The effects of screen ratio and order on information retrieval in web pages

Paul van Schaik*, Jonathan Ling

Psychology Section, School of Social Sciences, University of Teesside, Middlesbrough, TS1 3BA, UK

Abstract
Aesthetics can have a pivotal role in the success of Web sites [Interacting with Computers 15 (2003) 429], but need to be considered in conjunction with task performance. Typical Web pages divide the screen into a menu area for navigation and a content area to display information. The Golden Section can be applied to decide on the proportion of these two areas on the screen when designing this type of page. Using a computer-controlled experiment, we examined the role of aesthetics in the use of Web pages. Ninety-eight participants completed an information retrieval task. Analysis of variance showed that the effect of screen ratio (width of menu area to content area) on task performance and subjective outcomes was statistically significant. Application of the Golden Section was found to result in the worst screen ratio. The results are discussed in terms of theories of categorisation and aesthetic behaviour, and design recommendations are given.

q 2004 Elsevier B.V. All rights reserved.

Keywords: Aesthetics; Golden section; Web; Usability

1. Introduction

Over the last decade, the Internet has become increasingly important as a market place, with income generated from Internet-based transactions increasing year on year [1]. The greater commercial profile of the Web has made it imperative for businesses to improve their Web sites. Even for non-commercial Web sites, Web designers find themselves operating within an increasingly competitive market in which the information they have to offer needs to be easily accessible: the attention of users must be acquired quickly before they move on to better, or more appealingly designed sites. Indeed, usability needs serious consideration because when faced with a poorly usable Web site, users are likely to abandon these sites in favour of more usable ones [2]. It is important to develop accessible and visually appealing Web sites because these qualities can have a significant effect on the preferences of users.

Schenkman and Jonsson [3] list several reasons that may cause users to abandon a site: technical problems (such as slow download times), content-related (the site lacks the information required by the user) or form-related problems (badly designed home pages). Usability, particularly screen design, is of paramount importance for Web pages. Analysis of corporate sites has shown that users are significantly less likely to stay on a page that is badly designed for long enough to make a purchase; in other words, badly designed pages lose companies money [2]. Nielsen’s survey examining users’ ability to perform tasks on 20 US-based Web sites found the average completion rate was only 56%, due to the poor usability of the sites. On the basis of these considerations and findings, two conclusions can be drawn. First, that Web sites are not always easy to navigate because of poorly designed interfaces and second, that subjective impressions of a site, irrespective of a site’s usability, may influence the likelihood of a site being used.

An issue that has been identified in Web design is the tension between creativity and usability [4]. Nielsen [5], however, claims that this tension is only apparent and that there is ample scope for creativity once usability considerations have been carefully addressed. Unfortunately, although some empirical research on the use of Web sites has been published [6], many guidelines that currently exist are often not based on such investigations [7,8].

Screen layout has been the focus of a wide range of research; authors have concluded that screens should be structured in an organised manner to improve visual search and therefore usability [7,9–11]. Such organisation can be accomplished through the use of frames which divide web pages into two areas: one for navigation and one for content. A wide range of parameters impact on the usability of web pages in terms of the time take to perform tasks or in the accuracy of task performance [7]. Objective measures are commonly used to measure the usability of Web sites.
However, attention has moved towards using subjective measures to explore user satisfaction with web sites [12]. Subjective or preference data can be helpful in developing web sites that are visually appealing, as well as being used in conjunction with objective data to produce sites that are more successful—whether ‘success’ means an improvement in user satisfaction or higher levels of task performance or a combination of these. The aesthetic perceptions of users can be investigated using subjective measures. Researchers have only recently attempted to use psychological research to improve web pages [13], although the conclusion is clear: sites designed using principles derived from psychological research are superior to those that violate these principles.

