Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE DIRECT®

Information Sciences 152 (2003) 25–46

INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL

www.elsevier.com/locate/ins

Modelling interface aesthetics

David Chek Ling Ngo [a,*], Lian Seng Teo [a], John G. Byrne [b]

[a] *Faculty of Information Technology, Multimedia University, Cyberjaya 63100, Malaysia*
[b] *Department of Computer Science, Trinity College Dublin 2, Ireland*

Received 30 December 1999; received in revised form 20 June 2002; accepted 1 September 2002

Abstract

An important aspect of screen design is aesthetic evaluation of screen layouts. While it is conceivable to define a set of variables that characterize the key attributes of many alphanumeric display formats, such a task seems difficult for graphic displays because of their much greater complexity. This paper proposes a theoretical approach to capture the essence of artists' insights with 14 aesthetic measures for graphic displays. Our empirical study has suggested that these measures are important to prospective viewers and may help gain attention and build confidence in using computer system.

Keywords: Screen design; Interface aesthetics; Aesthetic measures; Aesthetic characteristics; Multi-screen interfaces

1. Introduction

The role of aesthetics in human affairs has been widely documented [1]. Certainly, it is related to our appreciation of computer systems as well. However, some [2,3] warn against a tendency among designers to emphasize

* Address for correspondence: 32 Jalan Tempua 5, Bander Puchong Jaya, Puchong, Selangor 47100, Malaysia. Tel.: +60-6-252-3485; fax: +60-6-231-8840.

E-mail addresses: dclngo@yahoo.com, david.ngo@mmu.edu.my (D.C.L. Ngo), lsteo@mmu.edu.my (L.S. Teo), john.byrne@cs.tcd.ie (J.G. Byrne).

0020-0255/02/$ - see front matter
doi:10.1016/S0020-0255(02)00404-8

the aesthetic elements of the user interface, because these might degrade usability. In fact, interface aesthetics play a greater role in affecting system usability and acceptability than we might be willing to admit. Careful application of aesthetic concepts can aid:

- *Acceptability*. Two recent studies [4,5] show that very high correlations were found between users' perceptions of interface aesthetics and usability.
- *Learnability*. Toh [6] found that aesthetically pleasing layouts have a definite effect on the student's motivation to learn. Aspillage [7] found that good graphic design and attractive displays help contribute to the transfer of information. Szabo and Kanuka [8] found that subjects who used the lesson with good design principles completed the lesson in less time and had a higher completion rate than those who used the lesson with poor design principles. Heines [9] found that a poorly designed computer screen can hinder communication. A study by Grabinger [10] indicated that organisation and visual interest are important criteria in judging the readability and studyability of the real screens. Screens that are plain, simple, unbalanced, and bare are perceived as undesirable.
- *Comprehensibility*. Tullis [11] found that redesigning a key display from a system for testing telephone lines resulted in a 40% reduction in the time required by the users to interpret the display. In a study of 500 displays, Tullis [12] found that the time it took users to extract information from displays of airline or lodging information was 128% longer for the worst format than for the best.
- *Productivity*. A convincing demonstration of design improvement has been reported by Keister and Gallaway [13]. Those authors describe a data entry application in which relatively simple improvements to user interface software—including selection and formatting of displayed data, consistency in wording and procedures, on-line user guidance, explicit error messages, re-entry rather than overtyping for data change, elimination of abbreviations, etc.—resulted in significantly improved system performance. Data entry was accomplished 25% faster, and with 25% fewer errors.

Although knowledge of the users' tasks and abilities is the key to designing effective screen displays, an objective, automatable metric of screen design is an essential aid. Tullis [14] developed four metrics for alphanumeric displays: overall density, local density, grouping, and layout complexity. Streveler and Wasserman [15] proposed an objective measure for assessing the spatial properties of alphanumeric screens. Sears' [16] developed a task-dependent metric called layout appropriateness to assess whether the spatial layout is in harmony with the users' tasks. Layout appropriateness is a widget-level metric that deals with buttons, boxes, and lists. In this paper, we attempt to synthesise the guidelines and empirical data related to the formatting of graphic displays into a well-defined framework. We develop fourteen aesthetic measures for graphic displays: balance, equilibrium, symmetry, sequence, cohesion, unity, proportion, simplicity, density, regularity, economy, homogeneity, rhythm, and order and complexity. (This paper is extended from [17–19]. This study differed from earlier studies in that the formulae were further refined to incorporate more relevant aesthetic components that might help explain our experience with, and evaluation of, computer aesthetics, and experiments were conducted using real screens copied from actual multimedia systems to provide stronger empirical support to the formulae.) The paper begins with an introduction to the aesthetic model, then carries out preliminary studies of the measures, and then summarises and reviews the contributions.

There are four basic ways to use windows [20].

- *Multi-window interfaces*. Examples are the IBM OS/2 operating system, most UNIX X-windows applications, and now Microsoft Windows 98.
- *Multi-document interfaces*. For example, Microsoft Word or Excel.
- *Multi-pane interfaces*. Microsoft Paint and Netscape Navigator are examples.
- *Multi-screen interfaces*. Mostly found in one-time GUIs and wizards, and multimedia applications.

This paper addresses primarily multi-screen interfaces. With some modification, some of the techniques presented can also be used for other screen types. Keep in mind that the following discussion does not focus on the words on the screen, but on the perception of structure created by such concepts as spacing.

2. Aesthetic measures

Many noteworthy texts discuss theories of design in both fine and commercial art. Arnheim [21] and Dondis [22] are good examples. From the literature on screen design, Galitz's book on design and layout, The Essential Guide to User Interface Design [23], presents an extensive list of very specific guidelines for the design of screens. This paper relies heavily on this and other similar works [24] to help demonstrate our approach. Observe that the range of the following measures is between 0 (worst) and 1 (best).

2.1. Measure of balance

Balance can be defined as the distribution of optical weight in a picture. Optical weight refers to the perception that some objects appear heavier than others. Larger objects are heavier, whereas small objects are lighter. Balance in screen design is achieved by providing an equal weight of screen elements, left and right, top and bottom.

Balance is computed as the difference between total weighting of components on each side of the horizontal and vertical axis and is given by

$$\mathrm{BM} = 1 - \frac{|\mathrm{BM}_{\mathrm{vertical}}| + |\mathrm{BM}_{\mathrm{horizontal}}|}{2} \in [0, 1] \tag{1}$$

$\mathrm{BM}_{\mathrm{vertical}}$ and $\mathrm{BM}_{\mathrm{horizontal}}$ are, respectively, the vertical and horizontal balances with

$$\mathrm{BM}_{\mathrm{vertical}} = \frac{w_{\mathrm{L}} - w_{\mathrm{R}}}{\max(|w_{\mathrm{L}}|, |w_{\mathrm{R}}|)} \tag{2}$$

$$\mathrm{BM}_{\mathrm{horizontal}} = \frac{w_{\mathrm{T}} - w_{\mathrm{B}}}{\max(|w_{\mathrm{T}}|, |w_{\mathrm{B}}|)} \tag{3}$$

with

$$w_j = \sum_{i}^{n_j} a_{ij} d_{ij} \quad j = \mathrm{L}, \mathrm{R}, \mathrm{T}, \mathrm{B} \tag{4}$$

where L, R, T, and B stand for left, right, top, and bottom, respectively; a_{ij} is the area of object i on side j; d_{ij} is the distance between the central lines of the object and the frame; and n_j is the total number of objects on the side.

