The Interplay Between Web Aesthetics and Accessibility

Grace Mbipom
School of Computer Science
University of Manchester
England, M13 9PL, UK
grace.mbipom@cs.manchester.ac.uk

Simon Harper
School of Computer Science
University of Manchester
England, M13 9PL, UK
simon.harper@manchester.ac.uk

ABSTRACT

Visual aesthetics enhances user experience in the context of the World Wide Web (Web). Accordingly, many studies report positive relationships between Web aesthetics and facets of user experience like usability and credibility, but does this hold for accessibility also? This paper describes an empirical investigation towards this end. The aesthetic judgements of 30 sighted Web users were elicited to understand what types of Web design come across as being visually pleasing. Participants judged 50 homepages based on Lavie and Tractinsky's *classical* and *expressive* Web aesthetics framework. A cross-section of the homepages were then manually audited for accessibility compliance by 11 Web accessibility experts who used a heuristic evaluation technique known as the Barrier Walkthrough (BW) method to check for accessibility barriers that could affect people with visual impairments. Web pages judged on the classical dimension as being visually clean showed significant correlations with accessibility, suggesting that visual cleanness may be a suitable proxy measure for accessibility as far as people with visual impairments are concerned. Expressive designs and other aesthetic dimensions showed no such correlation, however, demonstrating that an expressive or aesthetically pleasing Web design is not a barrier to accessibility.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces (D.2.2, H.1.2, I.3.6) - Evaluation/Methodology, User-centred design; K.4.2 [**Computers and Society**]: Social Issues-Assistive technologies for persons with disabilities

General Terms

Design, Experimentation, Human Factors

Keywords

User Experience, Visual Aesthetics, Web Accessibility

1. INTRODUCTION

The increasing levels of competition among organisations with an on-line presence has facilitated the need for sophisticated and fanciful visual designs. Sometimes, in a bid to attract on-line visitors, the visual appearance of a Web page is placed ahead of its functionality. How to maintain a balance between form and function has been a long standing issue for disciplines with elements of visual design [7]. The present study seeks to address a case of this design dilemma by investigating the interplay between visual aesthetics and accessibility considerations in the context of the Web.

Although visually attractive sites enhance Web experience for sighted users [10, 28, 30], these sort of sites are speculated to hinder people with disabilities, especially those with visual impairments[1] [31]. On the other hand, most Web designers perceive the accessibility initiative to be restrictive creativity-wise [24, 27]. Web accessibility practitioners' efforts to mediate between these two extremes have met with challenges, primarily because the current state of affairs seem not to support claims that appealing Web designs can go hand in glove with accessibility [27]. The nature of the interaction between the use of Web aesthetics and accessibility still remains unclear. A proper understanding of this interaction will help address some of the misconceptions that surround accessible designs. These misunderstandings have slowed down the advancement of inclusive Web design.

To aid to our investigation, the following research questions were considered: i) What types of Web design come across as aesthetically pleasing to sighted users? In order to answer this first question, Lavie and Tractinsky's Web aesthetics framework was used to classify Web pages. They found that perceived visual aesthetic aspects of the Web are "bi-dimensional", with classical and expressive aspects. The first dimension represents early visual design principles rooted in clarity and orderliness, while the second dimension highlights designs which showcase a Web designer's ingenuity [15]. ii) What is the relationship between Web aesthetics and accessibility? To answer this second question, manual accessibility audits were performed on Web pages whose aesthetic quality had already been determined in (i). Web accessibility experts examined a cross-section of the Web pages for accessibility barriers which could affect people with visual impairments. We did not employ user-testing, another effective manual accessibility evaluation technique for the following reasons: It is expensive [19]; It is also difficult to find people with disabilities of the same degree, who have

---

[1]Visual impairments here include blindness, low-vision and colour-blindness.

the same level of computing expertise [14, 25]; A person with disabilities may have more than one disability. As a result, such a person may be affected by multiple Web accessibility barriers which an evaluator may not be able to account for in one go [25, 34]; People with disabilities have different assistive technologies and personal adaptations. An evaluator may not be able to create the same environment [25]; The user's ability to use the assistive technologies available is another issue [14, 32]. We reckoned that the level of subjectivity would be greater if user-testing was employed, compared with the use of Web accessibility experts. It is important to note that no one evaluation method can identify all the accessibility issues on a Website [19, 29]. Hybrid approaches are agreed to be the most effective. In particular, expert reviews have been shown to be very effective when multiple evaluators are involved, and a combination of techniques are used [19]. We employed 11 experts, and they made use of a combination of accessibility evaluation tools and techniques.

The study reveals the relationship between Web aesthetics and accessibility to be a rather complex one. Web pages judged on the classical dimension as being visually *clean* were more readily accessible, while expressive Web designs (*fascinating* and *creative*) and other aesthetic dimensions showed no significant relationships with accessibility.

2. AESTHETICS AND WEB EXPERIENCE

The effect of visual aesthetics on user experience has been widely investigated in Web domains. Most studies report a positive relationship between Web aesthetics and the facet of user experience which they investigate. In general, Web users perceive visually attractive Websites to be more credible [13, 28], usable [1, 10, 26], useful [8, 12] and desirable [30], compared with sites that are not as visually adorned.

