Human-Computer Interaction -- INTERACT'03
M. Rauterberg et al. (Eds.)
Published by IOS Press, (c) IFIP, 2003, pp. 144-151

Designing for Visual Influence: an Eye Tracking Study of the Usability of Graphical Management Information

J.A. Renshaw[†], J.E. Finlay[†], D. Tyfa[‡] and R.D. Ward[‡]
[†]Leeds Metropolitan University, Beckett Park, Leeds, LS6 3QS, UK
[‡]University of Huddersfield, Queensgate, HD1 3DH, UK
t.renshaw@lmu.ac.uk

Abstract: In this paper we explore how specific visual designs influence eye movement and how eye tracking can help us to understand this process. We tracked the eye movements of participants whilst they carried out tasks using two contrasting styles of graph. One followed established design guidelines for facilitating efficient visual processing; the other flouted them. Traditional performance measures confirmed the superiority of the guideline-supporting graphs; users also preferred them. Eye tracking measures also indicated the superior usability of the well designed graphs, but, in addition, provided more specific information about the role of different graph areas in the assimilation of data. A new way of analyzing eye movements into vertical and horizontal gazes allowed us to compare the influence of specific design features on eye movement. Results suggest that certain design features influence eye movement in a predictable way and that eye tracking techniques are sensitive enough to detect them.

Keywords: graphs, eye tracking, design, usability, influence, gazes

1 Introduction

Eye movements are thought to provide an indication of the amount of cognitive processing a display requires and, hence, how easy it is to process (Rayner and Pollatsek, 1994) and this is the assumption behind the use of eye tracking as a way of contributing towards the assessment of usability.

Visual processing can be conscious and controlled (as is the case in reading) or unconscious and involuntary (Card, 1999). Specific design features, such as movement, pattern and proximity, can influence this visual processing. If such features are designed to stimulate efficient eye movements, then we might expect a lower cognitive load and greater usability. In this paper we explore how far eye tracking can identify the visual influence of specific design features, by examining the impact of particular design styles on the interpretation of graphical management information.

Graphs are an increasingly common way of conveying summary information. They provide a means of explaining relationships between variables and they are frequently used in presentations in many environments such as commerce and education.

The graphic facilities available in basic spreadsheet packages are numerous and the design formats diverse. They are frequently used by non-graphic specialists to communicate management information but, although there are many packages to produce graphs, these provide little advice on which design will be most effective, efficient and accurate in communicating information to the audience.

We compared the usability of two graphical formats, one designed in accordance with, and one in contravention of, established design guidelines. The usability of each format was evaluated by the conventional measures of success rates, time to completion, and user satisfaction together with more recently established eye movement based tests (Goldberg & Kotval, 1998; Cowen, Ball, & Delin, 2002). Conventional techniques can be used to establish the relative usability of the resultant designs, but eye tracking, we argue, can add new and interesting insights – the means of monitoring aspects of user cognitive processes and of determining how specific visual features influence eye movement.

We included a number of specific design features in the first graph type, which we hypothesised would influence eye movement in a predictable way (in this case vertically within a particular screen area). A new eye tracking measure – gaze orientation – was developed to categorise gazes as being either vertical or horizontal in relation to the display scene. This was used to determine whether eye movement patterns within specific graph areas conformed to our expectations.

We begin by locating our work within research on the visual design of graphical information and the use of eye tracking for assessing usability. We then present the experiment and our results and draw conclusions about some of the benefits of using eye-tracking to evaluate usability.

1.1 Visual Design of Graphical Design.

Graphs afford a means of summarising large quantities of data, providing visualisations of the information and insights into the relationship between variables, whilst at the same time being less demanding on the internal memory task than other presentation forms such as tables. The use of graphs helps visualisation and cognition by grouping information that is used together, so reducing search times and the demand on working memory (Larkin and Simon, 1987).

Many graphical tools provide a large variety of design features including various colours and forms, many of which have only an aesthetic purpose, if that, rather than being a driver of good design. However there are many established guidelines for the effective design of graphical information, as well as theoretical proposals for how visual information is processed.

