Usability News

October 2009, *Vol. 11 Issue 2*

| Volume 11 Issue 2 | Past Issues | A-Z List |

Usability News is a free web newsletter that is produced by the Software Usability Research Laboratory (SURL) at Wichita State University. The SURL team specializes in software/website user interface design, usability testing, and research in human-computer interaction.
Barbara S. Chaparro, Editor

Visual Appeal vs. Usability: Which One Influences User Perceptions of a Website More?

Christine Phillips* & Barbara S. Chaparro

Summary. This study examines the effects of visual appeal and usability on user performance and satisfaction with a website. Users completed search and exploratory tasks on sites which varied in visual appeal (high and low) and usability (high and low). Results indicate that first impressions are most influenced by the visual appeal of the site. Users gave high usability and interest ratings to sites with high appeal and low usability and interest ratings to sites with low appeal. User perceptions of a low appeal website were not significantly influenced by the site's usability even after a successful experience with the site. Another finding suggested users actively searching for information were more aware of usability issues than users who simply explored a site.

INTRODUCTION

Perceived usability of a website by a user is often more influential than the actual product efficiency and ease of use. The visual appeal of an interface appears to play a role in the user's rating on perceived usability. For example, Kurosu & Kashimura (1995) found that users reported an aesthetically appealing ATM interface easier to use than an unappealing or bland ATM interface. Tractinsky, Katz & Ikar (2000) also investigated ATM interfaces and found that the more aesthetically pleasing interfaces were judged to be more usable, despite actual usability. Brady and Phillips (2003) found that participants ranked websites with good balance and color as more usable than websites with unbalanced and poorly selected color schemes.

The importance of visual appeal of websites has shown that aesthetics play an important role in first impressions of a website and that they may form in as little as 50ms (Lindgaard, Fernandes, Dudek, & Brown, 2006). The role of first impressions on a website is very important as there is evidence that they may be long-lasting. Lindgaard & Dudek (2002) suggest that this occurs due to the confirmation bias (Mynatt, Doherty, & Tweeney, 1977, Klayman & Ha, 1987) which states that people tend to seek confirming evidence of their initial impressions and ignore disconfirming evidence. So, if a user likes the appearance of a website when they first see it, they may continue to like it regardless of how successful they are in using the site.

In the Lindgaard & Dudek (2002) study, participants completed tasks using a website with high aesthetic appeal and low usability. Participants did not change their ratings of aesthetic value after using the site with low usability; however, they did report lower satisfaction with the site. The authors suggested that this indicates that satisfaction and perception of aesthetic appeal of a website may be independent of one another. However, more research is needed as this study did not examine websites that were low in aesthetic appeal. Therefore, it is not known what would happen to user satisfaction or

ratings of appeal when working with a low appeal website with high usability. Likewise, it is unclear how much influence the tasks users do with the site have on perceptions of satisfaction and appeal. Does a site's usability influence users more if they are searching for specific information than if they are simply browsing?

The purpose of this study was to further examine the relationship between website usability, aesthetic appeal, and user satisfaction both before and after completing tasks that are directed (Search) vs non-directed (Browse). It was expected that the aesthetic appeal and usability of the website would impact user perceptions, performance, and satisfaction of the sites viewed. Specifically, it was expected that the participants would:

- Have more positive first impressions of the high appeal site than the low appeal site.
- Have more positive final impressions of sites with high usability than with low usability.
- Be influenced more by the site's usability when searching a site than when simply exploring the site.

METHODOLOGY

Two websites were selected by the authors from a prior study evaluating user perceptions of websites varying in aesthetic appeal. One site (Figure 1) was rated very high in terms of visual appeal and the other (Figure 2) was rated very low.

Figure 1. High appeal www.australia.com.

Figure 2. Low appeal www.renttoownrealestate.com.