Although the importance of aesthetics has long been recognised in the arts, researchers have only recently begun to investigate the implications of aesthetics for the development of visual displays [14–16] and web pages [3]. Users’ subjective perceptions of a Web site may be closely linked to the length of time they will spend on a site, as well as the likelihood of them returning to a site in the future. Indeed, Kurosa and Kashima [17] found a strong relationship between judgements of the aesthetic appeal of interfaces and their apparent usability. Given the competition between Web sites, particularly those of a commercial nature, there is a clear need to go beyond a sole focus upon usability and develop what [18] has termed ‘seductive user interfaces’ that will not only be easy to navigate but are also aesthetically appealing.

Analysis of the influence of aesthetic differences in web pages on user preferences is a new field of investigation. In one of the few studies to examine this topic, [3] asked participants to rate a series of 13 web pages in terms of their aesthetic appeal. Analysis of responses indicated that participants considered overall impression, beauty and meaningfulness as important determinants of the visual appeal of screens. Such factors, Schenkmans and Jonsson argue, should be taken into account when web pages are designed. They recommend that web designers make use of illustrations to create good impressions in users, as well as presenting overviews of sites to help orientation. Unfortunately participants in their study were not required to interact with the pages they saw, simply to subjectively rate each page. More robust design recommendations must be derived from the use of live web sites rather than simply viewing static images that require no user input.

One focus of researchers investigating the aesthetic appeal of user interfaces has been the Golden Section (GS). The golden section or golden ratio was defined by the Ancient Greeks as the most pleasing way to divide a line [19]. The GS has since had a significant impact on both art and architecture. The GS has a value of 0.618 (i.e. a line will be divided approximately 60/40). According to Fechner [20], who was the first to examine the GS from a psychological perspective, individuals will have a preference for rectangles that have a ratio of 1:1.618 for their width to height. There have been a number of attempts made to replicate Fechner’s findings [21]. For many researchers, the GS is ‘nothing more than a bit of ancient ‘metaphysical speculation’ that refuses to go away’ (Green, p. 953). However, other investigators have found some support for preference for the GS [21], or at least that it is the ‘least disliked’ ratio [22].

Benjafield and McFarlane [21] investigated the effect of context on perceptions of screen ratio. Participants were presented with three groups of rectangles, which had proportions of 0.56–1.00, 0.40–0.83 and 0.20–0.66. Rectangles were displayed on a video monitor and participants were able to modify their proportions by use of a joystick. Benjafield and McFarlane showed participants the full set of rectangles, although they varied the context of presentation in terms of the dominant proportions that were presented to these groups. They found that there was a significant effect of context on preference. In two of the three contexts, proportions close to the GS were preferred, however, when the first series of proportions presented were predominantly less than the GS, participants had a mean preference that was also lower than the GS. In other words, contextual determinants of preference appeared to be more dominant than other aesthetic influences; Benjafield and McFarlane concluded that the preferences of individuals were directly related to their prior exposure to stimuli.

Nonetheless, the GS appears to have some degree of promise as a means of improving aesthetic appeal, and now that the focus of usability researchers has turned to users’ subjective impressions of their sites, use of the GS and other aesthetic principles in screen design may be one way to improve user satisfaction.

Ngo and Ch’ng [15] reformatted six screens according to principles of dynamic symmetry, which involved altering the proportions of text boxes and pictures including ratios based on the GS. The original and reformatted screens were presented simultaneously and participants chose the screen they preferred. In all six cases, participants preferred screens that had been reformatted. Ngo and Ch’ng’s analysis of 150 screens from a variety of multimedia sources indicated that only 40% used dynamic symmetry in their construction. If the GS was found to be a useful means of improving the aesthetic appeal of web pages, use of this ratio should be encouraged in the development of relevant web sites.

The current study aims to investigate whether aesthetic considerations in terms of the ratio of the width of screen areas have an impact on the ability of users to find information in web pages. According to Nygren [23] search time can be up to 83% faster if a target object is made visually conspicuous, however, it remains unclear whether task performance is also enhanced if subjective or aesthetic improvements are made to web pages. The current study investigated the effects of screen ratio and distribution of ratios over time (order or context to use Benjafield and McFarlane’s [21] terminology) in Web pages.
on task performance and aesthetic judgements, using an information retrieval task.