When people look at a page or the environment around them they actively organise what they see. They resolve ambiguities, impose structure and make connections (Schriver, 1997). In effective graphical information, priority is given to clarification of the data and to the preservation of its integrity. Emphasis is placed on data whilst clutter (non-data) is minimised (Tufte, 1990). False perspectives and distortion on the scale on the graphical axes are avoided (Bertin, 1983). The most effective design is one in which the answer to the question, that the graph is designed to address, can be perceived in a single image (Bertin, 1983). While Bertin and Tufte offer no theoretical basis for their guidelines (basing them on practical experience, technology limitations and aesthetics) their advice reflects Gestalt principles of continuance, proximity and closure.

Vincow and Wickens (1993) demonstrated the benefits of designs that clustered information and did not demand complex integration activity, while Chandler & Sweller (1991) concluded that separating related information caused heavy cognitive loads as does the introduction of seemingly useful but non essential explanatory material.

Visual information is processed in two ways:

- Controlled processing (such as reading), where processing is detailed, serial, low capacity, slow, able to be inhibited, and conscious.
- Automatic Processing, which is superficial, parallel, high capacity, fast, cannot be inhibited, load independent, and unconscious.

Automatic stimulus-based “attention-shifting”, draws the visual mechanism towards either movement or areas where pre-attentive features have identified strong patterns of colour, intensity or size contrasts (Card,1999). Automatic processing can therefore be stimulated by the inclusion of these elements.

The two graphical styles used in our experiment were designed to support and frustrate these design guidelines respectively (see Section 2 for details). In addition, visual patterns and coding were incorporated into the first graph type in an attempt to influence automatic visual processing.

1.2 Eye Movement Analysis in Usability

Studies have established that the eye tracking technique can be used to measure usability in specific contexts (Goldberg and Kotval, 1998, Cowen et al. 2002). The technology has progressed steadily from highly invasive procedures involving the attachment of eye caps to the cornea using anaesthetics (Yarbus, 1967) to remote non-invasive tracking methods that do not involve any contact between the equipment and the participant. Furthermore, software developments have both improved the control of eye tracking equipment and facilitated the analysis and visualisation of the large volumes of data produced. However, there is still relatively little use of eye tracking to examine the influence of specific design decisions on eye movement.

Cowen et al. (2002) cites Buckingham (1931) as having doubts as to the experimental benefits of manipulating typographical layouts because of the problems associated with controlling variables in realistic settings. However, we would argue that usability must be assessed within context and eye tracking analysis must be able to evaluate the impact of realistic images on eye movement. In addition, early studies of graphical layout did not have the benefit of eye tracking technology, which provides more detailed insights into visual behaviour than other experimental methods.

The graphs used in this experiment have been designed using realistic data and an off the shelf application (Excel). They exploit the fact that details such as textual description are accessible by controlled processing, and that visual coding techniques can be deployed to aid search and pattern detection. It is hypothesised that at least some of the results of these cognitive processes can be identified by tracking eye movements, the mechanisms for which have been suggested in detail by several researchers (e.g. Rayner, 1994).

Name	Description
Fixation	A relatively stable position within some threshold of dispersion (typically ~ 2 degrees) over some minimum duration (typically 100-200 ms), and with a velocity below some threshold (typically 15-100 degrees per second).
Saccade	Rapid ballistic movements of the eye from one point of interest to another, whose trajectory cannot be altered once begun. Saccades take between 30 and 120 ms and may cover between 1-40 degrees of visual angle.
Gaze	A series of consecutive fixations within an area of interest.
Scan Path	A spatial arrangement of a sequence of fixations. The saccade-fixation-saccade sequence of eye movement defines a scan path during search and processing tasks.

Table 1: Types of Eye Movement

Several basic types of eye movements have been defined in the literature. The types used in this research – fixations, saccades, gaze and scan paths – are summarised in Table 1.

Goldbergh and Kotval (1998) identify a number of measures for assessing usability:

- Number of Fixations. The number of fixations overall is thought to be negatively correlated with search efficiency.
- Fixation Duration. Longer fixations are believed to be an indication of the difficulty a participant has in extracting information from a display.
- Scan Path. This indicates areas of interest, cognitive load, and various search strategies. The optimal scan path for a task is a straight line to a desired target with a short duration fixation at the target.
- Scan Path length. An unduly long scan path length could indicate non-meaningful representations or poor layout.

