The Influence of Website Category on Aesthetic Preferences

Eleftherios Papachristos, Nikolaos Avouris

HAL Id: hal-01497453

https://hal.inria.fr/hal-01497453

Submitted on 28 Mar 2017

The Influence of Website Category on Aesthetic Preferences

Eleftherios Papachristos and Nikolaos Avouris

University of Patras, Human-Computer Interaction Group,
Electrical and Computer Engineering Dept.,
GR-265 04 Rio Patras, Greece.
epap@ece.upatras.gr, avouris@upatras.gr

Abstract. This paper investigates whether users' aesthetic impressions about websites vary considerably across different domains. The assumption that aesthetic judgments about websites that belong to diverse domains are based on different visual design aspects has been investigated in three distinct studies in healthcare, tourism, and web design business. In these studies participants expressed their overall preference as well as their judgments on the constructs of visual appeal, perceived usability and novelty. In addition, descriptions about the test websites were obtained by expert panel and objective measures. Preference Mapping (PM), which is a data summarization and visualization technique, has been performed in each study. Attribute projection into the preference maps allowed for the identification of important driver of preference for each individual domain. Even though, visual appeal was the most important predictor of overall preference in all studies, appealing websites had different visual characteristics in each domain. Furthermore the importance of the evaluation constructs varied considerably among studies, indicating that aesthetic perceptions differed considerably across domains. These findings emphasize the need for flexible evaluation methods that can be used to identify important visual design factors within a specific website domain.

Keywords: Website design, aesthetic evaluation, website categories, visual appeal, preference mapping

1 Introduction

Since Tractinsky et al. [1] in 2000 published their seminal paper "What is beautiful is usable" much has changed in the HCI community in regard to aesthetics research. The lively debate that was initiated with the controversial suggestion that aesthetic design could influence perceptions about pragmatic qualities of user interfaces shifted attention to more subjective aspects of interaction. Particularly for websites the construct of *visual appeal* (used by some authors interchangeably with beauty) has been proven to be a very important factor determining users' overall impressions [2]. Furthermore, a series of experiments (e.g. [3]) have sown that users could form stable visual appeal judgments in time periods of as short as 50 msec. These judgments are based mainly on visual design since other aspects (e.g. content) cannot be recognized

during such sort time periods. Surprisingly, relatively stable judgments in the same exposure times could also be found for the constructs of perceived usability and credibility [3]. These results demonstrate the general importance of websites visual design. Although, actual website use could influence users' perceptions, it is nonetheless important to create positive first impressions, considering that it is more difficult to overcome negative ones.

These research findings emphasize the importance of website's visual design and therefore the need for appropriate evaluation methods. The most common evaluation approach is using one of the aesthetics oriented multiple-item instruments that are gaining steadily acceptance in the HCI community; for example the "Classical-Expressive" aesthetic scale [3], AttrakDiff [2] or the more recent visAWI [6] questionnaire. Among the advantages, of those multiple-item questionnaires is that they ensure fairly reliable and valid measurements and that they provide common ground for results communication and for between study comparisons. These questionnaires have been created to be sufficiently generic in order to be applicable to most - if not all - websites. This generality has been achieved by the inclusion of a variety of websites from different domains as test stimuli during questionnaire creation. Both the VisAWI and the "Classical–Expressive" aesthetic scale, for example, reported a similar website sampling procedure during questionnaire creation and validation. The result of using websites of different domains as test stimuli is that design factors pertinent to specific domains are canceled out. Visual design evaluation with one of the aforementioned questionnaires means having participants rate a design on a predefined set of factors that have been identified to be important for websites in general.

However, websites vary in terms of purpose, target user groups, and therefore visual styles. Although, there is no commonly agreed upon taxonomy, various categorizations schemes of websites have been proposed. Studies have shown that users have distinct mental models for different kinds of websites [7]. According to Norman [8] users form internal mental models of things with which they have interacted, which in turn creates expectation about similar objects they may encounter. These expectations are becoming stronger as the number of encounters increases and users are gaining more experience with a particular website domain. For example, users may have different expectations about the visual design of news websites or online shops. Tuch, et al. [9] showed that users' aesthetic judgment can be strongly influenced by their perceptions of website *prototypicality*. Prototypicality refers to the amount to which an object is representative of a class of objects and depends heavily on each individuals mental models that are build through experience. Designs that contradict what users typically expect of a website may lead to a negative first impression [9]. The fact that prototypicality is an important influencing factor in aesthetic judgments has been shown repeatedly in various empirical studies [9]. Thus, relying on predefined questionnaires could mean ignoring the visual design aspects that in the users mind have particular importance in the specific website domain.