A usability checklist was developed by selecting 36 guidelines from the Research Based Web Design and Usability Guidelines (Koyani, Bailey, & Nall, 2006). Initially, each site was scored by members of the SURL lab based on the compliance with each of the 36 guidelines. The sites then were manipulated to create two unique versions representing high usability and low usability. Finally, the sites were re-evaluated using the usability checklist to confirm that the manipulations created a high and low usability site. The site manipulations are described in Table 1.

Table 1. Usability Manipulations for Low and High Usability sites

Website	Manipulations for High Usability	Manipulations for Low Usability
High Appeal	Added alt tags to all clickable images.	Incorrect links on navigation menu. Distorted path to item. Broken links after homepage with next button as only navigation option. Small hotspots over links. Link location inconsistent throughout pages.
Low Appeal	Added menu navigation in consistent location at the top of the screen. Added page titles to inform users on location within site. Labeled embedded links clearly.	No navigation menu. Distorted path to item. Next button only navigation option. Small hotspot over links. Link location inconsistent throughout pages.

Two task types (Exploratory and Directed Search) were used to examine the impact of the task on user satisfaction. Participants were randomly assigned to one of eight conditions varying in task type, aesthetic appeal and usability. Participants completing the exploratory task were instructed to spend 15 minutes familiarizing themselves with the information presented on the website. Participants completing the search task were given 15 minutes to complete four directed search tasks (these tasks were comparable for both sites). Users were queried for their first impression of each website after a brief exposure, and then for a final impression of the site after completing the task(s).

Participants

One hundred-sixty participants participated in the study (66% female and 34% male). The age of the participants ranged from 18-75 with 82% of the participants between the ages of 18 and 29. The majority of participants reported online use at least several times a month with 79.4% reporting online use daily. Only eight participants reported not owning a computer, and all but one participant reported using a computer a few times a week.

Materials

A background survey was used to gather demographic information and content preferences from participants. Four sites (high appeal high usability, high appeal low usability, low appeal high usability, low appeal low usability) were used. Participant satisfaction was measured using the System Usability Scale (SUS) (Brooke, 1986). Time on task and mouse click data were gathered using the ErgoBrowser™ logging tool.

Procedure

After completing the background questionnaire, participants were informed they had ten seconds to examine the homepage of their first website. Participants were allowed to scroll vertically and horizontally but were not allowed to click on any of the links. At the end of the ten seconds, participants completed a first impression evaluation consisting of an adjective survey adapted from Lindgaard & Dudeck, 2006 (Table 2), a 10-item interest survey and the SUS. The interest survey included statements regarding the participant's interest in continuing to explore the website. Scores were averaged across all nine adjectives to generate a combined average adjective score and across all ten interest statements to create a combined average interest score. After finishing the first impression survey, users then completed a series of randomized search tasks or one exploratory task. The time allowed to complete the task(s) was 15 minutes, after which, participants then evaluated their experience with a final impression evaluation (the same surveys as the first impression survey).

Table 2. Adjective Ratings Scale

Simple	1	2	3	4	5	6	7	8	9	10	Complex
Interesting	1	2	3	4	5	6	7	8	9	10	Boring
Well Designed	1	2	3	4	5	6	7	8	9	10	Poorly Designed
Good Use of Color	1	2	3	4	5	6	7	8	9	10	Bad Use of Color
Good Layout	1	2	3	4	5	6	7	8	9	10	Bad Layout
Imaginative	1	2	3	4	5	6	7	8	9	10	Unimaginative
Very Attractive	1	2	3	4	5	6	7	8	9	10	Very Unattractive
Easy to Use	1	2	3	4	5	6	7	8	9	10	Hard to Use
Easy to Navigate	1	2	3	4	5	6	7	8	9	10	Hard to Navigate

RESULTS

First Impressions of Sites by Appeal

Adjectives. An independent samples t-test was conducted to evaluate the effects of aesthetic appeal on user first impression adjective scores. Participants scored the high appeal sites as more appealing than those viewing the low appeal sites $t(158) = 12.59$, $p=.001$, $d=1.93$, as shown in Figure 3.

