The influence of expectations on usability and visual appeal in a web environment

Milica STOJMENOVIC[a,1], Robert BIDDLE[a], John GRUNDY [b], and Vivienne
FARRELL[c]

[a]*Carleton University*
[b]*Monash University*
[c] *Swinburne University of Technology*

Abstract. The impact of verbal and textual influences that create expectations on
the perceived and objective usability and visual appeal of a website was examined.
A computer laboratory study, in the form of a user-based usability test, was done to
determine the effect of textual and verbal expectations on visual appeal and usability.
Results showed that the combination of textual and verbal expectations successfully
influenced participants. When told that the website was going to be hard and ugly,
participants were disinclined from using it, stating it was too hard to use, and
struggled more when using it. Similarly, participants thought that the website was
easier to use and prettier in the high expectations group than in the low expectations
group. Results suggest that web developers and project managers should focus on
investing in marketing and social media influencers as well as in the development
of an attractive and usable website.

1. Introduction

Many studies have been done on the relationship between usability and visual appeal but
the results vary vastly. There are many factors that influence the results, including
website domain, the type of task, if incentive is given, and metrics used to get the
usability and visual appeal measures. However, the impact of expectations on these two
variables is understudied. Therefore, this paper examines the influence of expectations
on usability and visual appeal, on an unfamiliar city council website [1] to control for
prior experience and expectations. This work is an extension of previous work [1-2].

According to ISO 9241/11 [3], "usability is the extent to which a given product can
be used by a specific group of users, to achieve specific goals with effectiveness,
efficiency, and satisfaction in a specific context of use". Effectiveness is how well users
achieve specific goals, efficiency is the time to complete a given task, and satisfaction is
the user's experience of acceptability. The context of use is a predefined group of users,
in a specific environment, who perform given tasks with the interface. Visual appeal is
the cognitive judgment of an object's aesthetic appearance [4]. Expectations are beliefs
on what is about to happen.

---

[1] Corresponding Author

2. A Brief Background

In Human-Computer Interaction (HCI), "what is beautiful is usable" [5], since a relationship between a user's first impression of visual appeal and perceived usability was found, pre- and post- system use. However, more recent literature reveals different findings on the relationship between usability and visual appeal. For example, Katz [6] found significant correlations between perceived visual appeal and perceived usability *before* system use, but not after. Tuch et al. [7] found that usability affected perceived visual appeal after use, and not the other way around. Thus, the relationship between usability and visual appeal has not yet been properly understood.

Textual expectations, in the form of textual word-of-mouth (WOM), affect visual appeal and usability in an unfamiliar website type (i.e. city councils) [2]. However, the effect of expectations on visual appeal and usability was smaller than anticipated since subjective usability was not affected, and visual appeal was not affected on a attractive and easy to use website. Therefore, the purpose of this study was to reinforce expectations by implementing them verbally as well as in text to see if it would have a greater effect.

The textual implantation of expectations used in an earlier study [2] mimicked textual online reviews. The verbal implementation of expectations used in this paper mimics in-person and video reviews, such as face-to-face communication between friends and online reviews such as those found in YouTube videos, respectively. To examine the impact of social media on the use and perception of websites, we examined both textual and verbal implementations of expectations in this paper.

3. Method

Five studies [1] were completed in order to get an easy and attractive website. Three levels of expectations were implemented: (a) easy and attractive (i.e. high expectations), (b) hard and ugly (i.e. low expectations), and (c) the control (i.e. no expectations). Expectations were implemented textually and/or verbally. As with an earlier study [2], textual expectations were nuanced task descriptions that participants read before they interacted with the website. The verbal expectations came from a confederate who acted like a participant and spoke to participants before the study began, giving them their 'opinion' of the website they supposedly just interacted with. This opinion was a rehearsed speech, similar to the task description (i.e. written expectation), reinforcing the expectations. The previous study [2] used only textual expectations. The present study used the same method from the earlier study [2] but added verbal expectations in an effort to strengthen the implementation, to see if more significant results could be obtained.

3.1. Participants

In the present study, 20 (16 males, 4 females; 16 aged 18-30 years, 4 aged 31+) participants were recruited. All volunteered and were screened for 20/20 or corrected to 20/20 vision and colour blindness. Thirteen participants were not familiar with the purpose of city councils (the subject of our test websites), and seven were somewhat familiar. No one was very familiar with the purpose of a city council. Participants were individually tested, approximately one hour per session, ten participants per condition.

In the analysis below, there are 30 participants which is the result of the addition of the control condition data from [2].

