Visual search and VDUS

Scott, Derek

VISUAL SEARCH AND VDUS

Derek Scott BSc, MSc (Eng.)
(Graduate Society)

Thesis submitted to the University of Durham in
candidature for the Degree of Doctor of Philosophy
September 1991

Statement of Copyright

Declaration

The work contained in this thesis was carried out by the author between October 1988 and September 1991 whilst a post-graduate student in the Department of Psychology at the University of Durham. None of the work contained in this thesis has been submitted in candidature for any other degree.

Details of Experiments 1 and 2 have previously been published in:
Scott, D. & Findlay, J.M. (1990). The shape of VDUs to come: A visual search study. *Contemporary Ergonomics 1990.* London: Taylor & Francis. pp. 353–358.

Details of Experiments 1, 2 and 3 have also previously been published in:
Scott, D. & Findlay, J.M. (1991). Visual search and VDUs. In: D. Brogan (ed.), *Visual Search II.* London: Taylor & Francis.
and in
Scott, D. & Findlay, J.M. (1991). *Visual Search, Eye Movements, and Displays.* IBM Laboratories (UK), Hursley.

Details of Experiment 4 have previously been published in:
Scott, D. & Findlay, J.M. (1991). Optimum display arrangements for presenting status information. *International Journal of Man-Machine Studies,* **35,** 399–407.

Details of Experiments 4 and 5 have also previously been published in:
Findlay, J.M. & Scott, D. (1991). Visual science and the user interface: What should go where? *Proceedings of the Annual Meeting of the Applied Vision Association,* Manchester: AVA. p. 8.

Details of Experiments 6 and 7 have previously been published in:
Scott, D. & Findlay, J.M. (1991). Role of peripheral information within a VDU text editing task. In: Y. Quiennec & F. Daniellou (eds.), *Designing for Everyone and Everybody.* London: Taylor & Francis. pp. 686–688.

Details of Experiments 8 and 9 have previously been published in:
Scott, D. & Findlay, J.M. (1991). Future displays: A visual search study of computer icons and words. In: E.J. Lovesey (ed.), *Contemporary Ergonomics 1991.* London: Taylor & Francis. pp. 246–251.

Details of Experiment 13 have previously been published in:
Scott, D. & Findlay, J.M. (1991). The role of density in antialiased fonts. *Ophthalmic and Physiological Optics,* **11,** 279.

Abstract

This wide-ranging study explored various parameters of visual search in relation to computer screen displays. Its ultimate goal was to help identify factors which could result in improvements in commercially available displays within the 'real world'. Those improvements are generally reflected in suggestions for enhancing efficiency of locatability of information through an acknowledgement of the visual and cognitive factors involved.

The thesis commenced by introducing an ergonomics approach to the presentation of information on VDUs. Memory load and attention were discussed. In the second chapter, literature on general and theoretical aspects of visual search (with particular regard for VDUs) was reviewed.

As an experimental starting point, three studies were conducted involving locating a target within arrays of varying configurations. A model concerning visual lobes was proposed.

Two text-editing studies were then detailed showing superior user performances where conspicuity and the potential for peripheral vision are enhanced. Relevant eye movement data was combined with a keystroke analysis derived from an automated protocol analyser.

Results of a further search task showed icons to be more quickly located within an array than textual material. Precise scan paths were then recorded and analyses suggested greater systematicity of search strategies for complex items.

This led on to a relatively 'pure' search study involving materials of varying spatial frequencies. Results were discussed in terms of verbal material generally being of higher spatial frequencies and how the ease of resolution and greater cues available in peripheral vision can result in items being accessed more directly.

In the final (relatively applied) study, differences in eye movement indices were found across various fonts used.

One main conclusion was that eye movement monitoring was a valuable technique within the visual search/VDU research area in illuminating precise details of performance which otherwise, at best, could only be inferred.

Acknowledgements

Thanks firstly go to John Findlay for demonstrating an ability to recognise talent and for support and guidance thereafter.

Vicente Ponsoda was a particularly valuable aid with statistical advice and in providing the impetus, rationale, programming and analyses behind the transition matrices studies.

All eye movement monitoring programmes were written by John Findlay. Steve Lavelle modified the screen-editor used in Experiments 6 and 7.

Name checks also go out to our technical staff: Malcolm Rolling for photography and slide productions; Chris Mullaney for the random characters generator programmes used in Experiments 1, 2, 12 and 13, setting up the T-scope projector timer programme, and general technical assistance; Ray Cookson, for mechanical constructions; and Neil Corr and Arthur Perry for electrical support. Particular thanks go to Shirley Whiteley and her steady hand for performing the tricky task of inserting/removing scleral search coils.