2.2. Measure of equilibrium

Equilibrium is a stabilisation, a midway centre of suspension. Equilibrium on a screen is accomplished through centring the layout itself. The centre of the layout coincides with that of the frame. (There are minor deviations from this definition, which we discuss in Section 4.)

Equilibrium is computed as the difference between the centre of mass of the displayed elements and the physical centre of the screen and is given by

$$\mathrm{EM} = 1 - \frac{|\mathrm{EM}_x| + |\mathrm{EM}_y|}{2} \in [0, 1] \tag{5}$$

The equilibrium components along the x- (EM_x) and y-axis (EM_y) are given by

$$\mathrm{EM}_x = \frac{2\sum_i^n a_i(x_i - x_c)}{n b_{\mathrm{frame}} \sum_i^n a_i} \tag{6}$$

$$\mathrm{EM}_y = \frac{2\sum_i^n a_i(y_i - y_c)}{n h_{\mathrm{frame}} \sum_i^n a_i} \tag{7}$$

where (x_i, y_i) and (x_c, y_c) are the co-ordinates of the centres of object i and the frame; a_i is the area of the object; b_{frame} and h_{frame} are the width and height of the frame; and n is the number of objects on the frame. Note that the maximum values of $|x_i - x_c|$ and $|y_i - y_c|$ are $b_{\mathrm{frame}}/2$ and $h_{\mathrm{frame}}/2$.

2.3. Measure of symmetry

Symmetry is axial duplication: a unit on one side of the centre line is exactly replicated on the other side. Vertical symmetry refers to the balanced arrangement of equivalent elements about a vertical axis, and horizontal symmetry about a horizontal axis. Radial symmetry consists of equivalent elements balanced about two or more axes that intersect at a central point.

Symmetry, by definition, is the extent to which the screen is symmetrical in three directions: vertical, horizontal, and diagonal and is given by

$$\mathrm{SYM} = 1 - \frac{|\mathrm{SYM}_{\mathrm{vertical}}| + |\mathrm{SYM}_{\mathrm{horizontal}}| + |\mathrm{SYM}_{\mathrm{radial}}|}{3} \in [0,1] \tag{8}$$

$\mathrm{SYM}_{\mathrm{vertical}}$, $\mathrm{SYM}_{\mathrm{horizontal}}$, and $\mathrm{SYM}_{\mathrm{radial}}$ are, respectively, the vertical, horizontal, and radial symmetries with

$$\mathrm{SYM}_{\mathrm{vertical}} = \frac{\begin{array}{l}|X'_{\mathrm{UL}} - X'_{\mathrm{UR}}| + |X'_{\mathrm{LL}} - X'_{\mathrm{LR}}| + |Y'_{\mathrm{UL}} - Y'_{\mathrm{UR}}| + |Y'_{\mathrm{LL}} - Y'_{\mathrm{LR}}| + \\ |H'_{\mathrm{UL}} - H'_{\mathrm{UR}}| + |H'_{\mathrm{LL}} - H'_{\mathrm{LR}}| + |B'_{\mathrm{UL}} - B'_{\mathrm{UR}}| + |B'_{\mathrm{LL}} - B'_{\mathrm{LR}}| + \\ |\Theta'_{\mathrm{UL}} - \Theta'_{\mathrm{UR}}| + |\Theta'_{\mathrm{LL}} - \Theta'_{\mathrm{LR}}| + |R'_{\mathrm{UL}} - R'_{\mathrm{UR}}| + |R'_{\mathrm{LL}} - R'_{\mathrm{LR}}|\end{array}}{12} \tag{9}$$

$$\mathrm{SYM}_{\mathrm{horizontal}} = \frac{\begin{array}{l}|X'_{\mathrm{UL}} - X'_{\mathrm{LL}}| + |X'_{\mathrm{UR}} - X'_{\mathrm{LR}}| + |Y'_{\mathrm{UL}} - Y'_{\mathrm{LL}}| + |Y'_{\mathrm{UR}} - Y'_{\mathrm{LR}}| + \\ |H'_{\mathrm{UL}} - H'_{\mathrm{LL}}| + |H'_{\mathrm{UR}} - H'_{\mathrm{LR}}| + |B'_{\mathrm{UL}} - B'_{\mathrm{LL}}| + |B'_{\mathrm{UR}} - B'_{\mathrm{LR}}| + \\ |\Theta'_{\mathrm{UL}} - \Theta'_{\mathrm{LL}}| + |\Theta'_{\mathrm{UR}} - \Theta'_{\mathrm{LR}}| + |R'_{\mathrm{UL}} - R'_{\mathrm{LL}}| + |R'_{\mathrm{UR}} - R'_{\mathrm{LR}}|\end{array}}{12} \tag{10}$$

$$\mathrm{SYM}_{\mathrm{radial}} = \frac{\begin{array}{l}|X'_{\mathrm{UL}} - X'_{\mathrm{LR}}| + |X'_{\mathrm{UR}} - X'_{\mathrm{LL}}| + |Y'_{\mathrm{UL}} - Y'_{\mathrm{LR}}| + |Y'_{\mathrm{UR}} - Y'_{\mathrm{LL}}| + \\ |H'_{\mathrm{UL}} - H'_{\mathrm{LR}}| + |H'_{\mathrm{UR}} - H'_{\mathrm{LL}}| + |B'_{\mathrm{UL}} - B'_{\mathrm{LR}}| + |B'_{\mathrm{UR}} - B'_{\mathrm{LL}}| + \\ |\Theta'_{\mathrm{UL}} - \Theta'_{\mathrm{LR}}| + |\Theta'_{\mathrm{UR}} - \Theta'_{\mathrm{LL}}| + |R'_{\mathrm{UL}} - R'_{\mathrm{LR}}| + |R'_{\mathrm{UR}} - R'_{\mathrm{LL}}|\end{array}}{12} \tag{11}$$