First Impression Adjective Score

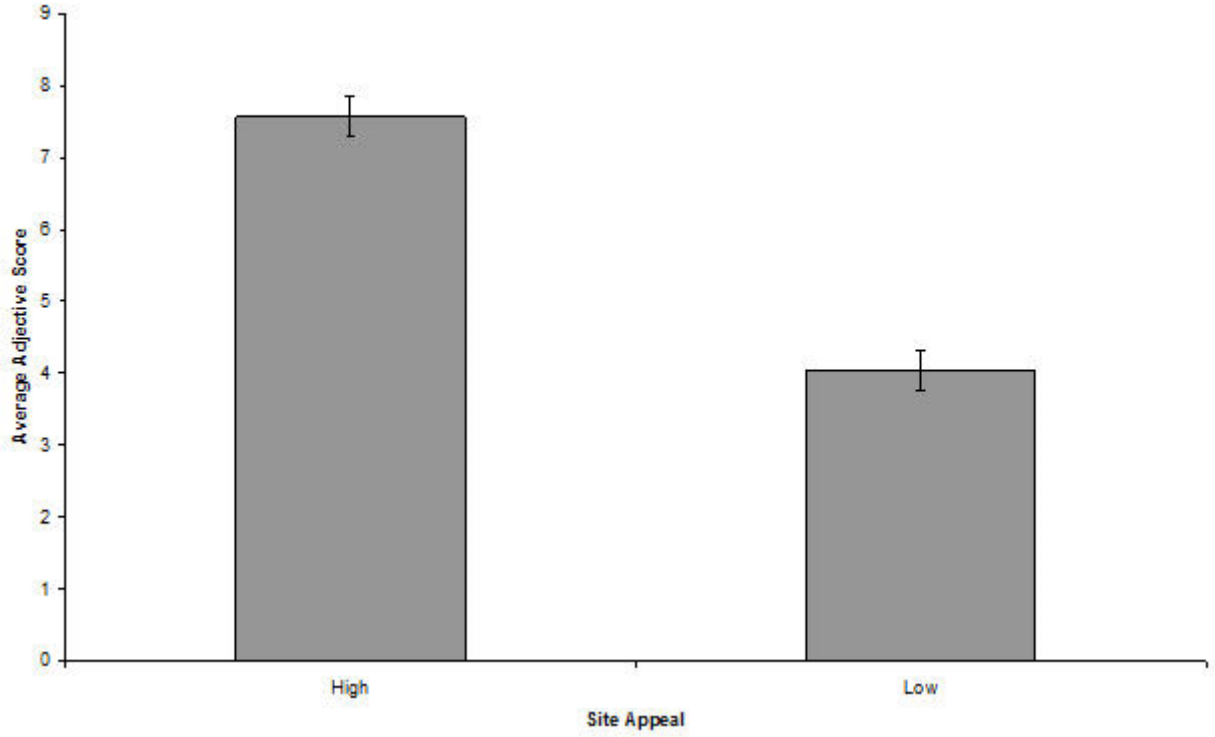

Figure 3. First Impression Adjective Scores for Appeal Mean (SD) (Scale 1-10 with a score closer to 1 = very positive and 10 = very negative) (N = 80 for each condition).

Interest. An independent samples t-test was conducted to evaluate the effects of aesthetic appeal on first impressions interest scores. Participant's average score across all variables was higher for those in the high appeal conditions than those in the low appeal condition $t(158) = 9.76$, $p<.001$, $d= 1.72$, presented in Figure 4.

