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

Paul van Schaik *, Jonathan Ling

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

Available online 2 April 2004

Abstract

Although many web pages consist of blocks of text surrounded by graphics, there is a lack of valid empirical research to aid the design of this type of page [D. Diaper, P. Waelend, Interact. Comput. 13 (2000) 163]. In particular little is known about the influence of animations on interaction with web pages. Proportion, in particular the Golden Section, is known to be a key determinant of aesthetic quality of objects and aesthetics have recently been identified as a powerful factor in the quality of human–computer interaction [N. Tractinsky, A.S. Katz, D. Ikar, Interact. Comput. 13 (2000) 127]. The current study aimed to establish the relative strength of the effects of graphical display and screen ratio of content and navigation areas in web pages, using an information retrieval task and a split-plot experimental research design. Results demonstrated the effect of screen ratio, but a lack of an effect of graphical display on task performance and two subjective outcome measures. However, there was an effect of graphical display on perceived distraction, with animated display leading to more distraction than static display, $t(64) = 2.33$. Results are discussed in terms of processes of perception and attention and recommendations for web page design are given.

Keywords: World Wide Web; Graphics; Animation; Aesthetics; Screen ratio; Information retrieval; Usability

1. Introduction

Over the last decade, the Internet has become increasingly important as a market place (Haig, 2002). The greater commercial profile of the web has led to a tension between the
needs of web sites that strive for clarity and interest, and those of advertisers whose primary aim is to attract attention away from the contents of the sites to their products (Flanders, 1998).

Visual perception is particularly sensitive to the detection of motion (Gregory, 1998). This is a sensitivity that has been used to good effect in web pages by designers and advertisers in order to draw the attention of users to particular areas of the screen. The influence of animations on performance needs to be established in order to find a balance between usability and advertising. Motion in web pages can be presented in a variety of forms, such as flashing or rotating text, animations and pop-up boxes. As conspicuous and common as such motion may be on web pages, little research has investigated the impact of visually intrusive information on usability. A focus on the usability issues related to the use of animation in web sites is crucial: when faced with a poorly usable web site, users are likely to abandon these sites in favour of more usable ones. 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 (Nielsen, 2001).

When the intruding information is meant to attract users’ attention, like pop-up advertising windows, performance will suffer (Zhang, 1999). However, less intentional intrusions that form part of web pages, such as flashing logos or revolving titles, may also impact on interaction. This latter area requires study – content providers may have little control over the formats chosen by advertisers, but they are able to make design decisions about elements present in their own pages.

Two studies have examined the effect of animation on task performance in simulated web pages. Although both of these studies suffer in terms of the realism of the tasks – in neither case were participants asked to navigate around a web site – they still offer insight into the impact of motion on user behaviour. Zhang (1999) examined the degree of interference in task performance when users are presented with animations. Participants were asked to count the number of times a random sequence of letters appeared on a web page which contained animations in either bright or dull colours which were either related to the task (strings of letters) or unrelated (pictures). Zhang found that when the task was difficult, animation did influence performance, with participants not exposed to animations having higher levels of performance than those who were. In addition, the use of dull colours and animations less related to the task led to lesser effects on performance. Unfortunately web design implications drawn by Zhang must be treated cautiously as the task presented was not a realistic one, performance was assessed using an unconventional measure (time and error data, which are normally analysed separately, were combined) and the reporting of statistical analyses was unclear. Nonetheless, Zhang’s results are useful in that they appear to indicate that animations do impact upon information seeking ability. This research was extended by Diaper and Waelend (2000).

In their paper, Diaper and Waelend (2000) also investigated the influence of animations on information extraction, however they used more realistic screens that contained blocks of text and animations similar to those commonly found on the Internet. Participants had to find answers to questions that were presented on a series of screens. In contrast to Zhang, Diaper and Waelend found no significant evidence of an effect of redundant graphical images on ability to extract information from text presented on web pages. The conclusions and recommendations made by these researchers should be treated with some
caution because of methodological issues relating to their study. The most significant flaw is in the confounding of experimental condition and page content, so that each page appeared in a different condition; this means that any differences observed could be due to the experimental manipulations, the content of particular pages, or a combination of the two. Furthermore, there were only two examples of each type of web page presented to participants – most likely too few to exclude random effects, especially given the small number of participants (12) tested in the study.