X'_j, Y'_j, H'_j, B'_j, Θ'_j, and R'_j are, respectively, the normalised values of

$$X_j = \sum_{i}^{n_j} |x_{ij} - x_c| \quad j = \mathrm{UL}, \mathrm{UR}, \mathrm{LL}, \mathrm{LR} \tag{12}$$

$$Y_j = \sum_{i}^{n_j} |y_{ij} - y_c| \quad j = \mathrm{UL}, \mathrm{UR}, \mathrm{LL}, \mathrm{LR} \tag{13}$$

$$H_j = \sum_{i}^{n_j} h_{ij} \quad j = \mathrm{UL}, \mathrm{UR}, \mathrm{LL}, \mathrm{LR} \tag{14}$$

$$B_j = \sum_{i}^{n_j} b_{ij} \quad j = \mathrm{UL}, \mathrm{UR}, \mathrm{LL}, \mathrm{LR} \tag{15}$$

$$\Theta_j = \sum_{i}^{n_j} \left| \frac{y_{ij} - y_c}{x_{ij} - x_c} \right| \quad j = \mathrm{UL}, \mathrm{UR}, \mathrm{LL}, \mathrm{LR} \tag{16}$$

$$R_j = \sum_{i}^{n_j} \sqrt{(x_{ij} - x_c)^2 + (y_{ij} - y_c)^2} \quad j = \mathrm{UL}, \mathrm{UR}, \mathrm{LL}, \mathrm{LR} \tag{17}$$

where UL, UR, LL and LR stand for upper-left, upper-right, lower-left and lower-right, respectively; (x_{ij}, y_{ij}) and (x_c, y_c) are the co-ordinates of the centres of object i on quadrant j and the frame; b_{ij} and h_{ij} are the width and height of the object; and n_j is the total number of objects on the quadrant.

2.4. Measure of sequence

Sequence in design refers to the arrangement of objects in a layout in a way that facilitates the movement of the eye through the information displayed. Normally the eye, trained by reading, starts from the upper left and moves back and forth across the display to the lower right. Perceptual psychologists have found that certain things attract the eye. It moves from big objects to small objects.

Sequence, by definition, is a measure of how information in a display is ordered in relation to a reading pattern that is common in Western cultures and is given by

$$\mathrm{SQM} = 1 - \frac{\sum_{j=\mathrm{UL,UR,LL,LR}} |q_j - v_j|}{8} \in [0, 1] \tag{18}$$

with

$$\{q_{\mathrm{UL}}, q_{\mathrm{UR}}, q_{\mathrm{LL}}, q_{\mathrm{LR}}\} = \{4, 3, 2, 1\} \tag{19}$$

$$v_j = \begin{cases} 4 & \text{if } w_j \text{ is the biggest in } w \\ 3 & \text{if } w_j \text{ is the 2nd biggest in } w \\ 2 & \text{if } w_j \text{ is the 3rd biggest in } w \\ 1 & \text{if } w_j \text{ is the smallest in } w \end{cases} \quad j = \mathrm{UL}, \mathrm{UR}, \mathrm{LL}, \mathrm{LR} \tag{20}$$

with

$$w_j = q_j \sum_{i}^{n_j} a_{ij} \quad j = \mathrm{UL}, \mathrm{UR}, \mathrm{LL}, \mathrm{LR} \tag{21}$$

$$w = \{w_{\mathrm{UL}}, w_{\mathrm{UR}}, w_{\mathrm{LL}}, w_{\mathrm{LR}}\} \tag{22}$$

where UL, UR, LL, and LR stand for upper-left, upper-right, lower-left, and lower-right, respectively; and a_{ij} is the area of object i on quadrant j. Each quadrant is given a weighting in q.

2.5. Measure of cohesion

In screen design, similar aspect ratios promote cohesion. The term aspect ratio refers to the relationship of width to height. Typical paper sizes are higher than they are wide, while the opposite is true for typical VDU displays. Changing the aspect ratio of a visual field may affect eye movement patterns sufficiently to account for some of the performance differences. The aspect ratio of a visual field should stay the same during the scanning of a display.

Cohesion, by definition, is a measure of how cohesive the screen is and is given by

$$\mathrm{CM} = \frac{|\mathrm{CM}_{\mathrm{fl}}| + |\mathrm{CM}_{\mathrm{lo}}|}{2} \in [0, 1] \tag{23}$$

$\mathrm{CM}_{\mathrm{fl}}$ is a relative measure of the ratios of the layout and screen with

$$\mathrm{CM}_{\mathrm{fl}} = \begin{cases} c_{\mathrm{fl}} & \text{if} \quad c_{\mathrm{fl}} \leqslant 1 \\ \dfrac{1}{c_{\mathrm{fl}}} & \text{otherwise} \end{cases} \tag{24}$$

with

$$c = \frac{h_{la55X\mathrm{fl}}/b_{\mathrm{layout}}}{h_{\mathrm{frame}}/b_{\mathrm{frame}}} \tag{25}$$

where b_{layout} and h_{layout} and b_{frame} and h_{frame} are the widths and heights of the layout and the frame, respectively. $\mathrm{CM}_{\mathrm{lo}}$ is a relative measure of the ratios of the objects and layout with

$$\mathrm{CM}_{\mathrm{lo}} = \frac{\sum_i^n t_i}{n} \tag{26}$$

with

$$t_i = \begin{cases} c_i & \text{if } c_i \leqslant 1 \\ \dfrac{1}{c_i} & \text{otherwise} \end{cases} \tag{27}$$

with

$$c = \frac{h_i/b_i}{h_{\mathrm{layout}}/b_{\mathrm{layout}}} \tag{28}$$

where b_i and h_i the width and height of object i; and n is the number of objects on the frame.

2.6. Measure of unity

Unity is coherence, a totality of elements that is visually all one piece. With unity, the elements seem to belong together, to dovetail so completely that they are seen as one thing. Unity in screen design is achieved by using similar sizes and leaving less space between elements of a screen than the space left at the margins.

Unity, by definition, is the extent to which the screen elements seem to belong together and is given by

$$\mathrm{UM} = \frac{|\mathrm{UM}_{\mathrm{form}}| + |\mathrm{UM}_{\mathrm{space}}|}{2} \quad \in [0, 1] \tag{29}$$

$\mathrm{UM}_{\mathrm{form}}$ is the extent to which the objects are related in size with

$$\mathrm{UM}_{\mathrm{form}} = 1 - \frac{n_{\mathrm{size}} - 1}{n} \tag{30}$$

and $\mathrm{UM}_{\mathrm{space}}$ is a relative measurement, which means that the space left at the margins (the margin area of the screen) is related to the space between elements of the screen (the between-component area) with

$$\mathrm{UM}_{\mathrm{space}} = 1 - \frac{a_{\mathrm{layout}} - \sum_i^n a_i}{a_{\mathrm{frame}} - \sum_i^n a_i} \tag{31}$$

where a_i, a_{layout}, and a_{frame} are the areas of object i, the layout, and the frame, respectively; n_{size} is the number of sizes used; and n is the number of objects on the frame.

2.7. Measure of proportion

Down through the ages, people and cultures have had preferred proportional relationships. What constitutes beauty in one culture is not necessarily considered the same by another culture, but some proportional shapes have stood the test of time and are found in abundance today. Marcus [3] describes the following shapes as aesthetically pleasing.