Last but not least, I would like to thank all those who participated in these experiments. Of special importance are those who acted as (unpaid) subjects in the rather aversive search coil studies: Andy Calder, John Findlay, Ala Hola, Bob Kentridge, Fiona Newell, Vicente Ponsoda, Robin Walker, and Matthew Wenban-Smith.

During this research, the author was in receipt of a SERC CASE award with IBM (UK) as the collaborating body.

Contents

Chapter One

An Ergonomics Perspective of VDUs

1.1 OVERVIEW

This first chapter introduces the complexities involved in an appraisal of the usability of VDUs and leads into a section on direct manipulation interfaces. This section forms a foundation for the studies on icons, typically a major component of such systems, as detailed in Chapter 6. There follows a discussion of windows, one of the other main aspects of direct manipulation. One of the cardinal characteristics of windows is their facility for changing shape and this points to the subject matter of Chapter 3. There follows a section on reminding and task switching, cognitive concepts which form the kernel of Chapter 4. The scene is then set for Chapters 7 and 8 with an introduction to spatial frequency analysis and antialiased fonts in connection with typography at the interface. This chapter also overviews the broad areas of cognitive ergonomics and presents the underlying general approach within this thesis of information processing. Basic vision ergonomics and vision psychophysics, again general approaches throughout, are presented too. Subsequent sections detail the issues of attention and peripheral vision, most notably major issues of Chapter 5, but also themes running throughout. As a supplementary experimental technique employed for all the studies included within this thesis is eye movement monitoring, there is then a review of the various techniques which may be used with particular emphasis on that employed here; the scleral search coil. Finally, a selection of previous studies relating to eye movements and VDUs is offered, followed by a brief indication of potential areas for exploration.

1.2 PROS AND CONS OF THE VDU

The visual display unit (VDU) is the primary source of feedback to the user from the computer. Classic texts concerning the ergonomics of VDUs, and human-computer interaction (HCI) more generally, include those of Cakir *et al.* (1980), Grandjean and Vigliani (1980), Helander *et al.* (1984), and Shneiderman (1987). Recent years have seen rapid technological developments. For instance, whereas a size greater 24 lines of 80 characters was not long ago unknown to most people, now 66 lines of 166 characters (and greater sizes) are available. Equally, resolution has increased, and although many people are working in the region of 320 × 400 pixels, 1664 × 1200 is now not unusual. The 'high resolution' monitor used in Experiments 12 and 13 of this thesis was of 1024 × 768 pixels. Unfortunately, this rapid advancement has also been, in part, responsible for many problems of HCI; some of these problems being dealt with within this thesis (e.g. antialias fonts and the 'icons with everything' trend.

The VDU has many important features which include: rapid operation (thousands of characters per second or a full image in a few milliseconds; quiet operation; no paper waste; relatively low cost (costs continue to decrease as quality continues to increase); reliability; highlighting, such as overwriting, windowing and blinking; graphics and animation; and special purpose features (Shneiderman, 1987). Health concerns such as visual fatigue (e.g. Miyao *et al.*, 1989), stress (e.g. Kanaya, 1990), and radiation levels (e.g. Rosner & Belkin, 1989) are being addressed by both manufacturers and government agencies, although some concerns do remain.

Monochrome displays Monochrome displays are perfectly adequate for many applications and in fact may be preferred, particularly if the monochrome display has a higher resolution than the colour display. Several different technologies are available, and below are described just three of the selection mentioned by Shneiderman (1987).

(1) **Raster scan cathode ray tube (CRT).** This, the most popular form of VDU, is similar to a television monitor and has an electron beam sweeping out lines of dots to form alphanumerals. Refresh rates (the reciprocol of the time required to produce a full screen image) vary from 30 to 65 per second. Higher rates reduce flicker and so are preferred. CRT displays are often green because the P39 green phosphor has a long decay time, which allows relatively stable images. The decay time of the P38 amber-orange phosphor is even longer. Another property of a phosphor is its low 'bloom' level, which allows sharp images because the small granules of the phosphor do not spread the glow to nearby points. The maximum resolution of a CRT is in the region of 100 lines per inch. Displays can have light letters against a dark background (negative contrast) or dark letters against a light background (positive contrast). CRT sizes, which are measured diagonally range from less than two inches to almost thirty inches, although typically they are within the nine to fifteen inch range. Screen size and shape are discussed more fully within Chapter 3.