3.2. Materials

The System Usability Scale (SUS; [8]) and the Visual Appeal of Websites Inventory – Short version (VisAWI-S; [9]) scales were used for perceived measures of usability and visual appeal, pre- and post-use. Objective usability was acquired per task: the number of clicks, the number of hovers, task completion time in seconds with a maximum of three minutes, and success (pass if the answer is correct and within 3mins). Pre-use measures were based on a 6-second exposure to the website's homepage, a page one click in, and a page two clicks in to the website (2sec/pg). Ten randomly ordered information retrieval tasks were given to participants. An example of a task is: "How many beaches are located in the Gold Coast?"

3.3. Procedure and Data Analysis

A confederate was in the experiment room, picking up their things and getting ready to leave as the participant entered the room. The experimenter asked the confederate if they were all done and the confederate responded that they were just leaving. The experimenter thanked them and told the participant to go ahead in and wait a minute while the experimenter left to set up the computers. The confederate then told them the usability and visual appeal expectations in the form of their experience with the website and left. The experimenter came back into the room.

The consent form was signed first. Then, the participant was given the written task descriptions according to the condition that they were randomly assigned to. The instructions were on the computer screen. Participants then briefly viewed the website slideshow and evaluated pre-use usability and visual appeal. Then, they were instructed to start each of the information-retrieval tasks from the homepage, told that the search bar would not work, to avoid using other websites or prior knowledge to answer the tasks, and asked to persist with a task until they got an answer or were told to move to the next one. The researcher then left the participant in room alone. As soon as the researcher and participant were both ready in their separate rooms, the usability test began. The participant and researcher were connected via a hands-free speaker. Participants then filled out the visual appeal and usability scales again. The researcher returned to the participant's room and asked the participant for feedback on usability, visual appeal, and if they believed and agreed with the confederate.

4. Results

4.1. Participant Feedback

None of the participants in the low expectations condition said anything positive about the usability, whereas 4/10 participants in the high expectations condition said that it was easy to use, consistent, and had a well-structured layout overall. Everyone in the low expectations condition was hostile about the website. For example, one said that he

would fire the developer, while another said during the test, "can I just put 'no' without looking for it in the website? It's so bad."

Seven participants from the high expectations group said that the website looked great, had great colours, and that it looked "easy on the eyes". Only four participants from the low expectations condition said that it looked good but they were less confident in their opinions. For example, one said that the website "looks modern, I guess."

When asked if they believed the confederate, three people from the high expectations condition said that they did and four in the low expectations group said that they did. One participant in the high expectations group mentioned that he did have high expectations of this website when usually he did not have any of city council websites. One participant in the low expectations group said that he "tried not to be biased but subconsciously, I was." Therefore, it might be the case that people heard the confederate and tried to be neutral, but were indeed influenced.

One participant in the high expectations group agreed with the confederate saying that it was a great website, while three agreed with the confederate's description of the 'bad' website in the low expectations condition. Lastly, two participants in the high and three in the low expectations condition said that they did not remember what the confederate. They never interacted with a city council website and had no expectations.

There was less disagreement with the confederate in this study than in disagreement with the textual expectations in [2], suggesting that the addition of the confederate influenced subjectively participants more so than just the written task description.

4.2. Statistical Testing

Assumptions Testing. The assumptions for normality and homogeneity of variance were checked for each variable across all conditions and were not unilaterally met. Shapiro-Wilk tests showed that clicks ($p<.05$), post-use visual appeal ($p<.05$), and post-use usability ($p<.05$), each with the low expectations condition, were not normally distributed. Post-use perceived usability for the low expectations group had a skewness of -1.627 (SE=0.687) and a kurtosis of 2.091 (SE=1.334), suggesting that it was not normally distributed as well. The non-parametric Levene's test revealed that the homogeneity of variance assumption was not violated. Given that assumptions for normality were not met, that a variable was binary (success), some were discrete (clicks and hovers) and others continuous (time), and that sample size per condition was small (n=10), Kruskal-Wallis and Fisher's Exact tests were applied where appropriate.

Main effects were found in pre-use usability ($p<0.01$), post-use visual appeal ($p=.01$), and post-use usability ($p<.01$). This means that a statistical difference was found amongst the high, low, and no expectation conditions. Paired comparisons showed that high and low expectation conditions differed in pre-use usability ($p<.01$), post-use visual appeal ($p<0.05$), and post-use usability ($p<.01$).

The average number of clicks per task ($p<.05$) differed within the attractive website conditions. Pairwise comparisons showed that the number of clicks were different ($p<.01$) between the high and low expectations conditions. Specifically, participants in the low condition, on average, clicked more often per task (3.82 clicks) than those in the high condition (2.8 clicks). Main effects were also found for task completion time ($p<0.01$) and the average number of passed tasks ($p<.05$). Pairwise comparisons found that the difference in time ($p<.01$) and passes ($p<.05$) was between the high and low conditions. Participants took over half a minute longer to complete a task in the low condition (i.e. 108sec in low versus 71sec in high). The significance of the comparison

of the average of passed tasks was confirmed (p<.01, one-tailed) with a Fisher's Exact test, where high had a better success rate (0.83) than the low group (0.58).