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 (see, e.g., Tullis, 1997). Objective measures are commonly used to measure the usability of web sites. However, attention has shifted towards the use of subjective measures to explore user satisfaction (Ahuja & Webster, 2001). Subjective or preference data can be helpful in developing web sites that are visually appealing. At the same time, this type of data can be used in conjunction with objective data to produce sites that are more successful – whether ‘success’ means an improvement in user satisfaction, higher levels of task performance or a combination of these. The aesthetic perceptions of users can be investigated using subjective measures.

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 (Ngo, 2001; Ngo & Ch’ng, 2001; Tractinsky, Katz, & Ikar, 2000) and web pages (Schenkman & Jonsson, 2000). Subjective perceptions of a web site may be closely linked to the length of time users will spend on a site, as well as the likelihood of them returning to a site in the future. In fact, Kurosu and Kashimara (1995) 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 focusing only upon usability and develop what Nielsen (1996) 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, Schenkman and Jonsson (2000) asked participants to rate a series of 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, Schenkman and Jonsson argue, should be taken into account when web pages are designed. Unfortunately participants in their study were not required to interact with the pages they saw, but 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 research investigating the aesthetic appeal of user interfaces has been the Golden Section (GS). The Golden Section or Golden Ratio has been defined since Antiquity as the most pleasing way to divide a line (Green, 1995). 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). Fechner (1872) assumed that individuals have a preference for rectangles that have a ratio of 1:1.618 for their width to height. There have been several attempts at replicating Fechner’s findings. 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.
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 (2001) reformatted six screens according to principles of dynamic symmetry, which involved altering the proportions of text boxes and pictures using ratios, some of which were 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. If the GS is 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 first whether graphic display, in particular the effect of irrelevant screen material, has an impact on human–computer interaction in web-based systems. It remains unclear whether task performance is enhanced if subjective or aesthetic improvements are made to web pages. Therefore, the second aim was to study the effect of aesthetics in terms of the ratio of the width of screen areas. The research used a series of realistic web sites that involved navigating across pages to find information, and as such this represents a significant step forward in the analysis of online user behaviour.

2. Method

2.1. Experimental design

A 3 × 5 split-plot experimental design was used with the independent variables of graphical display, used with independent measures, and screen ratio, used with repeated measures. Three types of graphical display were used: none, static logos and dynamic (animated) logos. 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; see Fig. 1). 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 GS. (Ratios are presented using rounded numbers below.) All 15 combinations of graphical display and screen ratio were presented in the experiment. Outcome measures included speed, accuracy, efficiency and subjective responses.

2.2. Participants

There were 98 undergraduate participants, consisting of 86 females and 12 males. Of these, 32 took part in the no logo condition, 34 in the static logo condition and 32 in the animated logo condition. Participants’ mean age was 24 years with a standard deviation of 8.4.

All participants had some experience of using the web. Participants had been using the web from 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, many participants (47%) used the web at least once a day, with a further 43% of participants using the web more than once a week but less than once a day.
2.3. Materials and apparatus

For the experiment a set of six web sites was produced. All sites consisted of a hierarchy of three levels. Each site was produced in fifteen versions, corresponding to the
five screen ratios combined with three levels of graphical display. For each web site a static logo was an image consisting of the first letter (font size 96 points) of the name of the site and was presented in black (colour code #FFFFFF) on a blue background.
The animated logo for each site was made up of two images that alternated every second. The first image was the same as the static logo and in the second image the colour of letter and background was reversed. Logos did not help with answering information retrieval questions and this would have been obvious to participants after the practice task. The first site used for the practice experiment consisted of 16 pages 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 total amount of empty space on any given page was kept constant over menu and content areas combined, regardless of the ratio used.