- Square (1:1)
- Square root of two (1:1.414)
- Golden rectangle (1:1.618)
- Square root of three (1:1.732)
- Double square (1:2)

In screen design, aesthetically pleasing proportions should be considered for major components of the screen, including windows and groups of data and text.

Proportion, by definition, is the comparative relationship between the dimensions of the screen components and proportional shapes and is given by

$$\mathrm{PM} = \frac{|\mathrm{PM}_{\mathrm{object}}| + |\mathrm{PM}_{\mathrm{layout}}|}{2} \in [0, 1] \tag{32}$$

$\mathrm{PM}_{\mathrm{object}}$ is the difference between the proportions of the objects and the closest proportional shapes described by Marcus with

$$\mathrm{PM}_{\mathrm{object}} = \frac{1}{n}\sum_{i}^{n}\left(1 - \frac{\min(|p_j - p_i|, j = sq, r2, gr, r3, ds)}{0.5}\right) \tag{33}$$

$$p_i = \begin{cases} r_i & \text{if } r_i \leqslant 1 \\ \dfrac{1}{r_i} & \text{otherwise} \end{cases} \tag{34}$$

with

$$r_i = \frac{h_i}{b_i} \tag{35}$$

where b_i and h_i are the width and height of object i. Note that the maximum value of $(p_j - p_i)$ is 0.5. $\mathrm{PM}_{\mathrm{layout}}$ is the difference between the proportions of the layout and the closest proportional shape with

$$\mathrm{PM}_{\mathrm{layout}} = 1 - \frac{\min(|p_j - p_{\mathrm{layout}}|, j = sq, r2, gr, r3, ds)}{0.5} \tag{36}$$

with

$$p_{\mathrm{layout}} = \begin{cases} r_{\mathrm{layout}} & \text{if } r \leqslant 1 \\ \dfrac{1}{r_{\mathrm{layout}}} & \text{otherwise} \end{cases} \tag{37}$$

with

$$r_{\mathrm{layout}} = \frac{h_{\mathrm{layout}}}{b_{\mathrm{layout}}} \tag{38}$$

where b_{layout} and h_{layout} are the width and height of the layout. Note that the maximum value of $(p_j - p_{\mathrm{layout}})$ is 0.5. p_j is the proportion of shape j with

$$\{p_{\mathrm{sq}}, p_{r2}, p_{\mathrm{gr}}, p_{r3}, p_{\mathrm{ds}}\} = \left\{\frac{1}{1}, \frac{1}{1.414}, \frac{1}{1.618}, \frac{1}{1.732}, \frac{1}{2}\right\} \tag{39}$$

where sq, $r2$, gr, $r3$, and ds stand for square, square root of two, golden rectangle, square root of three, and double square, respectively. While many guidelines documents recommend using these rectangles in both horizontal and vertical orientations, almost all of them exhibit no particular preference. For this reason, we assume that they are aesthetically indistinguishable, and therefore, we do not take into account the orientation of the proportion.

2.8. Measure of simplicity

Simplicity is directness and singleness of form, a combination of elements that results in ease in comprehending the meaning of a pattern. Simplicity in screen design is achieved by optimising the number of elements on a screen and minimising the alignment points. Tullis [12] has derived a measure of screen complexity for text-based screens based on the work of Bonsiepe [25], who proposed a method of measuring the complexity of typographically designed pages through the application of information theory. It involves counting the number of different rows or columns on the screen that are used as starting positions of alphanumeric data items. Information theory is then used to calculate the complexity of this arrangement of starting positions.

An easier method of calculation is

$$\mathrm{SMM} = \frac{3}{n_{\mathrm{vap}} + n_{\mathrm{hap}} + n} \in [0, 1] \tag{40}$$

where n_{vap} and n_{hap} are the numbers of vertical and horizontal alignment points; and n is the number of objects on the frame.

2.9. Measure of density

Density is the extent to which the screen is covered with objects. Density is achieved by restricting screen density levels to an optimal percent. A measure of density, derived by Tullis [11], is the percentage of character positions on the entire frame containing data.

Instead of looking at characters, our measure deals with objects with

$$\mathrm{DM} = 1 - 2\left|0.5 - \frac{\sum_i^n a_i}{a_{\mathrm{frame}}}\right| \in [0, 1] \tag{41}$$

where a_i and a_{frame} are the areas of object i and the frame; and n is the number of objects on the frame. Assume that the optimum screen density level for graphic screens is 50%.

2.10. Measure of regularity

Regularity is a uniformity of elements based on some principle or plan. Regularity in screen design is achieved by establishing standard and consistently spaced horizontal and vertical alignment points for screen elements, and minimising the alignment points.

Regularity, by definition, is a measure of how regular the screen is and is given by

$$\mathrm{RM} = \frac{|\mathrm{RM}_{\mathrm{alignment}}| + |\mathrm{RM}_{\mathrm{spacing}}|}{2} \in [0, 1] \tag{42}$$

$\mathrm{RM}_{\mathrm{alignment}}$ is the extent to which the alignment points are minimised with

$$\mathrm{RM}_{\mathrm{alignment}} = \begin{cases} 1 & \text{if } n = 1 \\ 1 - \dfrac{n_{\mathrm{vap}} + n_{\mathrm{hap}}}{2n} & \text{otherwise} \end{cases} \tag{43}$$

and $\mathrm{RM}_{\mathrm{spacing}}$ is the extent to which the alignment points are consistently spaced with

$$\mathrm{RM}_{\mathrm{spacing}} = \begin{cases} 1 & \text{if } n = 1 \\ 1 - \dfrac{n_{\mathrm{spacing}} - 1}{2(n-1)} & \text{otherwise} \end{cases} \tag{44}$$

where n_{vap} and n_{hap} are the numbers of vertical and horizontal alignment points; n_{spacing} is the number of distinct distances between column and row starting points; and n is the number of objects on the frame.

2.11. Measure of economy

Economy is the careful and discreet use of display elements to get the message across as simple as possible. Economy is achieved by using as few sizes as possible.

Economy, by definition, is a measure of how economical the screen is and is given by

$$\mathrm{ECM} = \frac{1}{n_{\mathrm{size}}} \in [0, 1] \tag{45}$$

where n_{size} is the number of sizes.

2.12. Measure of homogeneity

Entropy was developed in physics in the 19th century and was applied later in astronomy, chemistry and biology. Entropy influenced almost every science. We interpret the statistical entropy concept for screen design. The entropy equation is given by the following

$$S = k \log W \tag{46}$$

where S is the entropy of the screen; k is a constant, known as Boltzmann's constant; and W is a measure of the degree of homogeneity.