Other evaluation methods such us Repertory grid technique (RGT), Multidimensional Scaling (MDS), or Preference Mapping (PM) are based on multiple website evaluations and do not impose a set of predefined evaluation criteria on participants. In addition, conducting evaluation studies with one of these techniques allows for the

identification of design factors that are important in specific website domain by using multiple websites from the same domain as study stimuli. In the studies presented in this paper Preference Mapping (PM) has been used for evaluation of visual design of a number of examples of websites of a specific domain. This was repeated in the domains of healthcare, tourism and web design businesses. The main objective of this paper is to demonstrate that different design factors can be important driver of preferences in websites that belong to different domains, based on comparison of the findings of evaluation of the websites of these domains. This research is part of a broader research project that attempts to define guidelines for design and evaluation of various kinds of websites including social media presence of small medium organizations in various fields.

2 Method

Three evaluation studies have been conducted involving website designs from three distinct domains. In the first study (Healthcare domain) 15 hospital websites were evaluated by 34 participants (29 male, 4 female, mean age = 22.2). In the second study (Tourism domain) 32 participants (21 male, 11 female, mean age = 23.3) were asked to evaluate 18 hotel websites. In the third study (Web design business domain) 12 websites of web design companies were evaluated by 30 participants (17 male, 13 female, mean age = 28). These particular websites types were selected because they represent domains that differ both in terms of characteristics and in purpose. However, given that the goal was to identify important visual design characteristics that can shape user first impressions in each of these categories it was important to make sure that none of the test websites were previously known to our participants. Therefore, test websites were randomly selected from lists of top U.S. hospitals in the first study and New Zealand hotels in the second. Since all of our participants were of European origin it was assumed that these choices would minimize the possibility of prior familiarity with the specific test websites and thus influence the "first impression" effect. None of the participants reported previous experience with any of the test websites. In all cases participants volunteered to take part in the evaluation studies and did not receive any compensation. In all studies screenshots were used instead of actual websites, as our goal was to study the impact of various visual design aspects on participants' first impressions. Although, studies have shown [3] that perceptions about constructs such as visual appeal are relatively stable over time evaluations after actual website use could be biased by non-visual design related aspects (e.g. content).

2.1 Procedure

In each evaluation study participants first viewed screenshots of all the websites in a random order and then rated them according to their overall preference on a linear, unmarked scale (from 0 to 100) with the verbal anchors "least preferred" and "most preferred" at the two ends. In a subsequent evaluation phase participants were asked to rate the websites again on the constructs *visual appeal*, *credibility*, *perceived*

usability and *novelty*. Since participants had to rate multiple websites in each study only a limited number of evaluation constructs had been included in the studies in order to avoid participant fatigue. However, the identification of important design characteristics required a better profiling of the test websites that could be provided by these four constructs alone. For this reason all websites were rated on various descriptive attribute by an expert panel.

Nine experts (visual designers, HCI practitioners, and web developers) identified a list of visual design aspects which could possibly influence user preferences. The goal was to find a comprehensive list of descriptive attributes, such as *symmetry* or *complexity*, which could reflect variations in website designs. After a literature review and panel discussions, a preliminary list of attributes was tested on a set of generic websites in order to eliminate unsuitable attributes. The criteria for elimination were: limited discrimination ability and disagreement between assessors about meaning. In a subsequent session our experts rated the actual test websites on the final 15 attributes. Before finalizing the descriptive dataset, the attributes that did not discriminate significantly between our actual websites were identified through mixed model ANOVA's (websites as fixed and experts as random factors) and were excluded from further analysis.