2. Method

2.1. Experimental design

A 3 × 5 mixed experimental design was used with independent variables of order (distribution of screen ratios over time) used with independent measures, and screen ratio, used with repeated measures. Three orders were used: (1) web pages with ‘narrower’ navigation areas in the first half of trials followed by web pages with ‘wider’ navigation areas in the second half, (2) an equal distribution of ratios over all trials in the first half and in the second half and (3) web pages with wider navigation areas in the first half of trials followed by web pages with narrower navigation areas in the second half. Screen ratios were defined by the percentages of screen width occupied by the navigation area (on the left) and the content area (on the right). The following five ratios were used: 18.2:81.8, 23.2:76.8, 28.2:71.8, 33.2:66.8 and 38.2:61.8, with the last ratio corresponding to the Golden Section (see Fig. 1). Web pages were randomised per participant within the first half and second half of trials.

2.2. Participants

There were 98 undergraduate participants, consisting of 85 females and 13 males. Of these, 36 took part in the ‘narrow menus first’ (order 1) condition, 35 in the ‘equal distribution’ condition (order 2) and 27 in the ‘wide menus first’ condition (order 3). Participants’ mean age was 25 years with a standard deviation of 8.2.

Almost all participants (98%) had some experience of using the Web. Only two participants had never used the Web before. Participants had been using the Web from anywhere between less than one month to more than a year, however, the majority (86%) had been using the web for more than one year. In terms of frequency, the largest group of participants (44%) used the Web more than once a week but less than once a day, with a further 47% of participants using the Web approximately at least once a day.

2.3. Materials and apparatus

A set of six web sites was produced for the experiment. Each site consisted of a hierarchy of three levels. Each site was produced in five versions, corresponding to the five screen ratios. The first site consisted of 16 pages, was used for the practice experiment and represented the domain of gardening. The other five sites, each consisting of 30 pages, were used for the main experiment and represented five domains: sport, shopping, music, software and computer equipment.

The experiment ran on personal computers (Intel Pentium, 333 MHz, 64 Mb RAM, Microsoft NT4 operating system, 14 inch monitors). The screen dimensions were 800 × 600 pixels. In order to ensure both maximum clarity and replicability, contrast was set to maximum level and brightness to minimum level. Experimental software, written in Microsoft Visual Basic version 6, used the web browser control (based on version 5 of the Microsoft Internet Explorer web browser) to display pages, which enabled participants to navigate through the sites. The program recorded all participants’ computer interaction, including the sequence of web pages followed for each task (including time spent on each page) and all answers that they gave, as well as the time taken to answer questions. Participants were asked a series of 40 questions based on information available on the web sites. Answers to the questions were one (for 20 questions) or two (for another 20 questions) links away from the home page.

Two sets of rating questions were also presented to participants. In a first rating task, participants completed a set of 30 (six items × five screen ratios) questions (DES-W) based on the display evaluation scale (DES; [24]). These questions asked participants to rate their satisfaction with display quality in terms of the width of the navigation area and content area, the size of the navigation area and content area, and finally the combination of width of navigation area and content area and combination of size of navigation area and content area. All six questions were answered on a seven-point scale, where one was bad and seven was good. Reliability analysis showed that the DES-W reliable, Cronbach’s alpha = 0.82, 0.67, 0.78, 0.74 and 0.87 for screen ratios 18:82, 23:77, 28:72, 33:67 and 38:62, respectively. Scores on each of the items were averaged to give an overall rating.