Since increases or decreases of W are equivalent, respectively, to increases or decreases of S, we can conveniently work with W below rather than with S. The relative degree of homogeneity of a composition is determined by how evenly the objects are distributed among the four quadrants of the screen. The degree of evenness is a matter of the quadrants that contain more or less nearly equal numbers of objects.

Homogeneity, by definition, is a measure of how evenly the objects are distributed among the quadrants and is given by

$$\mathrm{HM} = \frac{W}{W_{\max}} \in [0,1] \tag{47}$$

W is the number of different ways a group of n objects can be arranged for the four quadrants when n_j is the total number of objects in quadrant j, that is

$$W = \frac{n!}{\prod_{j=\mathrm{UL,UR,LL,LR}} n_j} = \frac{n!}{n_{\mathrm{UL}}!n_{\mathrm{UR}}!n_{\mathrm{LL}}!n_{\mathrm{LR}}!} \tag{48}$$

where n_{UL}, n_{UR}, n_{LL}, and n_{LR} are the numbers of objects on the upper-left, upper-right, lower-left, and lower-right quadrants, respectively; and n is the number of objects on the frame.

W is maximum when the n objects are evenly allocated to the various quadrants of the screen, as compared to more or less uneven allocations among the quadrants, and thus

$$W_{\max} = \frac{n!}{\frac{n}{4}!\frac{n}{4}!\frac{n}{4}!\frac{n}{4}!} = \frac{n!}{\left(\frac{n}{4}!\right)^4} \tag{49}$$

2.13. Measure of rhythm

Rhythm in design refers to regular patterns of changes in the elements. This order with variation helps to make the appearance exciting. Rhythm is accomplished through variation of arrangement, dimension, number and form of the elements. The extent to which rhythm is introduced into a group of elements depends on the complexity (number and dissimilarity of the elements).

Rhythm, by definition, is the extent to which the objects are systematically ordered and is given by

$$\mathrm{RHM} = 1 - \frac{|\mathrm{RHM}_x| + |\mathrm{RHM}_y| + |\mathrm{RHM}_{\mathrm{area}}|}{3} \in [0,1] \tag{50}$$

The rhythm components are

$$\mathrm{RHM}_x = \frac{\begin{array}{c}|X'_{\mathrm{UL}} - X'_{\mathrm{UR}}| + |X'_{\mathrm{UL}} - X'_{\mathrm{LR}}| + \\ |X'_{\mathrm{UL}} - X'_{\mathrm{LL}}| + |X'_{\mathrm{UR}} - X'_{\mathrm{LR}}| + \\ |X'_{\mathrm{UR}} - X'_{\mathrm{LL}}| + |X'_{\mathrm{LR}} - X'_{\mathrm{LL}}|\end{array}}{6} \tag{51}$$

$$\mathrm{RHM}_y = \frac{\begin{array}{c}|Y'_{\mathrm{UL}} - Y'_{\mathrm{UR}}| + |Y'_{\mathrm{UL}} - Y'_{\mathrm{LR}}| + \\ |Y'_{\mathrm{UL}} - Y'_{\mathrm{LL}}| + |Y'_{\mathrm{UR}} - Y'_{\mathrm{LR}}| + \\ |Y'_{\mathrm{UR}} - Y'_{\mathrm{LL}}| + |Y'_{\mathrm{LR}} - Y'_{\mathrm{LL}}|\end{array}}{6} \tag{52}$$

$$\mathrm{RHM_{area}} = \frac{\begin{array}{l}|A'_{\mathrm{UL}} - A'_{\mathrm{UR}}| + |A'_{\mathrm{UL}} - A'_{\mathrm{LR}}| + \\ |A'_{\mathrm{UL}} - A'_{\mathrm{LL}}| + |A'_{\mathrm{UR}} - A'_{\mathrm{LR}}| + \\ |A'_{\mathrm{UR}} - A'_{\mathrm{LL}}| + |A'_{\mathrm{LR}} - A'_{\mathrm{LL}}|\end{array}}{6} \tag{53}$$

X'_j, Y'_j, and A'_j are, respectively, the normalised values of

$$X_j = \sum_{i}^{n_j} |x_{ij} - x_c| \quad j = \mathrm{UL, UR, LL, LR} \tag{54}$$

$$Y_j = \sum_{i}^{n_j} |y_{ij} - y_c| \quad j = \mathrm{UL, UR, LL, LR} \tag{55}$$

$$A_j = \sum_{i}^{n_j} a_{ij} \quad j = \mathrm{UL, UR, LL, LR} \tag{56}$$

where UL, UR, LL, and LR stand for upper-left, upper-right, lower-left, and lower-right, respectively; (x_{ij}, y_{ij}) and (x_c, y_c) are the co-ordinates of the centres of object i on quadrant j and the frame; a_{ij} is the area of the object; and n_j is the total number of objects on the quadrant.

2.14. Measure of order and complexity

The measure of order is written as an aggregate of the above measures for a layout. The opposite pole on the continuum is complexity. The scale created may also be considered a scale of complexity, with extreme complexity at one end and minimal complexity (order) at the other.

The general form of the measure is given by

$$\mathrm{OM} = g\{f_i(M_i)\} \in [0, 1] \tag{57}$$

with

$$\begin{aligned}\{M_1, M_2, M_3, M_4, M_5, M_6, M_7, M_8, M_9, M_{10}, M_{11}, M_{12}, M_{13}\} \\ = \{\mathrm{BM, EM, SYM, SQM, CM, UM, PM, SMM, DM, RM,} \\ \mathrm{ECM, HM, RHM}\}\end{aligned} \tag{58}$$

where $f_i()$ is a function of M_i and is functionally related to the measurable criteria which characterise $g\{\}$; and BM is given by (1), EM by (5), SYM by (8), SQM by (18), CM by (23), UM by (29), PM by (32), SMM by (40), DM by (41), RM by (42), ECM by (45), HM by (47), and RHM by (50).

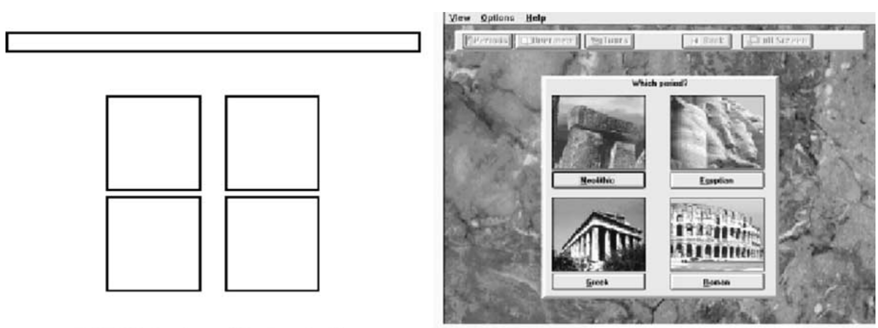
3. Empirical study

In this two-part exercise, we established aesthetic values for a group of real screens copied from multimedia systems using our measures for comparison with viewer judgements about the aesthetics of the actual screens and their screen models, and determined how consistent one was with the other. The specific goals were to confirm the practicality of the conventions we gave for determining the properties of screens, and compare judgements of real screens against model screens in order to determine if the same results carried across to more practical applications.

