<table>
<thead>
<tr>
<th>Author:</th>
<th>Haig, A. &amp; Whitfield, T. W. A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title:</td>
<td>Predicting the aesthetic performance of web sites: what attracts people?</td>
</tr>
<tr>
<td>Conference name:</td>
<td>1st International Symposium on Smart Graphics 2001</td>
</tr>
<tr>
<td>Conference location:</td>
<td>Hawthorne, New York, United States</td>
</tr>
<tr>
<td>Conference dates:</td>
<td>21-23 March 2001</td>
</tr>
<tr>
<td>Place published:</td>
<td>New York</td>
</tr>
<tr>
<td>Publisher:</td>
<td>ACM</td>
</tr>
<tr>
<td>Year:</td>
<td>2001</td>
</tr>
<tr>
<td>URL:</td>
<td><a href="http://www.acm.org/">http://www.acm.org/</a></td>
</tr>
</tbody>
</table>

Copyright: Copyright © ACM, 2001.

This is the author’s version of the work, posted here with the permission of the publisher for your personal use. No further distribution is permitted. You may also be able to access the published version from your library.
ABSTRACT
This paper focuses upon the visual aesthetic performance of web sites. It describes an experiment in which a site, designed at three controlled levels of visual ‘enrichment’, was evaluated on a number of measures by two subject groups. The measures used represent the two main facets of the Categorical-Motivation model of aesthetics, and included others specific to the performance of web sites. Using Multidimensional Scaling the results were consistent with predictions derived from this model. They indicate that the drivers of site evaluation were primarily exploratory, ‘collative’ variables representing ‘interest’, ‘novelty’ and ‘fun’.

General Terms

Keywords
Aesthetics, Design, Experimentation, Psychology.

1. INTRODUCTION
The visual performance of objects exerts a powerful influence upon our decisions to purchase or interact with them. This visual performance, traditionally defined in the context of ‘aesthetics’, permeates all cultures and the objects that they produce. So much so that it is difficult to identify a manufactured object that is without a designed visual aesthetic. This permeation stretches from the hand-made, unique object to the mass-produced, from commodities and clothing to media. Its most recent and challenging manifestation is in the new electronic Internet-based communications media.

The emergence of electronic communications for both products and services poses major problems of interface design, increasingly, with the proliferation of sites, the focus must move to visual performance as a means of product differentiation and brand identification, just as it has for other products and services. The usability of the site will soon become a given, just as the mechanical performance of a washing machine or hair-dryer is now taken for granted. It will become a nil product differentiator. ‘Style’ and ‘image’, those little understood features of visual performance, will emerge as primary forces.

Despite the sheer prevalence of aesthetic phenomena, no concerted effort has been made to understand aesthetics scientifically, model it, and therefore to predict its performance. Within psychology the dissociation of emotion from cognition by both Behaviourism and Cognitivism has not assisted in providing the theoretical resources for probing aesthetics. Similarly, the distinct fields that make up emotion research, complete with their own frames of reference and research paradigms, militate against coherence. After its illustrious empirical beginnings as the second published area of experimental psychology [3], aesthetics has suffered the same fate as research into emotion, namely, that it was ignored by behaviourists and subsequently by cognitivists. Perhaps more surprisingly, the recent emergence of emotion theories unaccountably neglect aesthetics [4], as do integrative theories that seek to couple emotion and cognition [5]. Aesthetics is even ignored in areas of applied scientific research where its impact would seem self-evident. Thus research in affective computing overlooks it [6], as largely does HCI research, despite the existence of an entire industry – the design industry – increasingly devoted to it. This oversight is inevitably due to the failure to identify aesthetics as either a category, and hence to position it within the cognitive research framework, an emotion, and hence to position it there, or both.

As yet, little attention has been given to the visual aesthetic performance of web-based material. The research described in this paper is an attempt to extend the research paradigm of experimental aesthetics into this domain. In the study reported here, a ‘commercial’ web site was designed at three controlled levels of ‘enrichment’, from full animation, sound and imagery to zero animation, sound and imagery (Figure 1). In order to provide a theoretical basis for the study, interest lay in the capacity of the recently formulated Categorical-Motivation model of aesthetics [8] to account for site preference. This model integrates two of the main - and opposing - empirically based theories of aesthetics to emerge from psychology, the Collative-Motivation model [1,2] and
the Categorical model [7,9]. The former equates aesthetic evaluation with arousal-based exploratory behaviour, and posits the key role of 'collative' variables (e.g. interest, novelty) in this process. The latter equates aesthetic evaluation with information processing demands, and posits the key role of categorical variables (e.g. typicality, representativeness) in this process. The theories make conflicting predictions: the former predicts preference as a function of moderate levels of processing demand (e.g. stimuli that are perceived as novel within a category), while the latter predicts preference as a function of low levels of processing demand (e.g. stimuli that are perceived as typical/prototypic within a category). The integrated theory, the Categorical-Motivation model, incorporates both sets of predictions, but allocates specific predictions according to the category of object cognitively accessed. Where a cognitive category is well formed and closed to further articulation (e.g. medieval cathedrals, Renaissance paintings to, say, a Western educated audience), preference would be a function of typicality and representativeness. Where a category is ill formed and open to further articulation, preference would be a function of novelty and interest. As a recent phenomenon, web sites would come into the latter category. ‘Collative’ variables therefore should account for their visual aesthetic performance. A number of measures were employed that represent both facets of the model, including others that appear relevant to site evaluation (e.g. ease of use).