2.4. Procedure

Participants were presented with instructions before completing a practice task, which consisted of a series of five information retrieval questions presented in random order, using the gardening site. They were told that a series of questions would appear at the top of the screen. After reading each question they had to click on the button labelled ‘Show Web site’. The home page of the first site was then displayed in the browser window (see Fig. 1). Participants were instructed to find the answer to each question using the site. Once they found the answer they had to click a button labelled ‘Your answer’. A dialog box then appeared into which participants entered their answers.

\[ * \text{When removing item five from the DES-W alpha increased to 0.71. The} \] 
\[ r \text{correlation between overall DES-W scores with five and six items for ratio} \] 
\[ 23:77 \text{ was extremely high, } r = 0.96. Based on these finding together with} \] 
\[ \text{high alphas of DES-W for the other screen ratios it was decided to calculate} \] 
\[ \text{an overall score from all six DES-W items for each screen ratio and use the} \] 
\[ \text{overall scores in subsequent analyses.} \]
Once they had typed their answer, they were shown the next question. Participants were instructed to work as quickly and as accurately as possible.

When all participants had completed the practice task, any questions were answered before they went on to the main experiment. The five other sites were used and participants completed a series of 40 further randomised questions (five sites X eight questions).

After the experimental trials, participants answered rating questions (DES-W). Next, participants indicated their preference for the relative width of the navigation and content areas for of the five screen ratios. For this purpose, they used a mouse to control vertically the proportion of the screen occupied by the navigation area on the left and the content area on the right. Finally, participants answered a number of demographic questions (age, sex, use of the web) that were also presented on screen. Participants took approximately 40 min to complete the experiment.

3. Results

For the purpose of analysis, the task performance trials in which a correct answer was given were examined in terms of speed (i.e. reaction time) and efficiency (i.e. number of pages loaded). Subjective measures included display quality and preferred screen ratio. Analysis of variance (ANOVA) and simple effect tests were conducted to test the effects of screen ratio and order on outcome measures. Post-hoc analysis tests were conducted to identify differences between screen ratios.

3.1. Task performance

3.1.1. Speed

Average time-on-task for correct answers was calculated. Times-on-task were transformed logarithmically, resulting in substantially improved distribution of scores and reduced kurtosis and skew. Transformed scores were used in subsequent analyses. The effects of screen ratio, \( F(4, 380) = 22.123, P < 0.001, MS_{\text{screen ratio}} = 1.546 \), and the interaction effect of screen ratio and order, \( F(8, 380) = 3.848, P < 0.001, MS_{\text{screen ratio}} = 0.269 \), were significant (see Table 1 and Fig. 2). The effect of order was not significant, \( F(1, 387) = 2.317, P = 0.130 \), \( MS_{\text{screen ratio}} = 0.784 \). Simple comparison tests with Sidak correction showed significant differences: of ratio 33:67 with both ratios 38:62 (both \( P < 0.001 \)) and 28:72 (both \( P < 0.05 \)) for order 1; of ratio 23:77 with ratios 28:72 (both \( P < 0.01 \)) and 33:67 and 38:62 (both \( P < 0.001 \)) and between ratios 18:7 and 32:68 (both \( P < 0.05 \)) for order 2, and of ratio 32:68 with ratios 18:782 and 23:77 (both \( P < 0.01 \)), 28:72 (both \( P < 0.005 \)) and 33:67 (both \( P < 0.05 \)) and of ratio 33:67 with ratios 18:72 (both \( P < 0.05 \)) and 23:67 (both \( P < 0.01 \)) for order 3. Testing a deviation contrast, ratio 38:62 was significantly different from the average of the other ratios, \( t(97) = 7.767, P < 0.001 \).