In the usability case, however, while some studies report a straightforward strong positive relationship, because Websites which users perceive to be aesthetically pleasing are also perceived to be easy to use (e.g [1, 12, 16, 26]). Other studies relay a more complex interaction, as such strong positive associations exist given certain conditions only. For example, when the aesthetic quality of Web pages are considered under 'classical' or 'expressive' dimensions as defined in [15], Web pages preferred on the classical dimension show stronger positive relationships with usability. Classical designs are simple and clear. Expressive designs which are usually more sophisticated in nature tend to be less closely related with usability [9, 11, 15]. This insight is particularly important for our study, since usability and accessibility are considered to be closely related facets of user experience.

Very few studies explicitly address the relationship between visual design and accessibility. In one of such studies [27] where the state of affairs in accessible design was discussed, the author highlighted the gaps between visual designs and Web accessibility, and the challenges which accessible designs present to Web designers. The study was based on the author's industrial design experiences, and the outcomes were mostly advisory in nature. No empirical data was gathered as such. In another closely related study [24] which was empirical in nature, 3 out of 100 Websites ascertained to be highly accessible by people with disabilities were also shown to have complex visual designs. The authors concluded that accessibility considerations do not prevent Web designers from creating Websites with complex visual designs. Along with the limited number of case studies

employed, their work highlights one aspect of visual designs which is complexity. Visual complexity, however, happens to be a rather weak indicator of aesthetic pleasure in Web domains [23, 33]. As such, questions regarding the relationship between the use of Web aesthetics and accessibility remain unanswered. The closest to our study is our earlier work [20] in which we established that there was a link between Web aesthetics and accessibility which required further investigations. In that preliminary study, an automated accessibility checker, Cynthia Says[2] was used to examine Web pages whose aesthetic quality had already been determined. We found that Web pages which were rated by sighted users as being *clean*, *clear* and *organised* significantly violated fewer checkpoints specified in Web Content Accessibility Guidelines (WCAG) 1.0 [6]. Web pages rated as being *beautiful* or *interesting* showed no significant relationships with the number of accessibility checkpoints violated.

The present study extends our previous work and closely related studies in a number of ways: i) More Web pages were examined here for their aesthetic quality, compared to our previous study. As such, a wider range of Web designs were taken into account considering the subjectivity of aesthetic preference ii) A more reliable method for measuring Web accessibility is employed here as well. Manual accessibility evaluations were conducted by Web accessibility experts who are well versed in the field iii) The manual evaluation technique employed here distinguishes between disability types, and takes into account the severity of accessibility barriers found a Web page, whereas automated checkers do not and finally iv) Our work extends the aforementioned closely related studies by adopting an empirical approach and a widely validated framework for Web aesthetics respectively, in order to explain the interplay between Web aesthetics and accessibility as it affects people with visual impairments.

3. STUDY 1 - AESTHETIC JUDGEMENTS

The aim of this study was to investigate how sighted users rate the visual quality of Web pages. This facilitated the classification of Web pages based on five design dimensions: *clean*, *pleasing*, *fascinating*, *creative* and *aesthetic*. Design features which moderated users' judgements were also elicited.

3.1 Participants

Thirty-two (32) participants, 25 males and 7 females were recruited for the study. Two persons (males) had invalid responses, so their data were not included in the final analysis. Consequently, there were 30 participants with valid responses, 23 males and 7 females with ages ranging from 16 to 41 and over. Twenty six (26) were undergraduate and postgraduate students from computing and life sciences departments, while 4 were professionals with nursing, teaching, veterinary medicine and engineering as their backgrounds. All the participants were frequent Web users. One person reported a mild case of colour blindness. Interviews before and after the task confirmed that the impairment he claimed did not affect the purpose of the experiment. Moreover, the inclusion or removal of his data did not influence the means or standard deviations for the participants' ratings significantly, hence, his data was included. From a cultural point of view, there were 16 Whites (British), 5 Blacks, 3 Asians, 2 Chinese, and 4 others (2 Arabics, 1 German, 1 Iraqi).

---

[2]Cynthia Says - http://www.cynthiasays.com/

3.2 Stimuli

Fifty (50) homepages were used for the study. This allowed for the investigation of a wide variety of Web designs. Homepages were used because they represent gateways to Websites. In agreement with Pandir and Knight [23], we judged that the use of visual aesthetics may be more crucial for homepages. The Web pages were randomly selected from the top 100 UK Websites as ranked by Alexa[3]. We also used randomly selected Web pages from the winners and nominees list for the Webby awards[4] under the best visual aesthetic design and home/welcome page categories for years 2005 to 2009. All the Web pages from Alexa were downloaded on the 10th of November, 2009, while the Web pages from the Webby award's Website were downloaded on the 23rd of November 2009, and then merged together to form a pool. The 50 Web pages were stored locally, and presented together with all their interactive features to the participants. The Web pages represented news, sports, entertainment, education, blogging, search, government, social networking, e-commerce and leisure genres among others.

3.3 Task and Procedure

Participants were first made to read an information sheet outlining the aim of the study. On agreeing to participate, a consent form was signed and a demographic form filled. The task was to judge the visual quality of 50 homepages. Participants were shown a homepage for 4 seconds, and were free to scroll up or down the page as they would normally do when browsing. However, they could not navigate away from the homepage. This was to ensure stimulus uniformity. After 4 seconds, another page was shown instructing participants to rate the Web page they had just seen based on the 5 design dimensions under investigation. A 7-point Likert scale was used, and the scores were written on a paper questionnaire. After rating a Web page, participants then clicked the 'next' button. This made another homepage appear for 4 seconds as before, after which a page with instructions to rate followed. The viewing time was set to 4 seconds because we were interested in visceral responses. Previous studies report viewing times ranging from 50 milliseconds [17] to 7 seconds [21] for gathering visceral responses. The pages were ordered in two different ways to counter balance any position effects, and participants had to rate twice.