(2) **Plasma panels.** Rows of horizontal wires are slightly seperated from vertical wires by a small glass-enclosed capsule of neon-based gases. When the horizontal and vertical wires on either side of the capsule receive a high voltage, the gas glows. These plasma displays are usually orange and flicker-free, although the size of the capsules limits resolution.

(3) **Liquid crystal displays (LCDs).** The reflectivity of tiny capsules of liquid crystals is influence by voltage changes, and this turns some spots darker when viewed by reflected light. LCDs are free from flicker, although the size of the capsules again limits the resolution. Smaller computers and calculators often operate with LCDs because of their light weight and low power consumption. Portable LCD computers are available with 24 lines by 80 characters.

The technology chosen affects numerous variables such as size, refresh rate, capacity to show animation, resolution, surface flatness, surface glare from reflected light, contrast between characters and background, brightness, flicker, line sharpness, character formation, and tolerance for vibration. Each display technology has advantages and disadvantages with respect to these variables. Shneiderman (1983) also notes that further consideration should be paid to the availability of features such as the following: user control of contrast and brightness; software highlighting of characters by brightness; underscoring; reverse video; character set (alphabetic, numeric, special and foreign characters); multiple fonts; multiple font sizes; shape, size, and blinking rate of the cursor; user control of cursor variables; blinking (possibly at several rates); scrolling mechanism (smooth scrolling is preferred); user control of number of lines or characters per line displayed; and support of negative and positive polarity.

1.3 DIRECT MANIPULATION

Since about 1981 a new class of user interface has emerged based on a high resolution flexible display screen and a pointing device with which the user, *ideally with ease*, points out the objects to be acted upon. Although the impressive display screens and direct methods of pointing out objects to be manipulated do greatly enhance the interaction, many of the present systems share certain problems: (a) awkwardness in handling structures that are not visible, (b) difficulty in user programming, and (c) a tendency to over-use menus (Roberts, 1984). Responding to these problems in a way that retains the attractiveness of these systems will be the user-interface designers' challenge in the immediate future. The style of user interface which is being built for these machines has been termed the 'direct-manipulation' interface (Shneiderman, 1983). What this implies is that users feel more as if they are touching the system's objects and doing things to them directly, rather than verbally instructing an intermediary to perform the operations for them.

Whilst direct manipulation interfaces deal with the obvious 'outward' or *peripheral* problems of users, important cognitive or *central* aspects of the user such as expectations, motivations and their models of operation must not be ignored. As Long (1989) has remarked, the discipline of HCI consists of two main sub-disciplines. One, software engineering (and computer science), primarily concerns the computer side of the interaction. Software engineering aims to develop specifications for both hardware and software aspects of the computer which can be implemented as an interaction. The second, ergonomics (or human factors), essentially concerns the human side of the interaction. Both sub-disciplines aim to support the optimisation of the interaction for effectiveness. Cognitive ergonomics is conceptualised as concerned with mental aspects of the interaction and thereby with developing specifications of the knowledge required by the human to interact with the computer to perform work effectively.

> To support the specification of knowledge required by humans to use computers, the discipline of cognitive ergonomics itself needs to acquire and apply principles concerning that knowledge. Knowledge involved in computer-based work can be conceived as consisting of representations and processes: representations of the work to be performed, of the computer, etc; and processes required to use the representations to perform the work, to use the computer, etc. The processes include the acquisition of the representations, their transformation and their expression in the form of behaviour. Cognitive ergonomics, then, seeks to support the specification of knowledge conceived as the representations and the processes required by the humans to interact with the computer to perform work effectively. Cognitive compatibility between the representations and the processes required by the interaction and those currently possessed or acquirable by the human has been advanced as one principle on which specifications can be based (Long, 1987). (Long, 1989, pp. 5–6)

1.3.1 General philosophy of direct manipulation interfaces

Prior to direct manipulation, the user of a computer had only a very restricted way of interacting with the computer. A language often had to be learned, although some databases could be interrogated with a minimum of learning by use of a database query language. The language necessary may have been a rather low level one whereby the user had to learn a language that was close to the computer's way of interpreting information.

Shneiderman (1983) has presented an 'integrated portrait' of the features provided by direct manipulation and summarised these four principles:

(1) Continuous representation of the object of interest. 'WIMP-talk' refers to things that appear on the screen as 'objects'.
(2) Physical actions (movement and selection using the mouse) instead of complex syntax.
(3) Rapid, incremental, reversible operations whose impact on the object of interest is immediately visible.
(4) A layered approach to learning that permits usage with minimal knowledge. Users are encouraged to 'expand their knowledge gracefully'.