The experiment ran on personal computers (Intel Pentium, 333 MHz, 64 Mb RAM, Microsoft NT4 operating system, 14 in. monitors). The screen dimensions were \(800 \times 600\) pixels and the monitors were set to optimum levels of contrast and brightness. 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 visited 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. Each screen ratio was presented eight times.
Two sets of rating questions were also presented to participants. In a first rating task, participants completed a set of 30 (6 items × 5 screen ratios) questions (DES-W), based on the DES (Display Evaluation Scale; Spenkelink, Besuijen, & Brok, 1993). 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 7-point scale, where 1 was bad and 7 was good. The DES-W was found to be reliable, Cronbach’s α = 0.88, 0.80, 0.85, 0.84 and 0.89 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.

The second rating task asked participants to give their assessment of the logos in relation to their task performance using three items adapted from Diaper and Waelend (2000); (‘When I was performing the tasks the logos diverted my attention from the text.’, ‘When performing the tasks it was easy to ignore the logos.’, ‘I would have preferred to have no logo whilst performing the tasks.’). All questions used a 7-point scale, ranging from 1 (strongly agree) to 7 (strongly disagree). The first two items of perception of logos were found to form a reliable logo distraction scale, Cronbach’s α = 0.75. An overall (average) distraction score was calculated (the higher, the less perceived distraction) and used in subsequent analysis. As in Diaper and Waelend (2000), students were also asked if they believed they paid attention to the logos.

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 and were told to take the most direct route possible to their answers. Once they found the answer they had to click a button labelled ‘Your answer’. A dialog box then appeared into which participants entered their answer. When they had entered their answer, they were shown the next question.

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 (5 sites × 8 questions).

After the experimental trials, participants answered rating questions (DES-W) before answering the distraction questions (the latter were only administered in the static and animated logo conditions). Next, participants indicated their preference for the relative width of the navigation and content areas for each of the five screen ratios. For this purpose, they used a mouse to control 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 on trials in which a correct answer was given was examined in terms of speed (i.e., time-on-task) and efficiency (i.e., number of pages loaded). Time-on-task for incorrect answers was not analysed because very few of these occurred. Subjective measures included display quality, preferred screen ratio and perception of logos. Analysis of variance (ANOVA) and simple effect tests were conducted to analyse the effects of screen ratio and graphical display on outcome measures. Post-hoc analyses were conducted to identify differences between screen ratios. T tests and \( \chi^2 \) tests were conducted to analyse participants’ perceptions of logos.

3.1. Analysis of 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 effect of screen ratio was significant, \( F(4,380) = 17.301, p < 0.001, \text{MS}_{\text{screen ratio}} = 1.398 \). The effect of graphical display, \( F < 1 \), and the interaction effect of screen ratio and graphical display, \( F(8,380) = 1.621, p > 0.05, \text{MS}_{\text{screen ratio}} = 0.131 \), were not significant (see Table 1). Multiple comparison tests with Sidak correction showed significant differences: of ratio 28:72 with ratios 18:82 (\( p < 0.01 \)), 23:77 (\( p < 0.001 \)) and 33:67 (\( p < 0.05 \)) and of ratio 38:62 with ratios 18:82, 23:77 and 33:67 (all \( p < 0.001 \)). Testing a deviation contrast, ratio 38:62 was significantly different from the average of the other ratios, \( t(97) = 7.074, p < 0.001 \).

3.1.2. Efficiency

The number of pages visited before a correct answer was given was calculated. The effect of screen ratio was significant, \( F(4,380) = 6.302, p < 0.001, \text{MS}_{\text{screen ratio}} = 4.448 \) (Table 2). The effects of graphical display 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 and 33:67 (both \( p = 0.001 \)), and between ratios 23:77 and 28:72 (\( p = 0.01 \)). The test of a deviation contrast showed that there was a significant difference between ratio 38:62 and the average of the other ratios, \( t(97) = 3.238, p < 0.005 \).