Judgements were based on Lavie and Tractinsky's Web aesthetic dimensions [15], and the adjectives *clean*, *pleasing* and *fascinating*, *creative* were selected to represent the two dimensions respectively. Participants also gave an overall score under the term *beautiful/aesthetic*. A close examination of Lavie and Tractinsky's framework will reveal that some of their terms are very much related. For example, the following pairs are synonyms: clean and clear; original and creative. Hence, we chose a subset of terms that were semantically disparate. Their framework has also been criticized for the inclusion of the term 'aesthetic', firstly as a design dimension of aesthetics [17], and secondly under the classical dimension only [22]. Consequently, we gathered the overall judgements of the participants under the term 'aesthetic'. Also, the word 'beautiful' was used alongside 'aesthetic' in the questionnaire to aid people who were not familiar with the word 'aesthetic'. As we had users with differing backgrounds in mind, we did not make use of a more technical design dimension like 'symmetrical' from their framework. Qualitative feedback was gathered after the experiment. Experiment sessions lasted between 30 minutes and 1 hour depending on the participants' judgement speed, and after task discussions. The participants were paid £10 for their time.

3.4 Results

Means and standard deviations were computed for each of the 50 homepages based on the participants' ratings for the pages, given the 5 design dimensions. This allowed the Web pages to be ranked under the different aesthetic dimensions. A Pearson correlation analysis was performed between each of the four design dimensions and the overall term 'aesthetic' using the mean ratings given by the participants. Figure 1 and Table 1 show that visual cleanness had the least positive correlation with the term 'aesthetic', suggesting the term 'clean' to be most unlike the other design dimensions. Clean homepages were found to have a simple and less-dense compositional layout, usually with one main image and a mostly white background. While cluttered homepages had a complex, dense look and feel, with heavy text presence and/or segmentation of the layout.

(a) Clean vs Aesthetic   (b) Pleasing vs Aesthetic

(c) Fascinating vs Aesthetic   (d) Creative vs Aesthetic
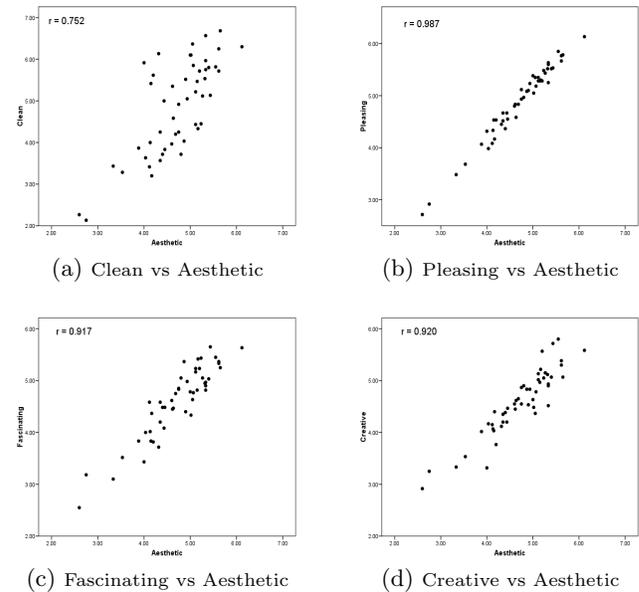
Figure 1: Relationship between the design dimensions investigated and the term 'aesthetic'

Table 1: Pearson correlation matrix for Figure 1

r	Clean	Pleasing	Fascinating	Creative	Aesthetic
Clean	1.000	0.781	0.471	0.513	0.752
Pleasing		1.000	0.898	0.891	0.987
Fascinating			1.000	0.958	0.917
Creative				1.000	0.920
Aesthetic					1.000

All correlations are significant at the 0.01 level (2-tailed)

Expressive aspects (fascinating and creative) had a stronger positive relationship with the term 'aesthetic'. Expressive Web designs made use of moderate to heavy animations and were very colourful. They also contained more images than

clean homepages. The term 'pleasing' had the strongest relationship overall with the term 'aesthetic'. Perhaps the participants perceived the term 'pleasing' to be a general design dimension as well. Recall that the participants were asked to give an overall visual quality score under the term 'aesthetic'. It was particularly difficult to determine Web page attributes which moderated participants' ratings for the design dimensions 'pleasing' and 'aesthetic'. The Web pages which topped the chart in these two categories exhibited a combination of characteristics found in the other design dimensions. A similar concern is raised in [22].

In general, if a page received a high rating for one design dimension, it was likely to receive similar ratings for other dimensions, suggesting a positive link between the terms. The correlation analysis in Table 1 confirms this. However, this was not always the case for homepages which topped the chart as being visually clean. Some of them performed poorly in the expressive design realm, hence the moderate correlations between the design dimensions *clean* and *fascinating* (r = 0.471), as well as *clean* and *creative* (r = 0.513).

4. STUDY 2 - ACCESSIBILITY AUDITS

The purpose of this study was to investigate the accessibility level of a cross-section of homepages whose visual aesthetic quality had been determined in Study 1. Based on insights gained from the usability case discussed earlier in the background and related work section (Section 2), we conjectured that Web pages preferred on the classical dimension were likely to have fewer accessibility barriers, compared with Web pages classed as being expressive in their designs.