In addition to using these, we have developed a further measure – the “gaze orientation”. This is derived from the coordinates of the fixations at either end of a gaze and is defined as the change in the vertical co-ordinates as a ratio of the change in the horizontal co-ordinates of a fixation pair. This is represented by the formula:

$$\text{Gaze orientation ratio} = |Y_{f1} - Y_{f2}/(X_{f1} - X_{f2})| \quad (1)$$

where X_{f1} and X_{f2} are the horizontal co-ordinates of the start fixation and the end fixation respectively and Y_{f1} and Y_{f2} are the vertical co-ordinates of the start and end fixations.

Gazes with a ratio of up to 1 were designated as horizontal whereas those with ratios greater than 1 were designated vertical.

2 Experimental Design and Hypotheses

The experiment was a repeated measures design with each participant viewing 8 graphs of two basic design formats (Graph Type 1 and Graph Type 2). Each participant viewed each graph one at a time in a random sequence until all 16 graphs had been displayed. Two levels of questions provided a cognitive load, including an “elementary” level question and a “global” level question (Bertin, 1983), with a unique question set for each graph.

Each graph was designed using standard Excel XP graphic facilities, and featured the same variables: sales, costs and profit on the vertical axis, with time on the horizontal axis.

The design of Graph Type 1 (see Figure 1) incorporated the following design guidelines:

putting the data element in the foreground and the axes in the background (Schriver, 1997; Ware, 2000)). The greater the contrast the more salient the effect (Kosslyn, 1994).

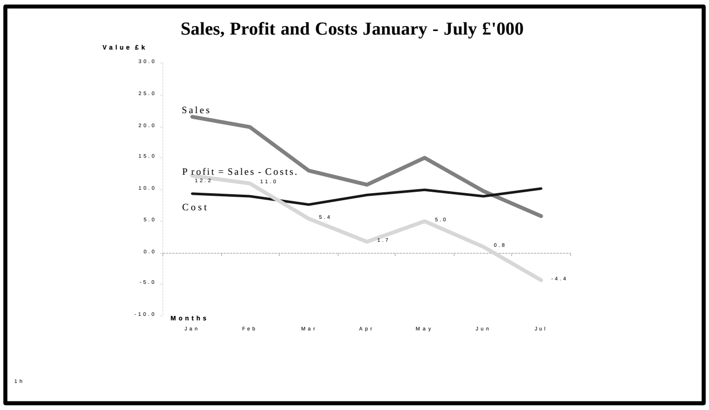

Figure 1: Example of Type 1 Graph

- Clutter was minimised (Tufte, 1990). For example, there were no gridlines, and a restriction on the number of data points labelled.
- The graph area was enclosed by a border indicating that it was to be viewed as a whole. This reinforced the relationship between Sales, Profit and Cost (Schriver,1997; Ware, 2000).
- The data lines were colour coded from black to light grey indicating the increase in relative importance of each data line (Bertin, 1983).
- The white background and black to grey data lines provided maximum contrast and emphasised the importance of the data lines by
- The proportion of black was low. Black should represent between 5-10% of the meaningful area of the plane for optimum retinal legibility (Bertin, 1983). Black was restricted to data lines and legend. Axes and scale were grey.
- The alignment of the headings and data labels reinforced their relationship. Their relative size and position indicated their relative importance and the sequence in which they should be read (Schriver,1997; Ware, 2000).
- The juxtaposition of the labels next to the data lines they represented emphasised their relationship and reduced the cognitive load on

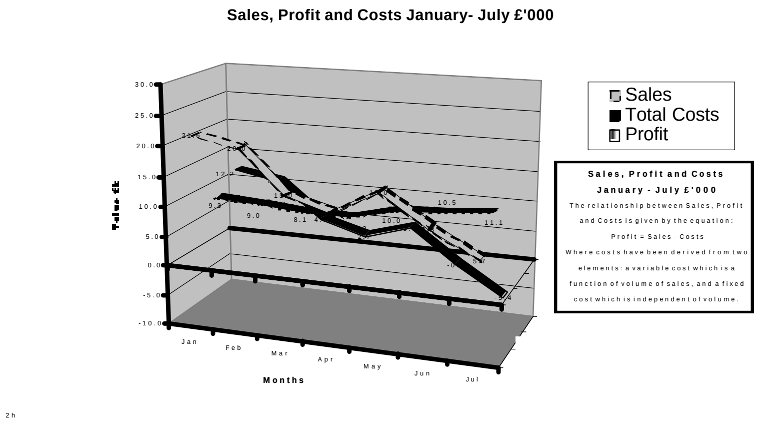

Figure 2: Example Type 2 Graph

working memory (Chandler and Sweller, 1991; Vincow and Wickens, 1993).