Table 1

<table>
<thead>
<tr>
<th>Graphic display</th>
<th>Screen ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>No logo</td>
<td>12.42 (4.58)</td>
</tr>
<tr>
<td>Static logo</td>
<td>12.47 (4.92)</td>
</tr>
<tr>
<td>Animated logo</td>
<td>13.87 (8.16)</td>
</tr>
<tr>
<td>Overall</td>
<td>12.91 (6.06)</td>
</tr>
</tbody>
</table>

Note: Descriptives for original (non-transformed) data are reported for ease of interpretation.
3.2. Analysis of subjective measures

3.2.1. Display quality

The effect of screen ratio on overall DES-S scores was significant, $F(4, 380) = 38.673$, $p < 0.001$, $MS_{\text{screen ratio}} = 27.415$ (see Table 3). The effect of graphical display and the interaction effect were not significant, both $F < 1$. Multiple comparison tests with Sidak correction revealed a significant difference in each pair of ratios; comparisons were significant with $p < 0.001$, except between ratios 18:82 and 33:67 ($p = 0.005$) and between ratios 23:77 and 28:72 ($p = 0.001$). When testing a deviation contrast, ratio 38:62 was found to be significantly different from the average of the other ratios, $t(97) = 9.024$, $p < 0.001$.

3.2.2. Preferred screen width

Preferred widths of the menu bar in twips (about 1444 twips to the inch) were filtered in order to include only values that would be realistic for screen design. The effect of screen ratio was significant, $F(4, 264) = 8.196$, $p < 0.001$, $MS_{\text{screen ratio}} = 2601406.531$ (see Table 3). Neither the effect of graphical display, $F(2, 66) = 1.547$, $p > 0.05$, $MS_{\text{graphical display}} = 2601406.531$ nor the interaction effect, $F < 1$, were significant. Multiple comparison tests with Sidak correction showed significant differences of ratio 18:82 with ratios 33:67 ($p < 0.01$) and 38:62 ($p = 0.001$) and between ratios 23:77 and 38:62 ($p < 0.01$). T tests with Bonferroni correction revealed that the difference between the preferred width and the Golden Section (4584 twips for the menu area, 7416 twips for the content area) was significant for each of the five screen ratios, all $p < 0.005$.

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 ratio was significant, $F(4, 264) = 322.739$, $p < 0.001$, $MS_{\text{screen ratio}} = 46571826.685$ (see Table 3). Trend tests demonstrated a significant linear trend in the effect of screen ratios, $F(1, 66) = 900.06$, $p < 0.001$, $MS_{\text{linear}} = 186017015.460$, 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 $p < 0.005$.

---

1 Values \( \geq 2184 \) twips (corresponding with the 18:82 ratio used in the experiment) and \( \leq 12000 \) twips (corresponding with a 50:50 ratio) were included. After filtering there were 23 cases in the no logo condition, 22 in the static logo condition and 24 in the animated logo condition.

2 Results for the effect of context and the interaction effect are necessarily identical to those presented above.
3.2.3. Perception of graphical display

The difference between distraction score and a neutral answer (a score of 4 on a seven-point Likert scale) was significant for both the static logo condition, \( t(33) = 7.998, p < 0.001 \), and the animated logo condition, \( t(31) = 2.531, p < 0.05 \).

The difference in perceived distraction between the two conditions was significant, \( t(64) = 2.33, p < 0.05 \), mean = 5.765 (SD = 1.286) and mean = 4.844 (SD = 1.886) for the static and the animated logo conditions respectively.

A large majority (71%) of participants believed they did not pay attention to the logos, \( \chi^2 = 11.879, p < 0.001 \). No difference existed between the two logo conditions in terms of participants' view of whether they paid attention to the logos, \( \chi^2(1) = 0.184, p > 0.05 \).