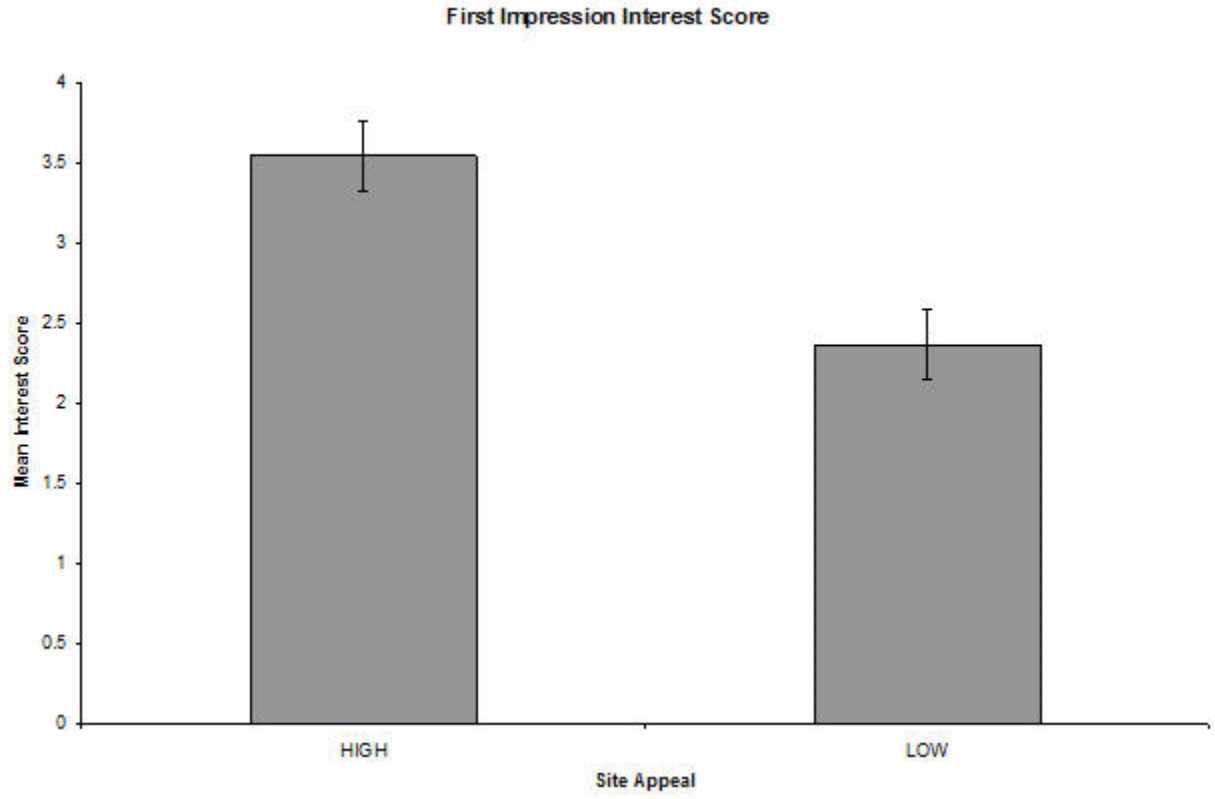

Figure 4. First Impression Interest Score for Aesthetic Appeal Mean (SD) (Scale 1-5 with 1= very uninterested and 5= very interested) (N = 80 for each condition).

Satisfaction. Participant satisfaction was measured using the System Usability Scale (SUS) which provides a score between 0 and 100. An independent samples t-test was conducted to evaluate the relationship between aesthetic appeal and participant first impression satisfaction scores. Participants scored their experience as more satisfying on the high appeal sites than the low appeal sites $t(158) = 9.87$, $p<.001$, $d= 1.62$ presented in Figure 5.