4.1 Participants

Nineteen (19) judges were contacted by e-mail. They were people who currently work in the accessibility area, and belong to research groups that focus on inclusive design across academia and industry. Seventeen (17) of them volunteered to do the study, but 12 judges carried out the accessibility audits and returned their completed evaluation results. Others sent in their apologies. One (1) out of the 12 had invalid questionnaire responses, hence the associated data was not used. There were 8 males and 3 females with ages ranging from 26 to 50 years (Mean = 36.2 and SD = 7.76). All the judges were fluent in English language except one judge who had intermediate skills. The judges had experience working in the accessibility area for number of years ranging from 1 to 15 (Mean = 7.5 and SD = 4.76). None of the evaluators were beginners to the Web accessibility area, they rated themselves as having intermediate or expert skills in the area. Six (6) rated themselves as intermediate, while 5 rated themselves as experts. Four (4) of the judges had worked as Web accessibility consultants in the past. The judges evaluated one or more Web pages, depending on how much time they were willing to spare.

4.2 Stimuli

Sixteen (16) out of 50 homepages which had previously been rated for their aesthetic quality were selected for Study 2. Ten (10) pages were first chosen by arranging all 50 homepages in descending order of their overall aesthetic quality, and choosing every fifth page. The selected pages therefore spanned the best, average and worst pages in terms of aesthetic quality. Six (6) extra pages were then added to the list. These were pages which were consistently in the top

or bottom positions under the visual design categories (i.e clean, pleasing, fascinating, creative and aesthetic) previously investigated. This was to enable us report findings on the interplay between various Web design dimensions and accessibility as well. Table 2 contains the selected Web pages, together with their study IDs and Web addresses.

Table 2: Web pages with their study IDs and Urls

Page Name	PageID	Web address
Villa San Michelle	ID1	http://www.villasanmichele.com/web/ovil/villa_san_michele.jsp
Google UK	ID2	http://www.google.co.uk/
Good things-Orange	ID3	http://awards.goodthingsshouldneverend.com/
Full Sail University	ID4	http://www.fullsail.edu/
Whalehunt	ID5	http://thewhalehunt.org/
BBC	ID6	http://www.bbc.co.uk/
Askjeeves	ID7	http://uk.ask.com/?o=312&l=dir
Solar System-NASA	ID8	http://solarsystem.nasa.gov/index.cfm
Wordpress	ID9	http://wordpress.com/
MSN	ID10	http://uk.msn.com/
Virgin Media	ID11	http://www.virginmedia.com/
Rapidshare	ID12	http://www.rapidshare.com/
Gumtree	ID13	http://www.gumtree.com/
Directgov	ID14	http://direct.gov.uk/en/index.htm
Money Saving Expert.com	ID15	http://www.moneysavingexpert.com/
Ezine Articles	ID16	http://ezinearticles.com/

4.3 Task and Procedure

On agreeing to participate, the judges were assigned a judge number and sent the study materials via e-mail. The study pack comprised of a Participant Information Sheet (PIS), demographic information sheet, Web page(s) to be evaluated, barrier-checklist spreadsheet(s) and a post-evaluation questionnaire(s). The judges were instructed to read the PIS which outlined the aim of the study and further instructions.

The following user categories were selected for investigation after discussions between 4 of the judges: *Blind:* people who cannot see and have to use screen readers to access the Web; *Low-vision:* people who see partially and require screen magnifiers, accessibility features offered by operating systems and maybe screen readers to access the Web; *Colour-blind:* people who cannot distinguish between certain colours [2]. The user categories chosen for investigation were restricted to people with visual impairments, because they represent a large population of people with disabilities who have access to the Web [19]. Other disability types were not considered, because we did not want to overburden the judges. Manual accessibility audits are very time consuming. The evaluators were asked to imagine that the user goal was browsing or information search. The evaluators judged the Web pages independently, and in their own work environment using the Barrier Walkthrough method.

The Barrier Walkthrough (BW) method: is an accessibility evaluation method adapted from heuristic evaluation techniques widely used in usability engineering [2]. The heuristics used as checkpoints in the usability case are replaced with *barriers* in the BW method. An *accessibility barrier* prevents a person with disabilities from achieving his/her set goals when interacting with a Web application. Barriers are derived from known accessibility guidelines, and can be described in terms of i) user category, ii) assistive technology used, iii) goal/task being hindered and iv) Web page features which trigger the barrier in question. A BW evaluator must take the following steps: i) define of the user category (e.g blind, deaf, mobile device user e.t.c), ii) define user goals (e.g casual browsing, e-shopping e.t.c), iii) check the selected Web pages for barrier presence and iv) determine the severity of each barrier as {0,1,2,3}, meaning *none*, *minor*, *significant* or *critical*. According to Brajnik [3], the
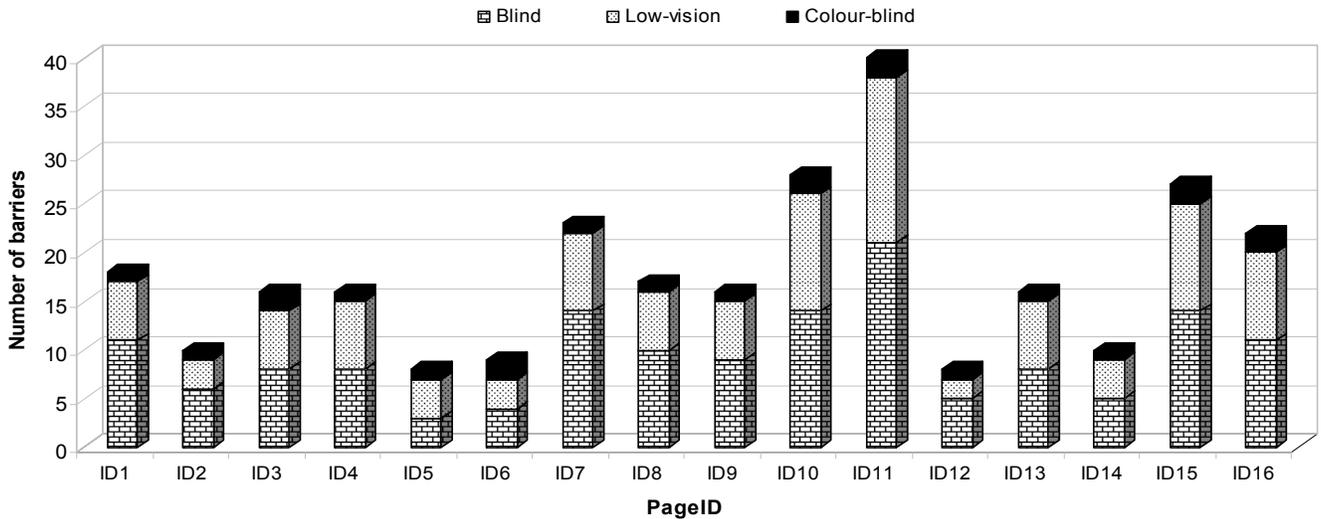
Figure 2: Accessibility barriers found on the Web pages based on average severity ratings