### 4. Discussion

There was no effect of graphical display on speed, efficiency, display quality, preference for screen ratio and perception of paying attention to logos. Furthermore, for both types of graphical display, participants' perceived level of distraction was different from a neutral answer, indicating that they did not feel distracted by logos. However, the difference was substantially greater for the static logo condition (1.4 SD difference from neutral answer) than for the animated condition (0.4 SD difference). In terms of Cohen's (1988) conventions for effect size, the first difference represents a very large effect (0.8 SD = large)
and the second difference represents less than a low to moderate effect size (0.5 SD = moderate). Moreover, when comparing the two conditions directly, static logos resulted in less perceived distraction than animated logos (0.6 SD difference) which was above a moderate effect size. These results indicate that although task performance and two subjective measures related to screen ratio (display quality and preference for screen ratio) were not affected by graphical display, a subjective measure directly related to graphical display (perceived distraction) did reveal a higher level of perceived distraction with animated rather than static logos.

Investigating the effects of graphics in web pages on task performance and perceived complexity of pages, Diaper and Waelend (2000) set out to address the ‘failure’ they perceived ‘in the application of theoretically oriented and applied psychological laboratory research to the real world of web page usage’ (p. 166). At first sight, these findings seem to confirm Diaper and Waelend’s findings that graphics do not have a negative impact on the quality of human–computer interaction in web pages. However, our results indicate that users do perceive animated graphics as more distracting than static graphics. The current study used graphics sparingly, by presenting a logo at the lower left of the screen, and more extensive use of graphics would naturally exacerbate their negative effect.

Task performance may have been affected by perceptual or attentional issues (Ware, 2000). In terms of perceptual distraction, participants may have learned during the practice task that logos always appeared in the same position (bottom left) on web pages that were not relevant to their retrieving information. Participants may therefore have intentionally directed their vision away from the potential distraction of logos that were presented in this location. This may have become more difficult if, for example, logo size was increased; however, in the present study the physical characteristics of the logos were not varied. Although stimuli may not always be processed consciously, negative priming of stimuli, when the processing of a target stimulus is impaired if a conflicting stimulus was presented on the previous trial, may occur (Driver & Tipper, 1989). However, in the experiment, as the logos did not help in retrieving information, negative priming would not occur.

Irrespective of their interpretation, more confidence can be placed in our findings than in those of previous studies (Diaper & Waelend, 2000; Zhang, 1999) as the current study avoided shortcomings in research design and analysis that reduces the value of these studies. Both studies employed a small number of stimuli per experimental condition and thereby potentially reduced reliability of results. Zhang used an unrealistic type of search task with arrays of meaningless letter strings and blocks of random words. Diaper and Waelend employed what was essentially a visual search task, using a modification of existing web pages consisting of text, but the task did not involve interaction with the pages. The interactive search task used in the current study is arguably more ecologically valid because it did involve interaction. Zhang’s accuracy measure is extremely suspect and her statistical analysis is problematic. Statistical power was inadequately low in Diaper and Waelend’s research (the authors admit this for one of their outcome measures) and unknown in Zhang’s study. Diaper and Waelend failed to analyse and interpret a statistical interaction effect, which reduces the validity of their conclusions, nor did they provide evidence of the psychometric properties of their subjective outcome measure (immediately perceived complexity).

Diaper and Waelend (2000, p. 179) conclude that first, designers should trust their intuition and second, empirical tests and the guidelines derived from these are neither
necessary nor relevant to web design. The strong statements made by Diaper and Waelend require careful scrutiny. Although Diaper and Waelend base their suggestions on empirical work, many recommendations made by human–computer interaction commentators are not (see Tullis, 1997; Venkatraman, 2000). In other words, although Diaper and Waelend use an empirical approach to study web page design, they encourage designers to ignore this and trust their instincts. Such a recommendation is surprising, given that web pages designed according to cognitive principles have been proved to be superior to those that are not (Dalal, Quible, & Wyatt, 2000). In conclusion, it is important to avoid sending incorrect and arguably dangerous messages to designers who require advice on designing their web pages; more specifically, it is undesirable to tell designers that they can use their own judgement when making decisions about choices of web design parameters. Diaper and Waelend draw conclusions that go far beyond the results that they obtained and thereby arguably commit the failure that they set out to address, with the risk of damaging the reputation of human–computer interaction research and credibility of this research with designers.