3.1.2. Efficiency

The number of pages visited before a correct answer was given was calculated. The effect of screen ratio was significant, \( F(4, 380) = 10.235, P < 0.001, MS_{\text{screen ratio}} = 8.749 \).
Table 1
Descriptives for time-on-task for correct answers

<table>
<thead>
<tr>
<th>Order</th>
<th>Screen ratio</th>
<th>18:82; M (SD)</th>
<th>23:77; M (SD)</th>
<th>28:72; M (SD)</th>
<th>33:67; M (SD)</th>
<th>38:62; M (SD)</th>
<th>Overall; M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrow menus first</td>
<td>15.96 (7.64)</td>
<td>15.87 (5.77)</td>
<td>16.80 (6.67)</td>
<td>13.45 (4.86)</td>
<td>19.34 (8.38)</td>
<td>16.28 (5.11)</td>
<td></td>
</tr>
<tr>
<td>Wide menus first</td>
<td>14.63 (7.57)</td>
<td>14.21 (6.75)</td>
<td>16.23 (8.00)</td>
<td>17.53 (8.93)</td>
<td>21.76 (13.11)</td>
<td>16.87 (7.97)</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>15.42 (7.36)</td>
<td>14.41 (5.63)</td>
<td>16.58 (7.56)</td>
<td>15.89 (7.13)</td>
<td>20.32 (10.25)</td>
<td>16.52 (6.06)</td>
<td></td>
</tr>
</tbody>
</table>

Note: original (non-transformed) results are reported for ease of interpretation.

The effects of order and the interaction effect were not significant, both $F < 1$. Multiple comparison tests with Sidak correction demonstrated significant differences of ratio 38:62 with ratios 23:77 ($P < 0.001$), 28:72 ($P < 0.05$) and 33:67 ($P < 0.001$). The test of a deviation contrast showed that there was a significant difference between ratio 38:62 and the average of the other ratios, $t(97) = 4.084$, $P < 0.001$.

### 3.1.3. Associations between performance measures

There was a positive correlation between speed and efficiency. Pearson’s $r = 0.714$, 0.582, 0.697, 0.689, 0.718, all $P < 0.001$, for ratios 18:82, 23:77, 28:72, 33:67 and 38:62, respectively.

### 3.2. Subjective measures

#### 3.2.1. Display quality

The effect of screen layout on overall DES-S scores was significant, $F(4, 380) = 32.524$, $P < 0.001$, $\text{MS}_{\text{screen ratio}} = 14.49$ (see Table 3). The effect of order, $F(2, 95) = 1.426$, $P < 0.05$, $\text{MS}_{\text{order}} = 4.735$, and the interaction effect, $F(8, 380) = 1.914$, $P < 0.05$, $\text{MS}_{\text{interaction}} = 0.735$, were not significant. Multiple comparison tests with Sidak correction revealed a significant difference in each pair of ratios, except between ratios 33:67 and 18:82 and between ratios 33:67 and 23:77; comparisons were significant with $P < 0.001$, except between ratios 23:77 and 28:72 ($P < 0.05$). When testing a deviation contrast, ratio 38:62 was found to be significantly different from the average of the other ratios, $t(97) = 7.314$, $P < 0.001$.

#### 3.2.2. Preferred screen width

Preferred widths of the menu bar in twips (about 1444 twips to the inch) were filtered$^2$ in order to include only values that would be realistic for screen design. The effect of screen layout was significant, $F(4, 288) = 8.364$, $P < 0.001$, $\text{MS}_{\text{screen ratio}} = 1244806.025$ (see Table 4). The effect of order, $F(2, 72) = 1.693$, $P < 0.05$, $\text{MS}_{\text{order}} = 3230602.178$, and the interaction effect, $F < 1$, were not significant.

Multiple comparison tests with Sidak correction showed significant differences of ratio 18:82 with ratios 28:72 ($P < 0.05$), 33:67 ($P < 0.001$) and 38:62 ($P = 0.001$) and between ratios 23:77 and 33:67 ($P < 0.05$). $T$ tests with Bonferroni correction revealed that the difference between the preferred width and the Golden Section (4584 twips) was significant for each of the five screen ratios, all $P < 0.005$. Correlations between preferred widths for different screen ratios were significant, all $P < 0.001$ (see Table 4).