In addition to participant and expert ratings we also used 16 objective measures that could be grouped into three categories: a) *text related metrics* (e.g. number of words, number of visible links), b) *area related metrics* (e.g. percentage of website used to display images or text), and c) *color related metrics* (e.g. average brightness, saturation). The text related measurements were taken with the help of optical character recognition (OCR) software and were double checked manually. This technique was used instead of html parsing since contemporary websites use graphics or flash instead of plain text very frequently. Websites fragmentation to specific areas (e.g. navigation, images) was done manually with the help of graphic editing software. For the color related measurements, a color recognition program has been written that parsed the website screenshots and calculated the metrics.

3 Analysis

In order to identify which design characteristics were most influential in preference formation, a common approach would be to perform multiple regression with *preference* as the depended variable and the various attribute ratings as the predictors (e.g. [2]). However, since a purely exploratory approach was followed in regard to the selection of predictors the number of independent variables (30-35) in most studies was larger than the number of observations. Using all predictors at once would over-fit the regression model. In addition most of the predictor variables are highly correlated to each other, which can lead to multicollinearity problems. Principal Component (PCR) or Partial Least Square (PLS) regression models in which a large amount of predictors are transformed into view orthogonal uncorrelated components are better suited for these circumstances. A data analysis method based on PCR called Internal Preference Mapping (IPM) has been used in our study in order to identify important design characteristics in each case.
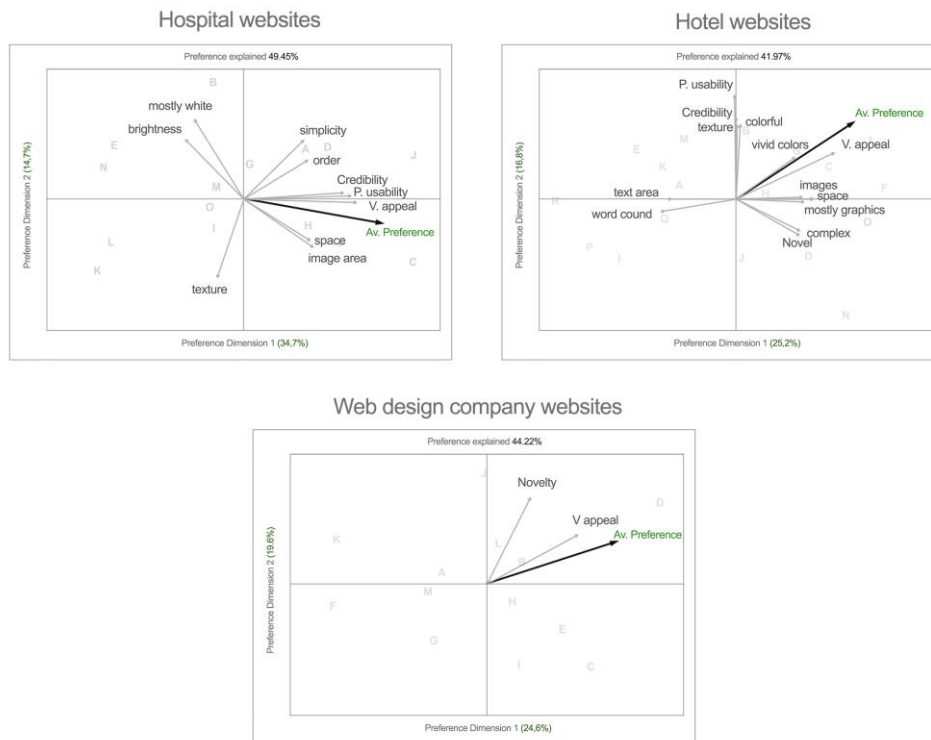
Fig. 1. Preference maps with attribute projection for each of the three studies. Light vectors represent attributes. Attributes that begin with a *capital letter* are participant construct ratings. The bold vector indicates the average preference direction of the participant sample.