Summary of statistical results. High and low expectation groups differed in pre-use usability, post-use visual appeal, and post-use usability. In addition, the average number of clicks per task, average task completion time, and proportion of passes (success rate) differed between these two groups. Therefore, participants rated the same website as prettier and easier to use when they were told that it was going to be well made, attractive, and usable. Moreover, they struggled more with the website when completing the information retrieval tasks when told that the website was hard to use. *Expectations influenced both how participants viewed and interacted with the website.*

5. Discussion

5.1. Results Summary

While the website presented to everyone was the same, the majority of participants in the high expectations condition thought that the website was attractive, whereas the majority of participants in the low expectations condition did not. People in the low expectations group were very critical of the usability and visual appeal levels. Yet, 8/10 in high expectation group said that they liked them.

Statistically, the low and high conditions differed in pre-use usability, post-use visual appeal, and post-use usability. This means that the same website was differently rated, depending on what the expectation was them before the experiment. Specifically, participants rated the website better when they were told it was going to be easy and attractive, and they rated it as worse when they were told the opposite. In addition, the average number of clicks, completion time, and the success rates differed between the high and low conditions. The low expectation group made more clicks, took longer, and had a lower success rate while doing the same tasks and using the same website.

5.2. Implications for Website Design

A bad reputation can turn people against your website, even if the reputation is not true. To overcome this, website operators might invest in marketing and social media influencers to give a website a more positive reputation right from the beginning. It will influence people before they use it and, according to the results of this study, last throughout use to influence their opinions after having used the website. In this study, participants were forced to use it, whereas in real life there are thousands of websites to choose from and competition can be fierce. If you advertise, people will (1) know about it, (2) know something good about it, (3) be willing to check your website out, and (4) like it a bit more than they would otherwise.

5.3. Limitations and Future Research

One limitation is the small sample size. However, we have significant results even with a small sample size which would only suggest that a larger sample would be even more successful in illustrating the influence of expectations.

This study was done using a confederate who acted like a participant just finishing the usability test, and either praised or complained about the website. However, this may not be the best way to do so given the unfamiliarity, untrustworthiness, and minimal exposure to the confederate and the expectation. Yet, again, the results showed that expectations did influence usability and visual appeal, and indeed, more so than in the earlier study [2].

Given that expectations influence visual appeal and usability when both are either high or low, the next study should examine what happens with this relationship when the usability and visual appeal are incongruent with each other.

6. Conclusion

Visual appeal and perceived usability were rated as higher when the expectation was high, and lower when the expectation was low. Participant performance was also affected by expectations. Comparing to the study in the previous study [2], the results suggest that verbally enforced expectations do impact the perception and use of a website, more so than just written task descriptions on their own. This suggests that while textual user reviews are effective, the combination of textual and verbal expectation setting is even more effective.

Thus, marketing and social media influencers have the power to say people towards and away from your website. It is important to give a website a more positive reputation right from the beginning. **If you advertise, people will (1) know about it, (2) know something good about it, (3) be willing to check your website out, and (4) like it a bit more than they would otherwise.**

7. References

[1] M. Stojmenovic, C. Pilgrim, & G. Lindgaard, Perceived and Objective Usability and Visual Appeal in a Website Domain with a Less Developed Mental Model. OZCHI'14, Dec2-5, Sydney, Australia (2014), 316-323.

[2] M. Stojmenovic, J. Grundy, V. Farrell, R. Biddle & L. Hoon. Does Textual Word-of-Mouth Affect Look and Feel? OZCHI'16, Nov29–Dec2, Launceston, Australia (2016), 257-265.

[3] ISO 9241/11. 1996. International Organization for Standardization. Retrieved June 2012.

[4] J. Blijlevens, Typically the Best? Perceived Typicality and Aesthetic Appraisal of Product Appearances. PhD thesis, Delft University of Technology, The Netherlands, 2011.

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

[6] A. Katz, Aesthetics, usefulness and performance in user-search-engine interaction. *Journal of Applied Quantitative Methods, 5*(3), (2010), 424-445.

[7] A. N. Tuch, S.P. Roth, K. Hornbaek, K. Opwis, & J. A. Bargas-Avila. Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. *Computers in Human Behavior, 28*(5), (2012), 1596-1607.

[8] J. Brooke. System Usability Scale (SUS). Digital Equipment Corporation, UK,1986.

[9] M. Moshagen, & M. Thielsch. A short version of the visual aesthetics of websites inventory. Behaviour and Information Technology (2012), 1-7.