In the following analyses, the difference between the preferred width of the menu bar and the width of the menu bar for the ratio that was presented as a stimulus at each trial in the adjustment task was used as an outcome measure. The effect of screen layout was significant, $F(4, 288) = 324.006$, $P < 0.001$, $\text{MS}_{\text{screen ratio}} = 48221441.0222$ (see Table 5). Trend tests demonstrated a significant linear trend in the effect of screen ratios, $F(1, 72) = 1194.750$, $P < 0.001$, $\text{MS}_{\text{linear}} = 192081201.191$, but no other significant trends, all $P < 0.05$. $T$ tests with Bonferroni correction showed that the difference between the preferred width and the initially presented ratio was not significant for ratio 28:72, $t < 1$, but was significant for the other ratios, all $P < 0.005$.

#### 3.2.3. Associations between performance and subjective measures

A significant correlation was found between display quality and preferred screen width for ratio 32:68, $r = 0.303$, $P < 0.05$, both not for the other ratios. No

---

$^2$ Values ≥ 2184 twips (corresponding with the 18:82 ratio used in the experiment) and ≤ 12000 twips (corresponding with a 50:50 ratio) were included.
consistent pattern of correlations was found between task performance and first, perceived display quality and second, preferred screen width.

4. Discussion

In summary, our results demonstrate an effect of screen ratio on task performance (both speed and efficiency) and subjective measures (display quality and preferred screen width). In particular, the GS (ratio 32:68) produced inferior results compared to the average of the other ratios in terms of speed, accuracy and display quality. Best ratios were 23:77 for speed, 23:77, 28:72 and 33:67 for efficiency and 28:72 for display quality. In the adjustment task, a linear relationship was found between screen ratio and deviation of preferred width from screen ratio, where only ratio 28:72 did not differ from preferred width.

No effect of order was found. However, in the adjustment task the effect of preferred screen width should arguably be interpreted as an effect of context because participants made an aesthetic judgement about screen width was made when being presented with a particular screen ratio. In this interpretation, an effect of context occurred in the adjustment task.

Correlations were found between efficiency and speed for all ratios, between deviation scores for preferred width and display quality for ratio 32:68 only and between preferred widths for all screen ratios. There are two possible explanations for the better performance of the 28:72 ratio over others. First, task performance may have been improved for this ratio because it represented a balance between degradation in performance due to increased eye movements between the menu and content areas and a performance impairments due to visual clutter caused by close proximity of the menu and content areas. The latter would make it difficult to distinguish the bodies of text in the two areas.

The failure to find any advantage for basing screen layout on the GS failed to support Ngo and Ch’ng [15], who found greater user preference for screens reformatted according to principles of dynamic symmetry. However, participants in Ngo and Ch’ng’s study did not have to interact with the visual displays they were shown. Had they also had to conduct an information retrieval task as in the present study, the results they obtained may have been closer to our own. Moreover, the GS was used to format text boxes and pictures rather than to balance out menu and content areas. In conclusion, Ngo and Ch’ng’s findings may give support to Nielsen [25] argument that users do not necessarily know what is good for them.

