site, SITEQUAL. Quarterly Journal of Electronic Commerce 2, 31–46.

Zhao, W., Massey, B., Murphy, J. & Fang, L., 2003. Cultural dimensions of website design and content. Prometheus, 21, 74-84.

Appendix 1. Content, usability, and aesthetics measures used in study 2 and 3.

	Study 2	Study 3
Content	Information on the website are useful to answer my questions. I can easily find out the costs and prices for the company's products. Product descriptions are meaningful an informative. Information appear too much like advertisements. (r) Apart from product information the website offers a lot of useful tips and services. Reading the headlines I know immediately which content is offered on that page. Overall texts are too long. (r) Texts contain to many special terms or abbreviations that are hard to understand. (r) Contact information on the website makes it easy to get in touch with the company.	The website piques my interest. I find the content of the website pleasant. I enjoy reading the website. The particular sentences are easy to read. The texts provide me information in a clear and concise manner. I find the language used in the texts to be current and easy to understand. The information is of great value. I find the information on the website to be useful. Contents of the website seem so important to me, that I would print or save them.
Cronbachs alpha original / current study; Source	α = .72 / .84 Nordlight Research (2008)	α = .85 / .88 Thielsch (2008)
Usability	Most important topics and services can be selected directly at welcome page. Navigation elements are distinct from other website-elements (e.g. text areas). The functional elements (links, buttons,) are placed where I expect them to be. By means of the link names in navigation I know exactly to which content I will get. The search-engine gives constructive results. While surfing the website gives clear information on which site I currently am. Important information can be found using only a few clicks.	In this website everything is easy to understand. This website is simple to use, even when using it for the first time. It is easy to find the information I need from this website. The structure of this website is easy to understand. It is easy to orient oneself within this website. The organization of the contents of this site makes it easy for me to know where I am. One can quickly reach the searched information.
Cronbachs alpha original / current study; Source	α = .93 / .93 Nordlight Research (2008)	α = .95 / .96 Adapted from Flavian et al. (2006), German version at Thielsch (2008)
Aesthetics	The website seems clearly arranged and not cluttered. To me there is too much flickering and flashing on the website. (r) The websites design appears professional. The fonts color provides an ideal figure ground contrast. Pictures and corresponding contents fit well. Website colors are comfortable to look at. The font is too small. (r)	The layout appears too dense. (r) The layout is easy to grasp. Everything goes together on this site. The site appears patchy. (r) The layout appears well structured. The layout is pleasantly varied. The layout is inventive. The design appears uninspired. (r) The layout appears dynamic. The design is uninteresting. (r) The color composition is attractive. The colors do not match. (r) The choice of colors is botched. (r) The colors are appealing. The layout appears professionally designed. The layout is not up-to-date. (r) The site is designed with care. The design of the site lacks a concept. (r)
Cronbachs alpha original / current study; Source	α = .68 / .82 Nordlight Research (2008)	α = .94 / .94 Moshagen & Thielsch (2010), German version in Moshagen & Thielsch (in press)

Note. Negatively-keyed items are indicated by (r) and are reverse-scored.

Appendix 2. Means and standard deviations of content, usability, and aesthetics ratings for each website in study 2.

	Content		Usability		Aesthetics	
Website URL	*M*	*SD*	*M*	*SD*	*M*	*SD*
http://www.enbw.com/content/de	2.94	0.51	3.17	0.72	3.23	0.54
http://www.eon.de	2.72	0.60	2.89	0.75	2.68	0.62
http://www.rheinenergie.com/de	2.97	0.61	3.20	0.69	3.01	0.67
http://www.rwe.de	2.82	0.66	2.90	0.76	3.15	0.65
Total mean	2.88	0.61	3.07	0.73	3.01	0.66

Note. Participants (N = 300) answered on a four-point Likert scale ranging from 1 (“disagree”) to 4 (“agree”).

Appendix 3. Means and standard deviations of content, usability, and aesthetics ratings for each website in study 3.