Preference *mapping* is referred to as a group of multivariate statistical techniques aimed at gaining deeper understanding of participants' preferences toward stimuli [10]. The method is a data summarization and visualization technique that creates low dimensional maps depicting stimuli and individual participant preferences simultaneously. This is usually accomplished by conducting Principal Component Analysis (PCA) on a data matrix consisting of stimuli in rows and participant preferences in columns. Since in PCA the first components account for the maximum possible amount of variance two or three dimensional spaces are usually sufficient to capture the majority of the underlying preference structure. Based on Euclidean distances conclusion can be drawn regarding website similarities as well as individual participant preferences towards them.Interpretation of the resulting dimensions as well as identification of important drivers of preferences can be accomplished by projecting additional website attributes into the preference map. Additional data about websites can be projected in the preference space by using average attribute scores as dependent and website factor scores as independent variables in a regression model. The regression coefficients represent the strength of the relations between the additional attributes and the preference dimensions. Attributes that have no relationship with any of the preference dimension cannot be used for preference

interpretation and should therefore be removed from further analysis. Thus, during website data projection insignificant attributes that cannot explain participant's preferences can be identified and discarded.

Participant preferences ratings were submitted to a PM analysis for each study individually. Figure 1 shows the three resulting preference maps. In these maps website designs are represented by capital letters while attributes are depicted as light vectors. The vectors indicate the general direction in which the intensity of each attribute increases. Attributes with vectors that point to a similar direction are positively correlated while attributes with vectors pointing to opposing directions are negatively correlated. Websites that lie in the general direction to which an attribute vector points have high intensities of that attribute while objects in the opposite directions have none or low intensities of the same attribute. The bold vector indicates the average preference direction of the participant sample.

From the initial list of 35 attributes only 9, 13 and 2 could be successfully projected in the 2-dimensional preference spaces of the healthcare, tourism, and web design business studies respectively. Although, analysis has been conducted in higher dimensional spaces the results presented here are primarily focused on the first two dimensions which represent the most important components in participant preferences.

In the *healthcare* study the first preference dimension represented 34.7% of the total preference variance and was highly correlated with the constructs of *visual appeal*, *perceived usability* and *credibility*. The second dimension which explained 14.7% of preference variance correlated with descriptive attributes such as *simplicity*, *order*, *brightness* on one site and *large image area*, *white space* and *texture* on the other.

In the *tourism* study the first dimension that captured 25.2% of preference variance differentiated among websites on the left site of the map which were *mostly graphics* based, had more *white space*, and *larger image areas* while designs on the right site were mostly *text based*. The later was confirmed by trained panel data (*mostly graphics* attribute) and by objective measures (*text area*, *word count*). The second dimension was primarily correlated with *perceived usability* and *credibility* as well as to the descriptive attributes *texture*, *colorful* and *dimensional*. *Novelty* was positively correlated with the first and negatively with the second preference dimension.

In the *web design business* study only two constructs could be projected in the 2-dimensional preference space. Interpretation of the preference dimensions solely based on these constructs indicate that the most important component is more related to *visual appeal* and the second to *novelty*. It is noteworthy that none of the descriptive attributes as well as none of the constructs *credibility* and *perceived usability* were useful in interpretation of the first two preference components in this study. The attributes *order, credibility* and *perceived usability* were actually found to be highly correlated with the third dimension that represented only 13.6 % of preference variance.

Considering the average preference direction of the participant sample the most important predictor of website preference was *visual appeal* in all studies (as in [2]). However, in the healthcare study *visual appeal* was highly correlated with *perceived usability* and *credibility*. In addition, visual characteristics that were common among preferred websites in this category were *simplicity*, *order large image areas* and *white*

space. In the tourism study visual appeal was not correlated with *credibility* or *perceived usability*. Preferred websites in this category had *vivid colors*, many *images*, plenty of *with space*, and a small *text area*. In the web design business study only *novelty* could be recognized as an important design factor apart from *visual appeal*.