Second, the preference for the 28:72 ratio may be explained in terms of McManus and Weatherby [22] theory of sensory aesthetics or in terms of ‘models’ of aesthetic behaviour [26]. McManus and Weatherby’s sensory aesthetics approach explains individual differences in ratio preferences in terms of preferences for particular, content areas. The former would make it difficult to distinguish the bodies of text in the two areas.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Descriptives for number of visited pages before correct answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order</td>
<td>Screen ratio</td>
</tr>
<tr>
<td>Narrow menus first</td>
<td>3.11 (0.88)</td>
</tr>
<tr>
<td>Equal distribution</td>
<td>3.28 (1.25)</td>
</tr>
<tr>
<td>Wide menus first</td>
<td>3.28 (0.97)</td>
</tr>
<tr>
<td>Overall</td>
<td>3.22 (1.04)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Descriptives for display quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order</td>
<td>Screen ratio</td>
</tr>
<tr>
<td>Narrow menus first</td>
<td>4.19 (1.39)</td>
</tr>
<tr>
<td>Equal distribution</td>
<td>4.00 (0.87)</td>
</tr>
<tr>
<td>Wide menus first</td>
<td>4.01 (0.88)</td>
</tr>
<tr>
<td>Overall</td>
<td>4.07 (1.09)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Correlations between preferred screen widths</th>
</tr>
</thead>
<tbody>
<tr>
<td>23:77</td>
<td>28:72</td>
</tr>
<tr>
<td>18:82</td>
<td>0.79</td>
</tr>
<tr>
<td>23:77</td>
<td>0.65</td>
</tr>
<tr>
<td>28:72</td>
<td>0.65</td>
</tr>
<tr>
<td>33:67</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Note: preferred screen widths when presented with screen ratio as stimulus.
The results of this study have implications for the design of Web pages. Our results indicate that the GS should not be used to design web pages; participants not only performed worse on objective measures of performance, but were also less likely to select the GS as their preferred (subjective) ratio. For optimal subjective outcomes, a ratio of 28:72 for the navigation frame and content area of Web pages should be used, but for optimal task performance 23:67 should be employed. It is possible to make calculations about the extent of reduction in task performance and aesthetic value by comparing screen ratios in terms of means presented in Tables 2–5. From these calculations it follows using screen ratio 28:72, in order to maximise subjective outcomes, would result in a 15% reduction in speed of task performance and 10% reduction of efficiency compared to 23:77. On the other hand, choosing screen ratio 23:67, in order to maximise subjective performance, would lead to in a 5% reduction in display quality and 4% reduction in preference compared to 28:72. Both ratios have large advantages compared to the GS (38:62). First, as a result of using the GS speed would be reduced by 41% and efficiency by 27% compared to ratio 23:77. Second, a consequence of adopting the GS would be a 15% reduction in display quality and a 4% reduction in preference compared ratio 28:72. Based on these considerations and previously reported statistical results, the 23:77 ratio is recommended over the other ratios. However, the GS is singled out as the worst design option both in terms of task performance and subjective outcomes.

The range of subjective measures to be used to measure aesthetic quality may be extended for more comprehensive assessment. The subjective measures used in the current proportion of investigations failed to establish such a relationship.

When considering the association between usability and subjective measures, an important issue is the distinction between ‘actual’ or objective, usability (typically measured by task performance) and ‘perceived’ usability (usually measured by psychometric scales or interviews). Previous research investigating associations between aesthetic judgement and usability has focused on perceived usability [16,26,32]. Unfortunately, these studies do not tell us anything about actual usability.

The effect of context in the adjustment task confirms the findings of Benjafie and Mcfarlane [21]. More specifically, correlations between preferred widths when presented with different screen ratios, together with the deviations for preferred widths from presented screen ratios, indicate that participants were consistent in their ratings, but still influenced in the direction of the presented ratio. Participants were most consistent in their judgement about the GS, the least preferred ratio: only for this ratio was the correlation between display quality and preference significant.

The lack of consistency between performance and preference that we obtained confirms the results of a meta-analysis performed by Neilsen and Levy [31]. They found that although a positive correlation between performance and preference existed in most studies, a substantial

Table 5
Descriptives for preferred screen width

<table>
<thead>
<tr>
<th>Order</th>
<th>Screen ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18:82; M, Mdiff&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Narrow menus first (n = 31)</td>
<td>3055, 871 (600)</td>
</tr>
<tr>
<td>Equal distribution (n = 27)</td>
<td>3105, 921 (533)</td>
</tr>
<tr>
<td>Wide menus first (n = 17)</td>
<td>3525, 1341 (799)</td>
</tr>
<tr>
<td>Overall</td>
<td>3179, 995 (648)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Difference between mean preferred width of the menu bar and width of the menu bar for the ratio that was presented as a stimulus.
study are based on previous research in Human–Computer Interaction (DES-W, based on [24]) and experimental aesthetics ([21]). However, a broader range of subjective measures may be considered, such as [16] aesthetics scales. More fundamentally, the vocabulary used to describe and measures to assess aesthetic experiences need further elucidation ([26]), including the relation between aesthetics and usability. Based on their results, Lindgaard and Dudek claim that perceived usability is an integral part of and independent of affective components of user satisfaction.