In contrast, Graph Type 2 (see Figure 2) was designed in contravention of standard guidelines by using

- Poor contrast, obscuring the data and frustrating interpretation.
- A false third dimension, which increased the cognitive workload.
- No discernable structure to the colour coding of the data lines.
- Legend and explanatory text in the form of notes within visually strong frames in a prominent position remote from the main body of the data (Chandler and Sweller, 1991).
- Strong non-data related frames in both the foreground and background, visually splitting up the data.

The study tested two groups of hypotheses. The first group addressed the general usability of the designs and predicted that the design features incorporated into Graph Type 1 would result in increased usability compared with Graph Type 2. Specifically, we hypothesised that this would be evidenced by greater speed and accuracy of the answers provided, shorter total fixation time and shorter scan path lengths. Participants would also rate Graph Type 1 as more usable.

The second hypothesis examined an element of design influence in more detail. It held that the design of Graph Type 1 would promote more efficient search patterns (vertical gazes in this case in the legend area) than the Graph Type 2 designs. In Graph Type 1, legend elements were vertically aligned both with each other and with the graph heading.

Legend descriptions were in close proximity to the data lines and positioned at the beginning of each data line. Data lines were progressively colour coded using black to grey. There were no boxes around the legend area and no clutter. Graph type 2 design afforded none of these advantages.

3 Method

3.1 Participants

24 volunteer participants (15 male, 9 female) took part in the experiment. Their ages ranged from 20 to 55 and they were drawn as an opportunity sample from the mathematics, computing, multimedia and health faculties of two universities. Sixteen wore no corrective lenses, five wore glasses and three wore contact lenses. Their dominant eye was identified and all were successfully calibrated. Two declared themselves to be novices in the use of graphs; the remainder were quite familiar: three stated they used graph very frequently, thirteen frequently, eight hardly ever.

3.2 Apparatus and Materials

The experiment used an ASL504 pan/tilt eye tracker system capable of detecting a bright back lit pupil image caused by the retinal reflection of a near infrared beam of light emanating from the eye tracker, which was placed immediately in front of the participant just below the display screen. The resolution of the eye tracker is better than one degree and it has a sampling rate of 60 Hz. The data was recorded using GazeTracker software, capable of recording fixation data and mouse clicks. ASL Eyepos software was used as a back up method of recording information on a second PC, in addition to the video recording of eye movements by means of a conventional video recorder. Participants were seated approximately 60 cms from the screen in a high backed chair – no other restraining mechanism was used. The graphs, created in Excel XP, were shown on a 17 inch PC colour monitor with a screen area of 1024 x 768 pixels. Software was used to select the images at random and to control the timed presentation of the cue for the first question, the “pause” splash screen, the second question screens and to record and display the participants’ answers at the end of each session. Time stamped fixation and mouse click data was exported to spreadsheets for subsequent analysis.

3.3 Procedure

A hybrid graph containing features of the two designs was created to explain the basic features of the graphs, and the responses to the display that would be required. A pre-prepared script was read to each participant.

Participants were asked to keep their head as still as possible and look at the centre of the screen during the experiment. The first graph was displayed and the participant allowed to gaze at it for a period of 15 seconds after which the test supervisor, cued by the disappearance of a coloured box placed at the lower edge of the graph display, asked the first question. The graph remained until the participant clicked with the mouse at the point on the display that represented the answer. The graph was then replaced by a screen that immediately presented the second question, which had four multiple choice answers. The participant was given 10 seconds to answer. This cycle was repeated for all 16 graphs. When the final question had been answered, the system displayed the answers recorded and the participant was asked to rate the graph formats on a scale of 1 to 5.

4 Results and Discussion

4.1 Usability Measures and Eye Tracking

The first hypothesis requires that the usability of each graph design be assessed by both performance metrics and established eye tracking measures. To this end the following performance data was collected: the total time taken to find the answer to the first question and its correctness, the correctness of answers to the second question and the participants' rating of each graph type's usability.

The direct eye tracking data recorded were the number of fixations, their "x and y" co-ordinates, start and finish times. Additional data was derived via computations on these measures: the scan path length between each fixation, the total scan path length for the duration of each graph's display, the gaze orientation, the area within which the fixations took place, the number and average length of gazes.