Our findings raise the issue of the sensitivity of outcome measures of human–computer interaction. Task performance may be less sensitive to experimental manipulations because participants will put more effort in when tasks become more difficult, and as a consequence, task performance will remain fairly constant over a range of experimental manipulations (Spenkelink & Besuijen, 1999). At the same time, specific subjective measures directly related to the design parameter that is investigated (in the current study: perceived distraction related to graphical display) are likely to be more sensitive than other subjective measures of the quality of human–computer interaction (display quality and preference for screen ratio herein) because users will be able to report their perceptions of task difficulty. Even if specific subjective measures are more sensitive, this does not in itself make them relevant as outcome measures. As argued previously, their relevance may become apparent when analysed together with measures of technology acceptance such as perceived ease of use (Davis, 1993), which will help determine the uptake and ultimately the success of web sites.

The results of the current study are based on logos, representing the web site from which participants retrieved information, which can serve as an orientation aid or context for users when using the web. This use of graphics is similar to that of Diaper and Waelend (2000), who used graphics that were ‘loosely semantically related’ to the textual content of their web pages (p. 169), where graphics provide a context for and may reinforce the content of pages. Our study did not use graphics representing web site functionality, such as an icon of a shopping cart representing the action of purchasing as well as providing a context for the content of a web page. Therefore, the conclusions from the current research should only be applied to the use of graphics to provide a context rather than functionality in web page designs.

Our results demonstrate an effect of screen ratio on task performance (both speed and accuracy) and subjective measures (display quality and preferred screen width). In particular, the Golden Section (GS; ratio 38:62) produced inferior results compared to the average of the other ratios in terms of speed, accuracy and display quality. Based on the results for speed and efficiency and for display quality together, 23:77 was the best ratio, with 28:72 equally good in terms of 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.
In establishing that the GS does not result in better outcomes in the use of web pages, these results confirm the findings of van Schaik and Ling (in press) and can be explained in terms of the categorical model of aesthetic behaviour (Lindgaard & Dudek, 2003). This model combines categorical and motivational drivers of aesthetic preference. The model fits our data well as the two screen ratios of 23:77 and 28:72 appear to match users’ prototype of a web page in terms of layout. Evidence for the operation of this model was found in typical screen ratios found in Web sites; we conducted a survey of 10 typical sites and found that mean screen ratio was 23:77 (SD = 4 for menu width). This mean ratio is far removed from the GS.

The findings of the current research have led to the following three design recommendations. First, static graphics can be used sparingly to provide context for content in web pages, as they do not appear to affect task performance and are not perceived as distractions. Second, the use of animated graphics is not recommended, as users will experience increased distraction by dynamic graphics compared to static ones; however, there is no evidence of an effect of sparingly used animated graphics on task performance. Third, use of the GS is not recommended in web page designs, due to its negative effect on task performance as well as on subjective outcomes. Instead, screen ratios in the range of 23:77 are recommended.

Considering the possible impact of perceived distraction on website acceptance, future research should investigate this distraction in relation to technology acceptance factors such as perceived ease of use, perceived usefulness and intention to use, in order to establish the potential influence of graphics on the acceptance of web sites. Another unanswered question remains the effect of physical characteristics of graphics, such as their size, on both task performance and subjective measures. Furthermore, comparison of the effect of graphics that represent web page functionality to those graphics that provide context of web page content or those that are gratuitous requires investigation.

The current study investigated the effects of two design parameters simultaneously, graphical display with screen ratio, where the latter previously had been shown to have a significant effect on the quality of interaction with web pages (van Schaik & Ling, in press). This made it possible to establish the relative effect of these two design parameters. Based on task performance measures and two subjective measures (display quality and preference for screen ratio) both the effect of screen ratio and the lack of an effect of graphical display (Diaper & Waelend, 2000) were confirmed. However, a specific subjective measure directly related to graphical display (perceived distraction) indicated that animation in web pages resulted in increased distraction. It is likely that distraction would be exacerbated with a greater use of animation and the acceptance of this type of web page reduced.

References