Table 3: Correlation between the aesthetic dimensions, number of barriers and other subjective quantities

r	Clean	Pleasing	Fascinating	Creative	Aesthetic	Barriers	Time	Confidence	Effort	Productivity
Clean	1.000	0.868^a	0.637^a	0.594^b	0.835^a	-0.501^b	-0.246	0.033	-0.417	0.047
Pleasing		1.000	0.930^a	0.898^a	0.994^a	-0.342	-0.228	-0.118	-0.184	-0.011
Fascinating			1.000	0.972^a	0.947^a	-0.158	-0.120	-0.218	0.011	-0.049
Creative				1.000	0.921^a	-0.129	-0.137	-0.266	0.008	-0.065
Aesthetic					1.000	-0.294	-0.218	-0.139	-0.163	-0.022
Barriers						1.000	0.378	-0.433	0.783^a	-0.439
Time							1.000	-0.220	0.517^b	-0.378
Confidence								1.000	-0.547^b	0.740^a
Effort									1.000	-0.655^a
Productivity										1.000

a is significant at the 0.01 level, while b is significant at the 0.05 level (2-tailed).

the BW method is believed to be educative for novice evaluators. It is also effective in identifying severe accessibility issues and reducing false positives in the evaluation process. However, like many other heuristic evaluation techniques, the BW method suffers from an evaluator effect [5, 35].

On completing the manual accessibility audits, the judges were required to fill a post-evaluation form. The form captured the time taken to do the evaluations, the judges' confidence in auditing the Web page(s), the effort required and their productivity levels. The completed accessibility audits were then e-mailed back to the principal investigator.

4.4 Results

We received 37 valid BW reports for the 16 Web pages investigated. Figure 2 shows the number of barriers found on a Web page which received an average severity rating of at least 1 from the judges which examined the page in question. The highest number of barriers were found on ID11, while the least numbers were found on ID5 and ID12. From the BW reports, 'spaced titles', 'ASCII art' and 'pages without titles' were the 3 accessibility barriers capable of affecting people with blindness that were not found on any of the examined Web pages. 'Widely formatted forms' were not found on any page examined for barriers which could affect

people with low-vision. The most common barrier found which could affect persons with blindness across the Web pages was 'generic or ambiguous links', while 'insufficient visual contrast' was the most common barrier found which could affect people with low-vision and colour-blindness. In general, more barriers were found capable of affecting people with blindness, followed by low-vision and colour-blindness.

Table 3 shows a Pearson correlation analysis which was performed between the aesthetic dimensions (ratings obtained from Study 1), the number of accessibility barriers found and other subjective ratings given by the judges. Although all the aesthetic dimensions showed negative relationships with the number of barriers found on the examined homepages, only visual cleanness had a moderate significant correlation with the number of barriers found. Visually clean homepages had fewer accessibility barriers. Furthermore, the number of barriers found on a Web page was positively related with the effort required to perform the accessibility evaluations. The time taken for the accessibility evaluations was positively related with the effort required to do the evaluations. The judges' confidence in identifying the accessibility barriers was negatively related with the effort expended, and positively related with productivity. Effort expended was negatively related with productivity.

Quality Assessment (QA): It is also important that we evaluate the effectiveness of the BW method here. One way of doing this is through a reliability test. Reliability is the extent to which independent accessibility audits produce the same results [4]. One way of measuring reliability is by the coefficient of variation (cv). It is defined as $\frac{SD}{M}$, where SD is the standard deviation and M is the mean of the number of correctly identified barriers. The smaller the *cv* value, the more reliable the audit results. Whenever the cv value exceeds 1, it depicts low reliability as SD is greater than M. Similar quality assessment measures are used in [4, 5, 35]. The cv value here measures the variation between the barrier severity ratings given by different judges for barriers on the same Web page, and for different disability types. It is represented by the triple (barrier type, disability category, page). Since this value can be influenced by the number of evaluators, we performed the reliability tests with the same number of judges per page. Two judges were chosen per page. For Web pages which had evaluations from more than two judges, two sets of audits were randomly chosen. ID4 and ID10 had lone evaluations due to invalid questionnaire responses from the second judge who was assigned the Web pages in question, hence no reliability checks could be performed. The results for the QA are shown in Table 4.