In our earlier studies [17–19] of the formulae (preceding versions), very high correlations were found between perceived and computed aesthetics of the interface. These studies were conducted using model screens. Model screens were used to control for content effects and to facilitate interpretation of the data analyses. Grabinger [10] found that the characteristics identified with model screens held up when viewers judged the readability and studyability of a selection of real screens from actual programs. However, there was no evidence to indicate whether people's perceptions of interface aesthetics would differ if the real screens were used rather than the screen models. Using real screens takes a further step in investigating whether the same characteristics used by viewers in the model screens emerge in judgements about real screens.

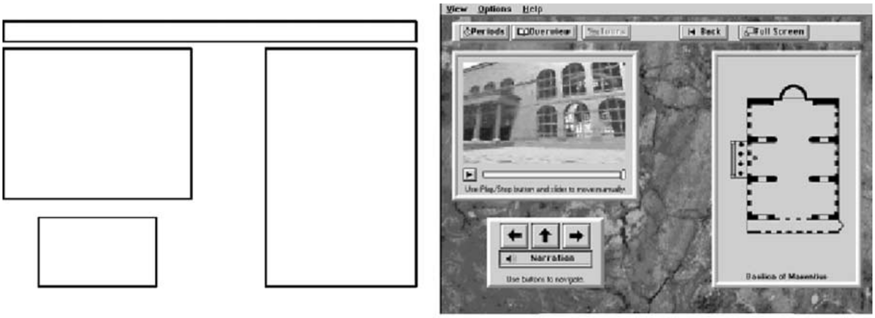
Samples of all screens are shown in Figs. 1–5. (The five screens and the corresponding models are side by side.) Table 1 presents the element configurations of the screens, and their aesthetic values, according to our formulae, are summarised in Table 2. (All values are in pixels.) To perform OM calculations for the screens, Eq. (57) is written as the linear summation of the weighted measures with

$$g\{\} = \frac{1}{m}\sum_{i}^{m}\alpha_i f_i(M_i) = \frac{1}{13}\sum_{i}^{13}\alpha_i M_i \quad \in [0,1], \quad 0 \leqslant \alpha_i \leqslant 1 \tag{59}$$



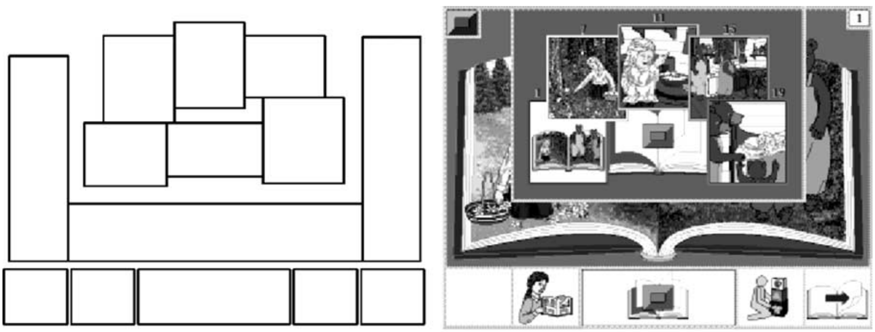
Model (structure without content) Real (structure with content)

Fig. 1. Exploring ancient architecture, by Medio Multimedia.



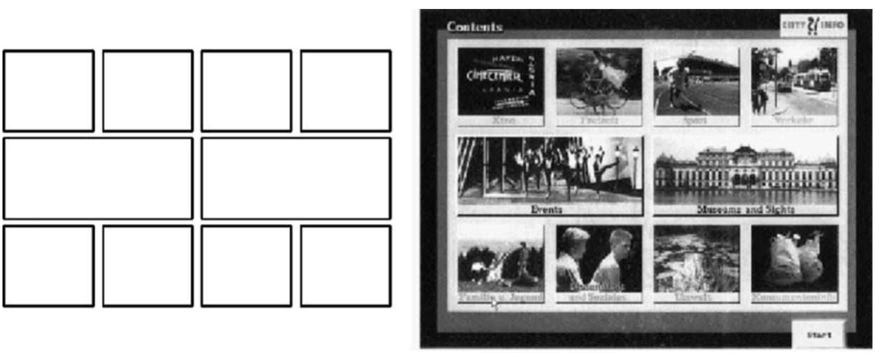
Model (structure without content) Real (structure with content)

Fig. 2. Exploring ancient architecture, by Medio Multimedia.



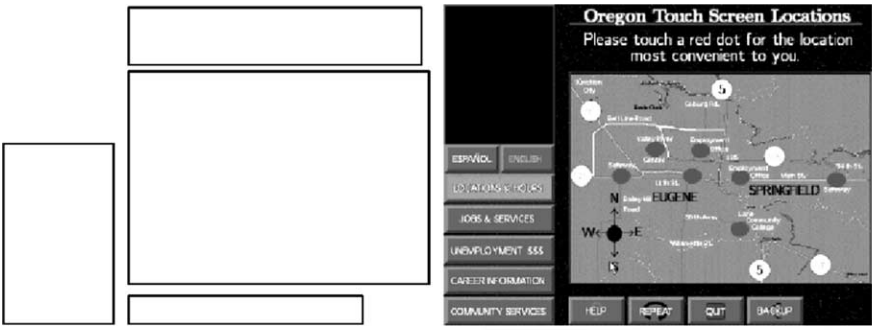
Model (structure without content) Real (structure with content)

Fig. 3. Goldilocks and the three bears.



Model (structure without content) Real (structure with content)

Fig. 4. The main menu of the CITY-INFO kiosk.



Model (structure without content)   Real (structure with content)

Fig. 5. A regional map showing oregon employment division kiosk locations.

Each aesthetic measure M_i has its own weighing component α_i, which is assumed to be a constant. (Determining weights is one of the multi-dimensional optimisation problems that are application specific. A paper presenting a solution using objective-based evolutionary programming is under review at the present time.) The weighting components are set to 1, supposing that all these measures are equally important. As indicated by the overall measure OM in Table 2, Figs. 1 and 4 are measured high, whereas Figs. 2, 3, and 5 medium. Of the layouts we consider, Fig. 4 has the highest aesthetic value and Fig. 5 the lowest.