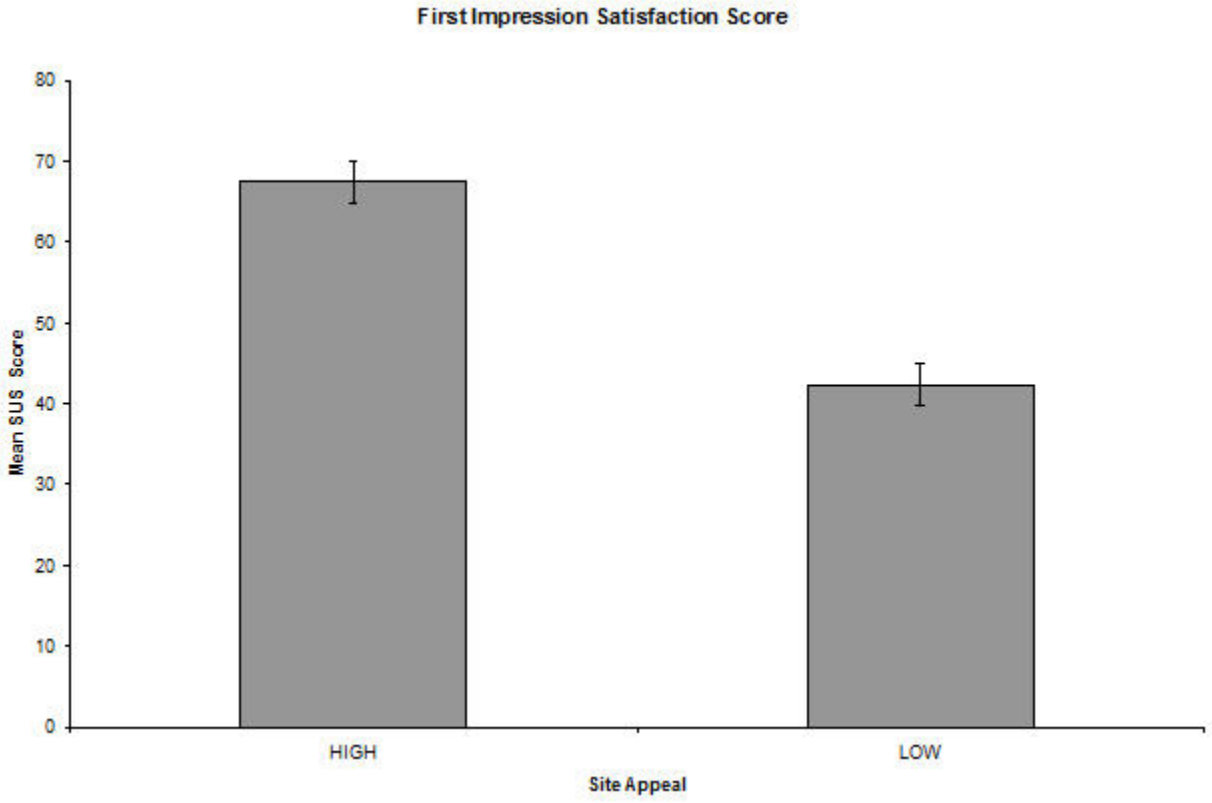

Figure 5. First impression satisfaction scores for high and low appeal sites.

Impressions after using the Website

Adjectives. A 2 X 2 X 2 between-subjects ANOVA (appeal x usability x task type) was conducted to evaluate the effects of factors on participant final impression adjective score. The means and standard deviations for final impression adjective ratings are presented in Table 3. There was a significant main effect of task type $F(1,152) = 7.15$, $p<.001$. $\eta^2=.044$ with search task groups ($M = 4.57$) scoring the sites less positively than those in the exploratory groups ($M = 5.40$). A significant main effect of site appeal was also found $F(1,152) = 111.50$, $p=.001$, $\eta^2=.423$ with participants in the high appeal site group ($M = 6.58$) scoring the sites more positively than those in the low appeal group ($M = 4.78$). There was no significant main effect of usability or any interactions.

Table 3. Final Impression Adjective Scores for Usability and Task Type Mean (SD) (Scale 1-10 with a score closer to 10 = very positive and 1=very negative) (N=20 for each condition).