A major question raised by Lindgaard and Dudek [26] is to what extent aesthetic preferences are influenced by the ‘mere exposure’ effect, where users’ judgements are determined by immediate impressions, whether positive or negative. Similarly, Schenkman and Jonsson [3] assumed that first impressions are an important determinant of continued use of a site. This is an assumption that requires testing. The findings presented by Tractinsky et al. [16] indicate that at least under certain conditions an initial mere exposure effect is likely to be responsible for perceived usability after use of an interactive computer system. However, this effect may depend on various factors such as type of Web site or other type of computer application, users’ goals and time available to perform a task.

In order to make further progress, future research on aesthetic behaviour should be guided by models of human–computer interaction and aesthetics. Potentially useful vehicles include models of aesthetic behaviour [26], product appeal [32] and technology acceptance (TAM; [33]). Concerning the models of aesthetic behaviour, although the categorical model was found to be useful in the current study, the bipolar categorical-motivational model applies more generally to aesthetic preference and the balance between the effect of categorical and motivational drivers is likely to depend on the nature of the stimuli that are judged. According to Hassenzahl’s model, product appeal is influenced independently by both ergonomic quality and hedonic quality. Unfortunately, his test of the model suffered from various defects in experimental design and, notably, statistical analysis. Attractions of TAM are that it incorporates a wider scope of factors in trying to predict users’ behaviour and its demonstrated robustness in a wealth of empirical studies over the past 15 years. Aesthetic factors could be included in TAM, making it possible to assess the strength of aesthetic factors relative to other factors, such as perceived usefulness, in influencing behaviour. A disadvantage of the models, with the exception of more recent versions of TAM, is that they do not include actual usability, despite being a vital aspect of human–computer interaction.

More generally, there is a need to improve the quality of empirical studies of aesthetics in human–computer interaction. For example, three recent studies [16,26,32] in this area all suffered from serious flaws in experimental design or data analysis, including insufficient sample sizes for the statistical analyses conducted and using a dependent variable as an independent variable. We agree completely and emphatically with [34] that, in order to make valid inferences and produce sound design advice, the application of sound methodology is essential. This should not be seen as a requirement to use complicated statistical techniques; rather that research methods should be selected and used correctly that are appropriate to the research questions.

Two conclusions can be drawn from this study, one theoretical, the other practical. First, the current research failed to support the idea that the GS is an important factor in aesthetic judgements. In fact, of all the ratios presented to participants we found that the GS was least preferred. This further strengthens the position put forward by Benjafiel and McFarlane [21], that preferences are most likely to be related to prior experience and is consistent with the notion of prototypical shapes [22] and categorical [28] and, more generally, categorical motivational [29] models of aesthetic behaviour. This is an issue recognised even by Fechner [20], who wrote ‘that most of his subjects declared that the most pleasing proportion depends on its purpose’ ([35], p. 236). Second, our results give clear guidelines to designers for the layout of Web pages. The present findings indicate that the best ratio is 28:72 for aesthetic value and 23:77 for performance. Given the lesser impact on aesthetic measures of 23:77, we recommend that this ratio be used. One of the problems for the future is determining what further factors makes a computerised system aesthetically pleasing; potential factors can be found in the work of Ramachandran [36] ‘eight laws of artistic experience’. This is the direction in which researchers now need to proceed.

References


[23] E. Nygren, From paper to computer screen: Human information-processing and user interface design, Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology, Almqvist and Wiksell International, Stockholm, 188, 1996.