On a visual basis, the formulae clearly separate Figs. 1 and 4 from Figs. 2, 3, and 5. The former group of screens shows strong aesthetic and organisational qualities. Figs. 1 and 4 are aesthetically balanced with well-defined areas, multiple columns of graphics, and with white space that is around the exterior screen margins. The other screens (2, 3, 5) show three characteristics that are viewed negatively: they are fragmented (mean UM $= 0.251$), irregular (mean RM $= 0.356$), and intricate (mean ECM $= 0.197$). All the screens (1–5) are rather complex (mean SMM ratings $= 0.226$). Fig. 3, the lowest rated screen (SMM $= 0.083$; see Table 2), appears cluttered as it contains overlapping screen elements. Galitz [23] was very much against using perspective for non-interactive elements. (Since our measures do not account for overlapping screen components, only the visible parts are included in the calculations.)

Part 1 was designed to test the robustness of the formulae using model screens. Participants in Part 1 were 79 undergraduate students from a university in Malaysia. Subjects were members of a variety of information technology classes and received credit for participating in the study. Although subjects were members of information technology classes, they were not familiar with screen design concepts. Part 1 collected viewer judgements using the screen aesthetics rating of model screens, i.e., screens that control for content effects. The five models (1–5) were displayed in a large classroom, using an overhead screen projector. Each layout was displayed for about 20 s. During

Table 1
Summary of layout properties

Layout	Object	X	Y	Width	Height
Fig. 1 (319 × 221)	1	80	53	70	70
	2	80	128	70	70
	3	168	53	70	70
	4	168	128	70	70
	5	6	5	306	16
Fig. 2 (319 × 221)	1	6	5	306	16
	2	6	25	140	112
	3	32	150	88	52
	4	200	25	112	177
Fig. 3 (320 × 240)	1	3	196	48	42
	2	53	196	48	42
	3	219	196	48	42
	4	269	196	48	42
	5	103	196	114	42
	6	196	68	61	65
	7	182	22	61	46
	8	130	12	52	65
	9	77	22	53	65
	10	63	87	61	48
	11	124	87	72	41
	12	7	37	44	154
	13	51	147	219	44
	14	270	23	44	168
Fig. 4 (320 × 240)	1	23	29	64	58
	2	93	29	64	58
	3	163	29	64	58
	4	233	29	64	58
	5	23	91	134	58
	6	163	91	134	58
	7	23	153	64	58
	8	93	153	64	58
	9	163	153	64	58
	10	233	153	64	58
Fig. 5 (320 × 240)	1	94	2	220	44
	2	0	104	84	136
	3	94	50	226	160
	4	94	218	176	22

that time, participants rated each layout on a low–medium–high scale regarding how beautify it was. The median for the ratings is calculated to show the relative aesthetics of the screens, according to the viewers (see Table 3).

By and large, the results resemble those obtained using the proposed measures. Figs. 1 and 4 were rated higher than Figs. 2, 3, and 5. Figs. 1 and 4

Table 2
Computation of aesthetic value for five layouts

	Fig. 1	Fig. 2	Fig. 3	Fig. 4	Fig. 5
BM	0.87412	0.63771	0.76319	0.99625	0.64216
EM	0.99368	0.98150	0.99570	1.00000	0.98134
SYM	0.66871	0.57828	0.73858	0.99850	0.44697
SQM	1.00000	0.50000	1.00000	1.00000	0.25000
CM	0.71578	0.74103	0.80032	0.80116	0.72686
UM	0.52435	0.32515	0.28404	0.87668	0.14543
PM	0.89779	0.83207	0.86032	0.86859	0.79443
SMM	0.27273	0.30000	0.08333	0.17647	0.30000
DM	0.69493	0.72407	0.53823	0.82188	0.40792
RM	0.51250	0.37500	0.31868	0.79722	0.37500
ECM	0.50000	0.25000	0.09091	0.50000	0.25000
HM	0.69444	0.00463	0.01531	1.00000	0.00714
RHM	0.66176	0.54721	0.76674	0.99840	0.46527
OM	0.69314	0.52282	0.55810	0.83347	0.44558

Table 3
Median of the sample (Part 1)

Fig. 1	Fig. 2	Fig. 3	Fig. 4	Fig. 5
High	Medium	Medium	High	Low

present two layouts that have high computed values and were rated high on aesthetics by the viewers. Fig. 5, the lowest measured screen, was rated low. Figs. 2 and 3 with moderate OM were rated medium. The findings are consistent with those obtained in the earlier studies that used simple, artificial models to analyse previous versions, reinforcing the patterns observed thus far.

While the use of models permits an examination of viewer perceptions in a controlled environment, generalising from the model screens to real applications may not be tenable [26]. Part 2 was designed to investigate whether judgements would differ if the real screens were used rather than the screen models. Participants in Part 2 were 180 undergraduate students from the same university. Subjects were members of a variety of information technology classes and received credit for participating in the study. None of the subjects participated in Part 1 and none was familiar with screen design concepts. The models employed in Part 1 were free from content, therefore enabling subjects to focus on the overall design of the screen. In contrast, Part 2 examined viewer judgements of real screens copied from existing multimedia systems. Part 2 proceeded in the same manner as Part 1.

Part 1's results are replicated here. No differences were found, with the results being given in Table 4. Figs. 1 and 4 were rated higher than Figs. 2, 3,

Table 4
Median of the sample (Part 2)

Fig. 1	Fig. 2	Fig. 3	Fig. 4	Fig. 5
High	Medium	Medium	High	Low

and 5. Part 2 indicated, as did Part 1, that our formulae are important evaluative metrics for accessing the aesthetics of the real screens. The actual content of the real screens did not seem to have any effect on judgement, nor did informal post-study interviews indicate that content was a factor in judging the screens. Note that the formulae do not focus on content, but on structure. The above findings are encouraging as they justify the formulae and the conventions in general.

3.1. Discussion

The purpose of this study is to discover principal determinants of the aesthetic evaluation of the interface. The results of Study 1, using model screens, confirm and further define the existence of evaluative constructs based on visual complexity and organization. However, this study differed from earlier studies in that more determinants were found. Specifically, this study defined symmetry, sequence, cohesion, regularity, homogeneity, and rhythm.

While the use of models permits an examination of viewer perceptions in a controlled environment, generalizing from the model screens to real applications may not be tenable. The purpose of comparing judgments of real screens against model screens was to determine if the same results carried across to more practical applications. Study 2 indicated, as did Study 1, that these thirteen characteristics are important determinants of system acceptability.

Our empirical study has been done for the one-to-one correspondence between a real screen (structure with content) and its model (structure without content). However, if the same information content is illustrated on the five layout structures shown in Figs. 1–5, the same empirical results as in Table 3 or 4 could be expected (assuming our findings are robust).

The findings of this research greatly help to clarify the choices that are involved. Many long-held beliefs have found detailed empirical confirmation for the first time; some new and perhaps surprising discoveries have been made. For example, the study indicated that cohesion and rhythm are important determinants in judging acceptability from the appearance of the screens. These and other determinants of system acceptability have not received extensive treatment in the screen design literature. This study calls attention to their potential contribution to our evaluation of system acceptability.