Table 4: Coefficient of variation (cv) among judges for the Web pages evaluated both at the page level and for the three disability types investigated - Blind (BL), Low-Vision (LV) and Colour-Blind (CB)

PageID	Page	BL	LV	CB
ID1	0.21	0.26	0.16	0.24
ID2	0.23	0.29	0.10	0.70
ID3	0.21	0.22	0.17	0.70
ID4	n/a	n/a	n/a	n/a
ID5	0.13	0.12	0.15	0.00
ID6	0.15	0.06	0.17	1.40
ID7	0.27	0.32	0.17	0.70
ID8	0.17	0.18	0.14	0.35
ID9	0.26	0.25	0.22	0.70
ID10	n/a	n/a	n/a	n/a
ID11	0.50	0.47	0.55	0.35
ID12	0.11	0.13	0.05	0.70
ID13	0.29	0.29	0.29	0.14
ID14	0.22	0.20	0.21	0.70
ID15	0.47	0.49	0.46	0.24
ID16	0.27	0.18	0.30	1.40

At the page level, we observe reasonable variations between the independent accessibility audits performed by the judges. The highest variation between judges was observed for ID11. The Web page in question had the highest number of barriers with an average severity rating of at least 1. Accordingly, we would expect greater variances in the opinions of the judges involved. When split between disability types, the results show reasonable variations for the blind and low-vision categories. The cv values are rather high for colour-blindness on some of the homepages. This is because the BW method specifies only two types of barriers capable of affecting people with colour-blindness, which are 'color is necessary to understand information' and 'insufficient visual contrast'. Whenever the judges disagreed on the barrier severity ratings for this user category the effect was greater. Going by

the page level reliability scores, the accessibility audits are quite reproducible among judges. This also boosts our confidence in the audits performed on ID4 and ID10, as the lone judge with valid questionnaire responses also helped in evaluating the accessibility levels of other Web pages.

In one study [5] where the role of expertise in Web accessibility evaluations was investigated, the authors found that a lone expert judge was capable of identifying 70% of the problems on a Website. With 2 judges 94% of the problems were covered and 3 judges covered all. While more than one expert judge per Web page is encouraged, they also found that the reproducibliity of the accessibility evaluations reduce as the judge numbers per page increase. This is due to the increased subjectivity arising from the many different opinions. Although the authors point out that not all of their study outcomes are generalizable, as they either apply to the experts they employed or the Websites evaluated. Their findings still give us an idea of the effect of expertise on our own studies, especially as we employed some of the Web accessibility experts that they used for their work.

5. DISCUSSION

The overall objective of this paper was to investigate the interplay between Web aesthetics and accessibility as it affects people with visual impairments. In the usability case discussed in the background section (Section 2), Web pages on the classical dimension were perceived to be easier to use, while expressive Web designs were perceived to be more difficult to use, because of their complex and sophisticated visual designs [9, 11]. We had speculated that a similar situation will hold for accessibility, since usability and accessibility are closely related facets of user experience. Results from Study 2 show the relationship between Web aesthetics and accessibility to be rather complex. The only design dimension that showed significant relationships with accessibility was visual *cleanness*. We observed a moderate significant negative correlation between the design dimension *clean* and the number of accessibility barriers present on a Web page. So, the cleaner the homepage, the fewer the accessibility barriers. Perhaps this outcome is due to the small number of HTML elements required to build visually clean Web pages. A Web page with a simple HTML code base is more likely to have fewer accessibility issues, because the Web designer typically has fewer bytes of code to mind. A similar outcome is reported in [18] where a correlation was observed between Web page complexity levels (measured by the number of HTML elements present) and accessibility quality. Web pages with fewer HTML elements had better accessibility quality. The authors believe that simplicity minimizes the occurrence of accessibility issues, and makes verification of accessibility compliance manageable during Web development.

The other classical design dimension investigated, 'pleasing' showed no significant relationships with accessibility. We had established in Study 1 that it leaned more towards the expressive, because it showed strong positive correlations with expressive design dimensions (*fascinating* and *creative*), and the overall term 'aesthetic'. In the original aesthetics framework developed by Lavie and Tractinsky, the design dimensions clean, pleasing and aesthetic all come under classical aesthetics. Notice that in Table 3, 'pleasing'(r = -0.342) and 'aesthetic' (r = -0.294) follow up closely behind 'clean' (r = -0.501) on their correlation with the number of barriers present. However, our study reveals the design di-

mension *clean* to be most unlike the other two investigated. Hence, its unique relationship with accessibility here. In our previous study [20], significant correlations were also observed between similar design dimensions (clean, clear and organised) and accessibility. This class of Web designs are characterised by simplicity and minimalism. They primarily address design clarity rather than appeal or affective aspects. It may be this simplicity that the design community struggles with or misunderstands. It is commonly believed that accessible Websites are boring, because of the perceived minimalistic undertone to their designs [24, 27]. Many Web designers fail to see simplicity as an aesthetic notion [13].