	Search Task		Exploratory Task	
Usability	**High Appeal**	**Low Appeal**	**High Appeal**	**Low Appeal**
High	6.39 (1.80)	2.52 (2.14)	7.22 (2.06)	3.65 (1.95)

Low	5.75 (1.62)	2.68 (1.94)	6.97 (1.83)	3.73 (1.88)

Interest. A 2 X 2 X 2 between-subjects ANOVA (appeal x usability x task type) was conducted to evaluate the effects on user final impression interest scores. The means and standard deviations for interest scores as a function of appeal, usability and task type are presented in Table 4. There was a significant main effect for site appeal $F(1,152) = 170.84$, $p < .001$. $\eta^2 = .529$ with participants in the high appeal group (M = 3.64) having a higher interest score than those in the low appeal group (M = 2.14) (Table 5). There was a significant main effect for task type $F(1,152) = 4.02$, $p = .05$. $\eta^2 = .026$ with participants in the search task group (M =2.77) scoring the site as less interesting than those in the exploratory task group (M = 3.00). There was no main effect for usability or any significant interactions.

Table 4. Interest Final Impression Scores for Appeal, Usability and Task Type Mean (SD) (Scale 1-5 with 1=very uninterested and 5= very interested) (N = 20 for each condition).

	Search Task		**Exploratory Task**	
Usability	**High Appeal**	**Low Appeal**	**High Appeal**	**Low Appeal**
High	3.68 (.68)	2.45 (.55)	2.87 (.90)	2.13 (.55)
Low	3.23 (.74)	1.95 (.60)	3.78 (.86)	2.23 (.62)

Table 5. Interest Final Impression Scores by Individual Variable for the Main Effect of Appeal Interest (Scale 1-5 with 1=very uninterested and 5= very interested).

	High Appeal Site	**Low Appeal Site**
I want to spend more time on this site **	2.75 (1.16)	1.54 (.94)
I want to move beyond the homepage on this website **	3.23 (1.21)	2.10 (1.34)
I think the design of this site positively influences my final impression **	3.41 (1.04)	2.49 (1.49)
I understand the purpose of this site **	4.16 (.91)	3.40 (1.30)
I would enjoy spending time on this site **	2.81 (1.15)	1.66 (1.03)
I like this site **	2.96 (1.08)	1.76 (1.16)
The homepage is an accurate representation of the site **	3.49 (.96)	2.90 (1.37)
The homepage is attractive **	3.96 (.88)	2.33 (1.23)
The information on this site is accurate **	3.72 (.88)	2.73 (1.16)
The design of this website looks trustworthy **	3.98 (.89)	2.45 (1.32)
First Impression Interest Mean Score **	3.64 (.68)	2.14 (.81)

** $p < .01$

Satisfaction. A 2 X 2 X 2 between-subjects ANOVA (appeal x usability x task type) was conducted to evaluate the effects on user final impression satisfaction scores. The means and standard deviations for satisfaction scores as a function of the three factors are presented in Table 7. There was a significant main effect for appeal $F(1,152) = 58.98$, $p < .001$. $\eta^2 = .280$ with participants in the high appeal group (M

=59.31) scoring their experience as more satisfying than those in the low appeal sites (M=35.53). There was a significant main effect for task type $F(1,152) = 10.89$, $p<.001$. $\eta^2=.067$ with exploratory groups (M=52.53) scoring the site as more satisfying than search groups (M=42.31). There was not a significant main effect for usability or any significant interactions.

Table 7. Satisfaction Final Impression Scores for Appeal, Usability and Task Type Mean (SD) (N = 20 for each condition).

	Search Task		Exploratory Task	
Usability	**High Appeal**	**Low Appeal**	**High Appeal**	**Low Appeal**
High	56.13(21.60)	36.25 (20.72)	68.63 (16.77)	36.88 (17.32)
Low	47.13 (20.17)	29.75 (18.46)	65.38 (19.40)	39.25 (21.61)

DISCUSSION

Results from this study indicated that, in terms of first impressions, users rated the sites with high appeal as more interesting, easier to use, easier to navigate, more accurate, more trustworthy, and overall more satisfying than the low appeal site. This was before they had even used the site. In fact, they had *only* viewed the site homepage for 10 seconds.