The descriptive statistics for each measurement are shown in Table 2. Two tailed, paired samples, t-tests were used to establish whether there were significant differences between the two styles at the 5% alpha level.

As discussed earlier, these measures are related to usability. All measures, except for the total number of fixations per graph and the error rate for the first question, showed a significant difference between the two formats, with Graph Type 1 being more usable according to these measures.

The total number of fixations measure is perhaps less meaningful than the total fixation duration when measuring cognitive load. The total number of fixations is only part of the story, the other element being the duration of each individual fixation and, if this varies significantly between fixations, measuring the total number only, may be misleading. A paired sample t-test on the fixation durations did in fact show a significant difference between the mean fixation durations of 387.12 ms for Graph Type 1 and 432.2 for Graph Type 2 [$t(23) = -7.383$, $p<.0005$].

Test	**Graph Type 1 Mean**	**Graph Type 2 Mean**	**t (23)**	**p (2 – tailed)**
Time taken to answer the first question (s)	22.88	24.60	-3.24	.004
Error rate for the first question	0.63	1.17	-1.97	.062
Error rate for the second question	2.25	3.96	-3.92	.001
The number of fixations for the duration of the graph display.	38.80	39.83	-0.83	.416
Average total fixation duration for the duration of the graph display.	14.91	17.05	-4.00	.001
The scan path length for the duration of the graph display.	5915.46	7856.54	-7.40	<.0005
Notes:				

1. The error rates for a question were determined from the number of errors made by each participant. An error being defined as either an incorrect answer or no answer recorded.
2. Eye tracking movements were collected at a sampling rate of 60 Hz. The minimum duration of a fixation was 100ms and the gaze point could not deviate by more than 60 pts.
3. The calculation of scan path length is derived from the co-ordinates of fixations. These co-ordinates have no physical unit (eg millimeters, pixels) but are related to the calibration area setting. The calibration area in the present experiment was 950 pt x 840 pts.

Table 2: Descriptive Statistics and t-test results

The lack of significant difference in the error rate in responses to the first question may be a result of the experimental method, as no time pressure was applied to the participants when answering the question. The experiment proceeded only when an answer was given and participants may have given themselves time to get the answers right.

The error rate for the second question is an indication of the comparative memorability of the two formats. Bertin (1983) asserted that that a good test of a graph was whether it could be recalled. The result of this test further establishes graph 1's superiority.

The usability assessments given by the participants showed a preference for Graph Type 1. In a scheme where usability was given a rating of between 1(low) and 5 (high) the average ratings given by participants overall for Graph Type 1 and Type 2 were 4.5 and 1.7 respectively. A paired sample t-test established this to be significant (t (23) = 15.91, p<0.0005).

discerned in the various areas of the graph, particularly the legend areas, the number of vertical and horizontal gazes were established by means of the gaze orientation described above. These were counted and their average length computed. The results are shown in Figure 3 which shows that the number of gazes originating in the data area exceeds those in the other areas in every case.

However, the number of gazes in the legend area is maintained only in the case of Graph Type 1. This is attributed to the design features of this graph format (unboxed, vertically aligned legend elements, descriptions in close proximity to colour coded data lines). These results support the second hypothesis. The design has influenced the search pattern by encouraging vertical gazes, enabling the more efficient covering of the legend area – key to the understanding of the graph.

To establish the significance of the difference in the number of vertical gazes in the legend areas of the two graph types, a two-way repeated measures

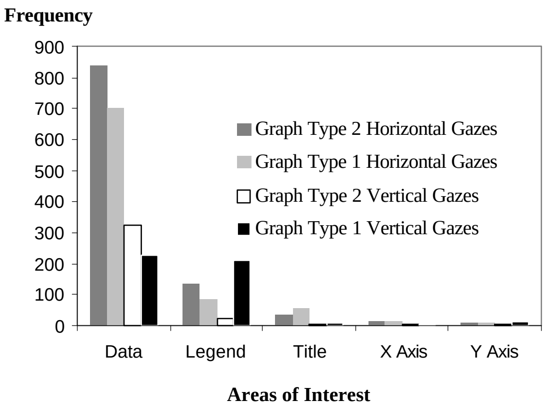

Figure 3: Frequency of Gazes by Area of Interest for each Graph.

In summary, data from performance measures, participant evaluations and standard eye movement records were analysed and found to support the first set of hypotheses.