The results also support the validity of the prediction system. Given a set of alternative formats, the system can quite accurately predict the relative subjective ratings for the formats.

4. Conclusions and future work

In this paper, we have presented a computational theory of evaluating interface aesthetics. In particular, we have introduced 14 aesthetic measures: balance, equilibrium, symmetry, sequence, cohesion, unity, proportion, simplicity, density, regularity, economy, homogeneity, rhythm, and order and complexity. Although the population samples used in our informal study limit generalisability, in view of consistency with previous studies [17–19], the findings have confirmed and further defined the existence of evaluative characteristics based on visual order and complexity. It should be emphasised that we have had to make two assumptions, namely, (1) that the interaction between the selected characteristics is linear, and (2) that all these characteristics are equally important. The main justification for these assumptions lies in the fact that they lead to results in complete agreement with experiment.

The study described here has suggested some improvements to enhance their usability. We can increase scope to include the colour, tone, and shape of objects in balance, for example. A designer can control some elements of composition to achieve balance. For instance, colour is visually heavier than black and white; big things are visually heavier than little things; black is visually heavier than white; irregular shapes are visually heavier than regular shapes. By controlling the colour, size, tone, and shape of objects in a design, one distributes the visual weight and thus influences balance.

A layout is in equilibrium when its centre corresponds approximately to the centre of the frame. Practically speaking, there are however, minor deviations from this definition. Owing to the visual gravitational pull, the balancing centre of the layout will lie somewhat higher than the centre of the frame, thereby compensating for the greater weight of the area's upper half. But such discrepancies are small.

There are many interesting research topics involving the computation and use of our formulae. First, experiments must be conducted to provide additional empirical validation of the formulae and the conventions. Future research should focus on investigating the interplay between the selected characteristics, which, contrary to our original assertion, may be nonlinear. Additional research is also necessary to evaluate the effects of different weighting strategies. (Weighting deals with the problem that we care about some characteristics more than others.) Characteristics that are common to the feeling which gives one an aesthetic experience should not be limited to the few, more accordant ordering principles with appropriate design conventions must be found if this approach is to be improved.

This study has lent initial support to the validity of our formulae as an evaluative metric for interface aesthetics. The most exciting possibility for these formulae would be to use it in conjunction with an automated layout program that incorporates additional details about information content and style guidelines. The formulae could be used when generating the initial layout. Subsequently, as the designer makes changes, the formulae and other metrics could provide feedback concerning the possible effects of the changes on the user's ability to interpret the information being presented. The formulae are intended to supplement other metrics, not to stand-alone.

Acknowledgements

The research reported here was supported by IRPA grant 191/9636/0012. Thanks to Boon Leong Law for his help in conducting the experiments and analysing the data.

References

[1] J. Maquet, The Aesthetic Experience, Yale University Press, New Haven, CT, 1986.
[2] J.D. Foley, A. van Dam, S.K. Feiner, J.F. Hughes, Computer Graphics: Principles and Practice, second ed., Addison-Wesley, Reading, MA, 1990.
[3] A. Marcus, Graphic Design for Electronic Documents and User Interfaces, ACM Press, New York, 1992.
[4] N. Tractinsky, Aesthetics and Apparent Usability: Empirically Assessing Cultural and Methodological Issues, in: CHI '97 Conference Proceedings, Association for Computing Machinery, New York, 1997.
[5] M. Kurosu, K. Kashimura, Apparent Usability vs. Inherent Usability, in: CHI '95 Conference Companion, Association for Computing Machinery, New York, 1995.
[6] S.C. Toh, Cognitive and motivational effects of two multimedia simulation presentation modes on science learning, Ph.D. Dissertation, University of Science Malaysia, Malaysia, 1998.
[7] M. Aspillaga, Screen design: a location of information and its effects on learning, Journal of Computer-Based Instruction 18 (3) (1991) 89–92.
[8] M. Szabo, H. Kanuka, Effects of Violating Screen Design Principles of Balance, Unity and Focus on Recall Learning, Study Time, and Completion Rates, in: ED-Media/ED-Telecom 98 Conference Proceedings, Association for the Advancement of Computing in Education, Charlottesville, VA, 1998.
[9] J. Heines, Screen design strategies for computer-assisted instruction, Digital Press, Bedford, MA, 1984.
[10] R.S. Grabinger, Computer screen designs: viewer judgements, Educational Technology Research and Development 41 (2) (1991) 35–73.
[11] T.S. Tullis, An evaluation of alphanumeric, graphic, and colour information displays, Human Factors 23 (1981) 541–550.

[12] T.S. Tullis, Predicting the usability of alphanumeric displays, Ph.D. Dissertation, Rice University, Kansas, 1984.
[13] R.S. Keister, G.R. Gallaway, Making software user friendly: an assessment of data entry performance, in: HFS 27th Annual Meeting Proceedings, Human Factors Society, Santa Monica, CA, 1983.
[14] T.S. Tullis, Screen design, in: M. Helander (Ed.), Handbook of Human–Computer Interaction, Elsevier Science Publishers, Amsterdam, The Netherlands, 1988, pp. 377–411.
[15] D.J. Streveler, A.I. Wasserman, Quantitative measures of the spatial properties of screen designs, in: INTERACT '84 Conference Proceedings, North Holland, Amsterdam, 1984.
[16] A. Sears, Layout appropriateness: guiding user interface design with simple task descriptions, IEEE Transactions on Software Engineering 19 (7) (1993) 707–719.
[17] D.C.L. Ngo, Visit: visitor information system implementation tool, Ph.D. Dissertation, Trinity College, Dublin, 1994.
[18] D.C.L. Ngo, J.G. Byrne, Aesthetic measures for screen design, in: OZCHI '98 Conference Proceedings, IEEE Computer Society, Los Alamitos, CA, 1998.
[19] D.C.L. Ngo, A. Samsudin, R. Abdullah, Aesthetic measures for assessing graphic screens, Journal of Information Science and Engineering 16 (1) (2000) 97–116.
[20] L. Arlov, GUI design for dummies, IDG Books Worldwide Inc., Foster City, CA, 1997.
[21] R. Arnheim, Art and visual perception, University of California Press, Berkeley, CA, 1954.
[22] D.A. Dondis, A primer of visual literacy, The MIT Press, Cambridge, MA, 1973.
[23] W.O. Galitz, The essential guide to user interface design: an introduction to GUI design principles and techniques, John Wiley & Sons, Inc., New York, 1997.
[24] S.S. Reilly, J.W. Roach, Improved visual design for graphics display, CG&A (February) (1984) 42–51.
[25] G. Bonsiepe, A method of quantifying order in typographic design, Journal of Typographic Research 2 (1968) 203–220.
[26] S.M. Ross, G.R. Morrison, In search of a happy medium in instructional technology research: issues concerning external validity, media replications, and learner control, Educational Technology Research and Development 37 (1) (1989) 19–33.