After using the websites, the impressions of the users did not change significantly even though the site they were using was manipulated to have poor usability. Users still gave more positive feedback, reported more interest, and higher satisfaction for the high appeal site. Users who merely explored the sites (both high and low appeal; low and high usability) also were more positive than those who searched for specific information. Participants appear to be more critical of sites when they are attempting to find specific information in a limited amount of time than participants that were allowed to browse the site with no specific goal. These results support the confirmation bias theory proposed by Lingaard and Dudek (2002) in that the users' first impressions held true even after using a site that was not usable. In their study, however, they found different effects of usability on satisfaction than aesthetic appeal. In our study, all measures (satisfaction, aesthetic appeal, and interest) were not significantly affected by usability. That said, the participants in our study were using a prescribed site and prearranged tasks. It is not known whether the effects of usability would be more noticeable under circumstances where the users may be more personally engaged.

Future research should focus on what exactly makes a homepage aesthetically appealing. One of the main differences between the high appeal and low appeal site used in this study was the proportion of text vs. images found on the homepage. The high appeal site homepage consisted of a large landscape with minimal text. The low appeal site homepage had no images and contained a lot of text. It is not known if the same results would have been obtained if the sites had an equal amount of text and graphics, but still maintained the low and high appeal.

CONCLUSION

An attractive homepage entices users to view more of the site and creates feelings of interest and initial satisfaction. If the homepage is unattractive, users do not appear to be interested, nor do they desire more interaction with the site. Designers must develop a homepage that not only attracts user' attention but also engages them. This research suggests that an attractive site is more likely to pull in users than an unattractive site regardless of how well it is designed from a usability standpoint. An unattractive site, despite high usability, does not attract user interest and maintains low satisfaction.

*Note: This article presents a portion of findings from a comprehensive study investigating the influence of appeal, usability, and task type on website satisfaction by Dr. Christine Phillips. Please contact Dr. Phillips for more information.

REFERENCES

Brady, L., & Phillips, C. (2003). Aesthetics and Usability: A Look at Color and Balance. Usability News, 5.1, retrieved September 14, 2009 http://www.surl.org/usabilitynews/51/aesthetics.asp

Brooke, J. (1996). SUS: A Quick and Dirty Usability Scale, in P. Jordan, B. Thomas, B. Weerdmeester, & I. L. McClelland (eds.), Usability evaluation in industry, (pp. 189-94). London, UK: Taylor & Francis.

Koyani, S, Bailey, R., & Nall. (2006). Research-Based Web Design and Usability Guidelines. www.usability.gov.

Klayman, J. & Ha, Y. (1987). Conformation, Disconfirmation, and Information in Hypothesis Testing. Psychological Review, 94(2), 211-228.

Kurosu, M., & Kashimura, K. (1995). Apparent usability vs. inherent usability experimental analysis on the determinants of apparent usability. Paper presented at the annual meeting of CHI. ACM Press.

Lindgaard, G. & Dudek, C. (2002). High appeal versus high usability: Implications for user satisfaction, HF2002 Human Factors Conference, Melbourne, Australia, November 25-27.
Lindgaard, G. Fernandes, G., Dudek, C. & Brown, J. (2006). Attention web designers: you have 50 milliseconds to make a good first impression!. Behaviour and Information Technology, 25, 115-126.

Mynatt, C.R., Doherty, M.E., & Tweney, R.D. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inference. Quarterly Journal of Experimental Psychology, 29, 85-95.

Tractinsky, N., Katz, A.S., Ikar, D. (2000). What is beautiful is usable. Interacting with Computers¸ 13, 127-145.