4.2 Vertical and Horizontal Gazes

The second hypothesis required the analysis of gazes. There were five areas of interest: the legend, data, title and the two axis areas. The surface areas of the legend area for both graph types were different but the vertical height subtended the same visual angle (approx 7 deg.). Other areas were within 5% of each other. In order to assess whether different patterns of eye movement could be

ANOVA (analysis of variance) (graph type x area of interest) was performed. This showed a significant interaction between graph type and area [F (1, 23) = 37.22,p <.0005]. A post-hoc t-test showed that the number of vertical gazes in the legend area were significantly lower for Graph Type 2 [t (23) = 6.56, p<.0005].

4.3 Average Gaze Length

The average gaze length was computed. The lengths of vertical gazes in the legend area for Graph Type 1 are greater than those for the legend area in Graph Type 2. A two-way repeated measures ANOVA (graph type x area of interest) performed on the mean length of vertical gazes showed a significant interaction [F (1, 23) = 17.82,p <.0005]. A post-hoc t-test showed that the length of vertical gazes in the legend area were significantly longer for Graph Type 1 [t (23) = 4.80, p<.0005]. The shorter gaze lengths, when viewing Graph Type 2, may be caused by the poor contrast, remote legend and general graphical clutter.

Possibly as a result of the legend being more remote from the data in Graph Type 2 than Graph Type 1 designs, there was a significant difference in the number, but not length, of horizontal gazes in the legend [t (23) = -2.484, p = .021].

In summary, eye movement data has established a significant difference between the number and length of search efficient vertical gazes in the legend area of graph type 1 and those found in the legend area of graph type 2. It is argued that this indicates the influence of the design features which guided eye movements over the important legend area, in the case of graph type 1 but not in the case of graph type 2 and this is likely to have contributed to the greater usability of the type 1 graph.

5 Conclusion

This experiment has shown that it is possible to use eye tracking in a realistic scenario to provide a detailed assessment of the usability of graphs. The legend and data areas of a graph have been established as being pre-eminent in the comprehension of graphs and their usability. Eye movement data have been found to be consistent with more traditional performance measures and have helped establish that graphs constructed in line with design criteria have a statistically significant advantage over other designs in terms of time taken to complete tasks, accuracy and user appeal.

In addition, a new measure, the “gaze orientation”, has enabled the exploration of some of the influence design features have on eye movements.

Eye tracking evidence indicates that openness (lack of boxed text), alignment and proximity induce eye movement that can be utilised to accelerate the assimilation of graphical content.

We believe the orientation of gazes is influenced by design features. This knowledge has application in many areas such as information visualisations, instructional text and web site design and we intend to test this hypothesis and the relative power of various design features in further experiments in other display genres.

References

Bertin, J. (1983) *Semiology of Graphics,* University of Wisconsin Press, Madison.

Card, K., Mackinlay, J. D. and Shneiderman, B. (1999) Readings in Information Visualization - Using Vision to Think, Morgan Kaufman Publishers Inc.

Cowen, L., Ball, L. J. and Delin, J. (2002) An Eye Movement Analysis of Web Page Usability *People and Computers XVI-Memorable Yet Invisible***,** 317-335.

Goldberg, J. H. and Kotval, X. P. (1999) Computer Interface Evaluation Using Eye Movements: Methods and Constructs *International Journal Of Industrial Ergonomics,* **24,** 631-645.

Kosslyn, S. M. (1994) *Image and Brain: The Resolution of the Imagery Debate,* MIT Press, Cambridge, M.A.

Larkin, J. and Simon, H. A. (1987) Why a diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science,* **11(1),** 65-99.

Raynor, K. and Pollatsek, A. (1994) *The Psychology of Reading,* Lawrence Erlbaum Associates.

Schriver, K. A. (1997) *Dynamics in Document Design,* J.Wiley & Sons, New York.

Tufte, E. R. (1990) *Envisioning Information,* Graphics Press, Cheshire, CT.

Vincow, M. A. and Wickens, C. (1993) Spatial Layout of Displayed Information: Three steps towards developing quantitative Models. *Human Factors Society - Thirty Seventh Annual Meeting***,** 348 - 352.

Ware, C. (2000) *Information Visualization Perception for Design,* Morgan Kaufmann Publishers.

Yarbus, A. L. (1967) *Eye Movements and Vision,* Plenum Press, New York.