Our findings do not fully support a closely related study where highly accessible Websites were also shown to have complex visual designs [24]. In our case, it was more common for simple Web designs to have fewer accessibility issues. A similar situation to ours is reported in [18]. Expressive designs which would normally come across as being more complex and sophisticated showed no significant relationships with accessibility. However, further studies would be required to draw stronger conclusions across related studies. Consequently, this does not rule out the fact that fanciful or expressive Web designs cannot be created in an accessible fashion. Our data here, and indeed from our previous study [20] only shows this to be less the case presently. Expressive designs are not necessarily a barrier to accessibility. Since the relationship between classical aesthetics and functionality is more consistently established across studies, Web designers need to embrace simplicity in visual design. We also acknowledge the fact that there may be other factors moderating the interplay between Web aesthetics and accessibility which we did not investigate here. Some examples include the tools used for Web development, the Web designer's background and level of accessibility awareness, and an individual's or organisation's adherence to accessibility guidelines, either for business or legal reasons.

Furthermore, our findings suggest a link between the visual appearance of a Website and its underlying functionality. At a very high level, the way a Web page looks could be used to predict how easy or difficult Web users would find the Web page in question. Existing studies establish links between perceived aesthetic aspects of Web pages and their perceived usability [1, 12, 16, 26]. However, none of these studies explore the possibility of predicting functionality based on visual appearance. Such a system is diserable, as it could provide a rough overview of any functionality issues, before actual user evaluations are employed. Our study suggests visual cleanness to be a suitable proxy measure for accessibility, compared with other aesthetic dimensions. It is important to note, however, that we do not suggest that such a system replaces manual accessibility evaluations.

6. CONCLUSION

Using two sets of studies, this work provides empirical evidence for the interplay between visual aesthetics and accessibility in the context of the Web. In the first study, we elicited the aesthetic judgements of sighted Web users in order to understand what types of Web pages were aesthetically pleasing or displeasing. The adjectives *clean*, *pleasing*, *fascinating*, *creative* and *aesthetic* were used to categorize Web pages. In the second study, a cross-section of the Web pages from the first study were examined by experts for accessibility barriers which could affect people with visual

impairments. This was to enable us understand the interactions between the use of Web aesthetics and accessibility.

Our results show only one aesthetic dimension, *clean* to be significantly related with accessibility. Clean Web designs were characterised by simplicity and minimalism. They had fewer accessibility barriers compared with the other aesthetic dimensions investigated. Our data suggests this design dimension to be a suitable proxy measure for accessibility as far as people with visual impairments are concerned. This research contributes to our knowledge by further lending support to the existing relationship between classical aesthetic aspects and functionality, in this case accessibility.

As a next step, we intend to build a tool which is capable of predicting the accessibility level of Web pages based on the clarity of their visual designs. A further evaluation will also be undertaken to investigate the efficacy of our tool. Further details on the studies can be found in their associated technical reports available on-line at http://weleprints.cs.manchester.ac.uk/view/subjects/eivaa.html

8. REFERENCES

[1] L. Brady and C. Phillips. Aesthetics and Usability: A Look at Color and Balance. *Usability News, February*, 5 (1), 2003.

[2] G. Brajnik. Barrier Walkthrough: Heuristic Evaluation Guided by Accessibility Barriers. *Available Online at http://sole.dimi.uniud.it/ giorgio.brajnik/projects/bw/bw.html (Accessed 08 September, 2010).*

[3] G. Brajnik. Web Accessibility Testing: When the Method is the Culprit. *Lecture Notes in Computer Science, Springer Berlin / Heidelberg*, pages 156–163, 2006.

[4] G. Brajnik. A Comparative Test of Web Accessibility Evaluation Methods. *In Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility, Nova Scotia, Canada*, pages 113–120, 2008.

[5] G. Brajnik, Y. Yesilada, and S. Harper. The Expertise Effect on Web Accessibility Evaluation Methods. *Human Computer Interaction, Taylor and Francis*, 2011.

[6] W. Chisholm, G. Vanderheiden, and I. Jacobs (Eds.). Web Content Accessibility Guidelines 1.0. *W3C Recommendation 5th May, 1999. http://www.w3.org/TR/WAI-WEBCONTENT/*, 1999.

[7] R. Crozier. Manufactured Pleasures: Psychological Responses to Design. *Manchester University Press*, 1994.

[8] D. Cyr, M. Head, and A. Ivanov. Design Aesthetics Leading to M-loyalty in Mobile Commerce. *Information and Management, Elsevier B.V.*, 43 (8):950 – 963, 2006.

[9] A. De Angeli, A. Sutcliffe, and J. Hartmann. Interaction, Usability and Aesthetics: What Influences Users' Preferences? *In Proceedings of the 6th Conference on Designing Interactive Systems, University Park, PA, USA, ACM Press*, pages 271 – 280, 2006.

[10] J. Hartmann, A. Sutcliffe, and A. D. Angeli. Investigating Attractiveness in Web User Interfaces. *In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, San Jose, California, USA*, pages 387 – 396, 2007.

[11] J. Hartmann, A. Sutcliffe, and A. D. Angeli. Towards a Theory of User Judgment of Aesthetics and User Interface Quality. *ACM Transactions on Computer-Human Interaction (TOCHI), Article No. 15, ACM New York, NY, USA*, 15(4), 2008.

[12] H. Heijden. Factors Influencing the Usage of Websites: The Case of a Generic Portal in the Netherlands. *Information and Management, Elsevier Science B.V.*, 40 (6):541 – 549, 2003.

[13] K. Karvonen. The Beauty of Simplicity. *In Proceedings Conference on Universal Usability, Arlington, Virginia, USA, ACM Press.*, pages 85 – 90, 2000.