A further interest lay in possible differences in evaluation of the sites between different subject groups. Previous research involving non-web material found differences between the evaluations of designers and non-designers [10]. The present study therefore used both a design-trained group and a computing-trained group, with the intention of expanding the sampling to other groups. At the time of submitting this paper only two subject groups in Australia have been included in the study. For this reason only a preliminary analysis is available. The results, however, appear of sufficient interest to warrant presentation.

2. METHOD
2.1 Subjects
These were 64 students from Swinburne University of Technology, Melbourne: 37 were undergraduate Multimedia design students (‘design group’) and 27 were students from an undergraduate Multimedia software engineering and networks course (‘computing group’). All subjects were from Years One and Two of their respective programs and took part in the study at the end of their academic year. Subjects were volunteers and recruited largely on the incentive of satisfying their curiosity. While attempts were made to recruit equal numbers of males/females per course, the numbers obtained largely reflect the unequal gender distribution of students in the two departments - Design Group M=25/F=12 and Computing Group M=14/F=13.

2.2 Stimulus material
This consisted of three web sites designed for a hypothetical mobile phone company. Site 1 used Flash-based animation extensively, plus sound and imagery. Site 2 used neither animation nor sound, but used imagery. Site 3 used neither animation, sound nor imagery. All three sites were designed with identical layout, text, typefaces and colours, the only variables being those listed above. The sites were viewed on Apple Macintosh iMac computers via the Microsoft Explorer 5.0 browser.

Figure 1: The three web sites

A whitfield@swin.edu.au
2.3 Procedure
Subjects were tested as members of their course group. Accordingly, five sessions were run. Subjects were provided with a seven page survey, the top page of which was for biographical data and to inform them of the broad purpose of the investigation. The rating system was also explained here, a ten point interval scale with disagree and agree at the poles, and a description was given of the type of evaluations required. Subjects were asked to indicate their level of agreement with a number of statements covering ‘collative’, ‘categorical’, and site related measures. Examples of these are: “This site is novel in its visual design”; “This site is fun to use”; “This site is appropriate for a mobile phone company”; “I have seen sites similar to this one”; “This site is easy to use”; “I like this site”. The three sites were evaluated by each subject using the same measures. In order to counter possible site order effects, two orders of presentation were used – first order and reverse. Subjects were allocated at random to either order. Also, to provide masking between the site presentations, two sets of pre-loaded sites of approximately five minutes duration were viewed between the target sites 1 and 2, and 2 and 3. The masking sites were unrelated in content to the target sites.

3. RESULTS AND DISCUSSION
For both subject groups, site preference increased with site enrichment (Figure 2). Thus, site 1, with full animation, sound and imagery, was preferred to site 2, with neither animation nor sound but imagery, while site 2 was preferred to site 3, which used neither animation, sound nor imagery. Figure 2 also reveals that in absolute terms liking was higher for all sites by the computing group.

An intriguing feature of the results is that while both subject groups show very similar patterns of response, the computing group evaluated each site higher than did the design group. This is not limited to isolated measures, but permeates almost all of the data set (Figures 3,4). This suggests that designers may underestimate the aesthetic performance of sites relative to non-designers: alternatively, non-designers may overestimate the aesthetic performance. Clearly, data from other subject groups representative of the public are required to clarify the meaning of this difference.

In order to gain an overview of the relationships amongst response measures, Multidimensional Scaling (MDS) was performed on each subject group’s data. MDS represents objects – or in this case response measures – as a set of points in multidimensional space. The locations of the objects are estimated from matrices of differences or similarities between all pairs of objects, and the distances between the points reflect the empirical relationships of the data. Essentially, the closer together two points are in the MDS space, the closer they are in their degree of similarity or association. Using ordinal measures in an Euclidean space restricted to two dimensions, both Kruskal’s Stress and the RSQ...
indicate strong fits between the data and their spatial representations (Kruskal’s Stress/RSQ: Design Group .08/.97: Computer Group .09/.97).

3.1 Design Group
Site preference is closely associated with a cluster of measures denoting ‘interesting…. ’ and ‘fun…. ’, and also includes ‘visually novel’ and ‘liking for the colour’ (Figure 5). Measures disassociated from site preference are those denoting ‘similarity to other sites’, ‘typical of phone sites’ and ‘ease of use’. As with the design group, these results can be interpreted as indicating ‘collative’ properties as the drivers of site preference. Where they differ from the design group is in the closer association of ‘appropriate for a mobile phone site’ and ‘novel to navigate’ with preference. ‘Visual complexity’ is sufficiently removed from the other measures as to question its meaning for the computing group.

4. CONCLUSIONS
The high level of consistency obtained in the results suggests order and coherence in the site evaluations. Furthermore, the correspondence between the Categorical-Motivation model predictions and the Multidimensional Scaling outcomes lends support to this model as a useful theoretical framework for investigating visual interfaces. The incorporation of further subject group data (in process) will enhance interpretation, as would the extension of the research to include actual usability measures.

5. REFERENCES