4 Discussion

Preference mapping analysis revealed that diverse design characteristics can be important drivers of preferences for designs that originate from different website domains. Although, the most important preference predictor in all studies was *visual appeal,* what constitutes an appealing website differed considerably among the three studied cases. In the healthcare study the constructs *perceived usability* and *credibility* were found to be equally important as *visual appeal*. Generally preferred and appealing websites in this study were *simple*, *ordered*, *spacious* designs with large *image areas*. In the tourism study participants showed an aversion towards websites that were primarily text based. Designs that used *mostly graphics, vivid colors* and large *images* of hotel rooms were the most preferred and *appealing* websites. Perceptions of *usability* and *credibility* were exclusively related to the second dimension and were therefore less important drivers of preference than in the healthcare study. Finally in the web design business study only *visual appeal* and *novelty* could be found to be important drivers of preferences in the two dimensional space. Websites in this study varied on more design factors and could be characterized as unusual or extreme compared to designs in the others studies. This was generally expected since visual design in this domain serves as a first showcase of the company's ability to produce cutting edge design. Creativity and unconventionality cannot be appropriately captured by descriptive attributes and therefore none of the expert panel or objective measures were useful in preference interpretation in this study. The constructs of *perceived usability* and *credibility* that were found to be important drivers of preference in the other two website domains were less important in this one.

These results demonstrate that different design aspects play a determining role in preference creation towards websites within a specific domain. Therefore, misleading conclusions can be drawn by relying on a fixed set of evaluative or descriptive attributes for evaluation purposes of websites in general. Use of generic questionnaires for website design evaluation could lead to consideration of less relevant attributes into the evaluation process while design aspects that are central to a specific website domain could be ignored. For example, *novelty* was found to be one of the most important drivers of preference in the third (web business), fairly important in the second (tourism), and not important at all in the first (healthcare) study. Furthermore, *symmetry* which is an item in the classical aesthetic dimension in the questionnaire of Lavie and Tractinsky [4] was not found to be an important preference attribute in any of the studies presented in this paper. In addition, use of *images* that has been found to be an important driver of preference in two out of the three studies is a design aspect that is ignored by all aforementioned multiple item questionnaires [4][5][6].

5 Conclusion

This paper reports results from three different studies in which the influence of website domain on users' aesthetic preferences has been investigated. The results suggest that diverse design aspects can influence participants overall impressions of website designs in different domains. A different set of descriptive and evaluative constructs could be identified as important drivers of preference in three distinct studies involving healthcare, tourism and web design business websites. *Visual appeal* was the most important predictor of participant preferences towards websites in all case studies. However, appealing websites had different visual characteristics in each domain. To conclude, the results in this paper demonstrate the influence of website domain in shaping users' aesthetic preferences. This finding emphasizes the need for flexible evaluation methods that do not ignore the visual design aspects that are important in specific website domains.

Acknowledgements

References

1. Tractinsky, N., Katz, A. S., & Ikar, D. What is beautiful is usable. Interacting with Computers 13, 2 (2000) 127-14.
2. Schenkman, B. N., & Jönsson, F. U. Aesthetics and preferences of web pages. Behaviour & Information Technology 19, 5 (2000), 367-377.
3. Lindgaard, G., Dudek, C., Sen, D., Sumegi, L., & Noonan, P. An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. ACM Transactions on Computer-Human Interaction 18, 1 (2011).
4. Lavie, T., & Tractinsky, N. Assessing dimensions of perceived visual aesthetics of web sites. International Journal of Human-Computer Studies 60, 3 (2004), 269-298.
5. Hassenzahl, M., Burmester, M., & Koller, F. AttrakDiff: A questionnaire to measure perceived hedonic and pragmatic quality. Mensch & Computer. Interaktion in Bewegung (2003).
6. Moshagen, M., & Thielsch, M. T. Facets of visual aesthetics. Int. Journal of Human-Computer Studies, 68, 10 (2010), 689-709.
7. Roth, S. P., Schmutz, P., Pauwels, S. L., Bargas-Avila, J. A., & Opwis, K. Mental models for web objects: Where do users expect to find the most frequent objects in online shops, news portals, and company web pages? Interacting with computers, 22, 2, (2010), 140-152.
8. Norman, D. Some observations on mentai models. Mental models 7 (1983) 7–14.
9. Tuch, A. N., Presslaber, E. E., Stöcklin, M., Opwis, K., & Bargas-Avila, J. A. The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. Int. J. of Human-Computer Studies 70, 11 (2012), 794–811.
10. Meullenet, J., Xiong, R., Findlay, C., & Knovel. Multivariate and probabilistic analyses of sensory science problems. Wiley Online Library (2007).