[14] B. Kelly, D. Sloan, L. Phipps, H. Petrie, and F. Hamilton. Forcing Standardization or Accommodating Diversity? A Framework for Applying the WCAG in the Real World. *In Proceedings of the International Cross-disciplinary Workshop on Web Accessibility (W4A), Chiba, Japan, ACM*, pages 46 – 54, 2005.

[15] T. Lavie and N. Tractinsky. Assessing Dimensions of Perceived Visual Aesthetics of Web Sites. *International Journal of Human-Computer Studies, Academic Press, Inc. Duluth, MN, USA*, 60(3):269 – 298, 2004.

[16] Y.-M. Li and Y.-S. Yeh. Increasing Trust in Mobile Commerce Through Design Aesthetics. *Computers in Human Behaviour, Elsevier Ltd.*, 26 (4):673 – 684, 2010.

[17] G. Lindgaard, G. Fernandes, C. Dudek, and J. Brown. Attention Web Designers: You Have 50 Milliseconds to Make a Good First Impression! *Behaviour and Information Technology. Taylor and Francis*, Vol. 25, No. 2:115–126, 2006.

[18] R. Lopes, D. Gomes, and C. Luis. Web Not For All: A Large Scale Study of Web Accessibility. *In Proceedings of the International Cross Disciplinary Conference on Web Accessibility (W4A)*, 2010.

[19] J. Mankoff, H. Fait, and T. Tran. Is Your Web Page Accessible?: A Comparative Study of Methods for Assessing Web Page Accessibility for the Blind. *In Proceedings of the SIGCHI conference on Human factors in computing systems, Portland, Oregon, ACM Press*, 41 - 50, 2005.

[20] G. Mbipom. Good Visual Aesthetics Equals Good Web Accessibility. *ACM SIGACCESS Newsletter, January 2009*, 93:75 – 83, 2009.

[21] E. Michailidou, S. Harper, and S. Bechhofer. Visual Complexity and Aesthetic Perception of Web Pages. *In Proceedings of the 26th ACM International Conference on Design of Communication, Lisbon, Portugal, ACM*, pages 215 – 224, 2008.

[22] M. Moshagen and M. Thielsch. Facets of Visual Aesthetics. *International Journal of Human-Computer Studies, Elsevier Ltd*, 68 (10):689–709, 2010.

[23] M. Pandir and J. Knight. Homepage Aesthetics: The Search for Preference Factors and the Challenges of Subjectivity. *Interacting with Computers, Elsevier Science Inc. NY, USA*, 18(6):1351–1370, 2006.

[24] H. Petrie, F. Hamilton, and N. King. Tension, What Tension? Website Accessibility and Visual Design. *In Proceedings of the international cross-disciplinary workshop on Web Accessibility (W4A), New York city, New York*, 63:13 – 18, 2004.

[25] H. Petrie, F. Hamilton, N. King, and P. Pavan. Remote Usability Evaluations with Disabled People. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montréal, Québec, Canada*, pages 1133 – 1141, 2006.

[26] C. Phillips and B. Chaparro. Visual Appeal vs. Usability: Which one Influences User Perceptions of a Website More? *Usability News, October*, 11 (2), 2009.

[27] B. Regan. Accessibility and Design: A Failure of the Imagination. *In Proceedings of the international cross-disciplinary workshop on Web accessibility (W4A), New York, ACM Press*, 63:29 – 37, 2004.

[28] D. Robins and J. Holmes. Aesthetics and Credibility in Web Site Design. *Information Processing and Management: An International Journal*, 44(1):386–399, 2008.

[29] M. Rowan, P. Gregor, D. Sloan, and P. Booth. Evaluating Web Resources for Disability Access. *In Proceedings of the 4th International ACM Conference on Assistive Technologies, Arlington, Virginia, USA, ACM*, pages 80 – 84, 2000.

[30] B. Schenkman and F. Jönsson. Aesthetics and Preferences of Web Pages. *Behaviour and Information Technology*, 19 (5):367–377, 2000.

[31] D. Sloan, P. Gregor, M. Rowan, and P. Booth. Accessible Accessibility. *In Proceedings Conference on Universal Accessibility, Arlington, Virginia, USA*, pages 96 – 101, 2000.

[32] D. Sloan, A. Heath, F. Hamilton, B. Kelly, H. Petrie, and L. Phipps. Contextual Web Accessibility - Maximizing the Benefit of Accessibility Guidelines. *In Proceedings of the International Cross-disciplinary Workshop on Web accessibility (W4A), Edinburgh, UK, ACM*, 134:121–131, 2006.

[33] A. Tuch, J. Bargas-Avila, K. Opwis, and F. Wilhelm. Visual Complexity of Websites: Effects on Users' Experience, Physiology, Performance, and Memory. *International Journal of Human Computer Studies*, 67 (9):703 – 715, 2009.

[34] M. Vigo, A. Kobsa, M. Arrue, and J. Abascal. User-tailored Web Accessibility Evaluations. *In Proceedings of the eighteenth conference on Hypertext and Hypermedia, Manchester, UK*, pages 95 – 104, 2007.

[35] Y. Yesilada, G. Brajnik, and S. Harper. How Much Does Expertise Matter?: A Barrier Walkthrough Study With Experts and Non-experts. *In Proceedings of the 11th international ACM SIGACCESS Conference on Computers and Accessibility, Pittsburgh, Pennsylvania, USA*, pages 203 – 210, 2009.