	Content		Usability		Aesthetics	
Website category and website URL	*M*	*SD*	*M*	*SD*	*M*	*SD*
Download & software websites						
http://de.selfhtml.org	2.88	1.30	3.56	1.57	2.87	0.72
http://www.freeware.de	3.69	1.03	3.73	1.23	3.98	1.01
http://www.heise.de/software	3.88	1.50	3.78	1.50	3.70	1.02
http://www.java.de	3.53	1.06	4.21	1.76	3.79	1.44
E-commerce websites						
http://www.amazon.de	4.56	1.40	4.76	1.67	4.09	1.58
http://www.ebay.de	3.95	0.91	4.52	1.85	4.33	1.30
http://www.golop.de	3.34	0.99	4.33	1.31	3.44	0.88
http://www.kelkoo.de	3.31	1.06	3.93	1.18	3.51	1.15
http://www.mobile.de	3.53	0.78	4.50	1.30	3.98	0.93
Entertainment websites						
http://de.youtube.com	3.92	1.40	4.51	1.67	4.49	1.21
http://www.myspass.de	3.42	1.20	4.04	1.48	4.25	1.16
http://www.promi-star.de	3.50	1.04	5.31	1.34	3.79	1.16
http://www.spiele-zone.de	3.98	1.18	5.70	1.13	4.90	1.33
http://www.wow-europe.com/de/index.xml	2.72	0.83	4.17	1.56	4.10	1.19
E-learning websites						
http://ihk.elearningspace.de	3.83	0.97	4.24	1.96	3.70	1.02
http://www.bildung.at	3.58	1.02	4.20	1.87	4.19	1.30
http://www.elearningeuropa.info	3.61	1.47	4.53	1.37	4.78	1.42
http://www.moodle.de	4.41	0.71	5.24	0.85	4.40	1.06
http://www.sgd.de	4.14	1.15	4.80	1.63	3.85	1.46
E-recruiting websites						
http://www.jobpilot.de	4.18	1.00	5.01	1.36	4.14	1.20
http://www.jobscout24.de	4.16	1.26	4.89	1.43	4.13	1.29
http://www.jobware.de	4.03	0.81	5.00	0.84	4.51	0.31
http://www.monster.de	3.45	1.06	3.45	1.27	3.42	1.41
http://www.stepstone.de	3.15	1.32	2.61	1.26	2.72	1.00
Information and news websites						
http://de.wikipedia.org	5.60	0.90	4.91	1.41	4.63	1.20
http://www.faz.net	4.67	0.95	5.11	1.26	4.74	1.44
http://www.n-tv.de	4.07	0.85	3.79	1.07	3.88	0.74
http://www.sueddeutsche.de	4.50	0.89	4.83	1.83	4.39	1.66
http://www.stern.de	3.95	1.12	4.09	1.55	3.96	0.99
Web portals						
http://www.aol.de	2.78	0.65	3.71	1.50	3.28	1.00
http://www.freenet.de	2.45	0.82	2.81	1.15	3.05	1.01
http://www.gmx.de	3.45	1.03	4.08	1.42	4.08	0.98
http://www.t-online.de	2.85	0.66	2.95	1.30	2.96	0.71
http://www.web.de	3.51	0.72	4.07	1.50	3.86	0.95
Corporate websites						
http://www.bayer.de	3.44	1.07	3.26	1.58	4.20	1.26
http://www.bertelsmann.de	3.66	0.94	4.49	1.45	4.25	0.98
http://www.daimler.com/dccom/home/de	3.76	0.94	5.14	0.81	4.86	1.02
http://www.dseurope.ag	3.04	1.39	4.32	1.41	4.22	0.95
http://www.hochtief.de	3.58	0.98	4.21	1.48	4.49	1.20
Search engines						
http://www.de.altavista.com	3.62	1.06	5.49	1.00	4.09	1.29
http://www.fireball.de	3.84	0.82	5.43	1.53	3.57	1.12
http://www.google.de	4.64	1.13	6.04	1.16	4.55	1.12
Total mean	*3.69*	*1.17*	*4.35*	*1.58*	*3.99*	*1.21*

Note. Participants (N = 512) answered on a seven-point Likert scale ranging from 1 (“strongly disagree”) to 7 (“strongly agree”).