1.3.2 Hidden structure

Some direct manipulation systems are often described as 'What-you-see-is-what-you-get' or WYSIWYG systems. For systems where the primary emphasis is on creating something that will be printed, the screen will display as far as possible exactly how the hardcopy will appear. Direct manipulation is concerned with the look of the result. Since the user and the computer communicate mainly in terms of appearance, there is much less emphasis on structure. When the user positions one object near another, the system has no way of discerning the intent behind the move. Are the two objects intended to remain near each other in later manipulations? Do they remain a fixed distance apart or is their distance relative to something else. With systems such as TEX (the electronic 'mark-up' type-setting programme used to print this thesis), the user has no other choice but to instruct the system in terms of structure, so such problems never arise (Roberts, 1984).

Once again, however, whilst discussing 'hidden structure' in terms of the machine's understanding of the operator's wishes, the obverse must not be forgotten. The more *central* or cognitive issues involved are well described in chapters such as those of Barnard *et al.* (1989), Long (1989) and Whitefield (1989).

1.3.3 User programmability

A convenient feature of interacting with a system through typed command languages is that such languages are easily converted to programming languages. If the user wishes to invent a new command which performs the action of several old commands, the programme can be written in a language with which the user is familiar. However, if a language is based on pointing to objects and on pointing to (possibly graphic) indications of what is to be done to the object, it will be more difficult to programme in. Without a user programming language, Roberts (1984) points out, the user cannot tailor the environment to his own tasks to the extent that many users would prefer. A 'programming-by-example' scheme in which the system converts user actions to a conventional programming language is a promising solution to this (Halbert, 1984).

1.3.4 Arrays of choices

People who create direct manipulation systems often apply other psychological principles to their system design, with the result that direct manipulation systems often share features that are not direct manipulation as such. One such cognitive principle would be 'Recognition is easier than recall' (Anderson, 1980), or alternatively 'seeing and pointing is better than remembering and typing'. This has led to designs which include menus for commands, property sheets for specifying objects' attributes, and visible 'folders' for filing. This sort of menu can lead to several problems:

(a) they can overwhelm novices with their large array of choices;

(b) progressive-disclosure schemes which attempt to minimise the overwhelming aspects sometimes make a multi-step process out of what is simply one operation to the user; and

(c) they require waiting for a display of information and pointing to a place which might be far from the user's primary focus of attention.

For an expert, such a style of interaction may be considerably slower than merely typing a few remembered characters. 'Seeing and pointing' interfaces must be carefully designed so that they are either extremely efficient in their context, or provide bypasses for users who will not require their full richness. Roberts (1984) concluded that these direct manipulation systems have not yet matured. Although this was stated several years ago, considerably more work is still required to solve certain problems concerning how to fully maximise their advantages. Again, here lies the challenge for interface designers.

1.3.5 A cognitive account of directness

The concept of direct manipulation holds great promise, but as yet there is no full explanation of how the particular properties produce the feeling of directness. At the heart of the cognitive ergonomics approach is the assumption that the feelings of directness result from the use of fewer cognitive resources. Conversely, the need to commit additional cognitive resources in the use of an interface leads to the feeling of indirectness. The sensation of directness is always relative and is often due to the interaction of a number of factors. Costs are associated with every factor that increases the sensation of directness. There is no way at present to measure the tradeoff values, but Hutchins *et al.* (1986) have attempted to provide a framework within which one can say what is being traded off against what.

Two seperate and distinct aspects of the feeling of *directness* are *distance* and *engagement.* The former involves the notion of a distance between one's thoughts and the physical requirements of the computer system. A short distance implies that the transition is simple and straightforward. Thoughts are readily translated into the physical actions required by the system and the system output is in a form readily interpreted in terms of the goals of interest to the user. the term 'distance' emphasises that directness is never a property of the interface alone, but involves a relationship between the task the user has in mind and the way that task can be accomplished via the interface. The critical issues involve minimising the effort required to bridge the gulf between the user's goals and the way they must be specified to the system. These are Norman's (1986) *Gulf of Execution* and the *Gulf of Evaluation* (Fig. 1.1). The Gulf of Execution is bridged from the psychology side by the user's formation of intentions relevant to the system and the determination of an action sequence. It is bridged from the system side when the designer of the system builds the input characteristics of the interface.

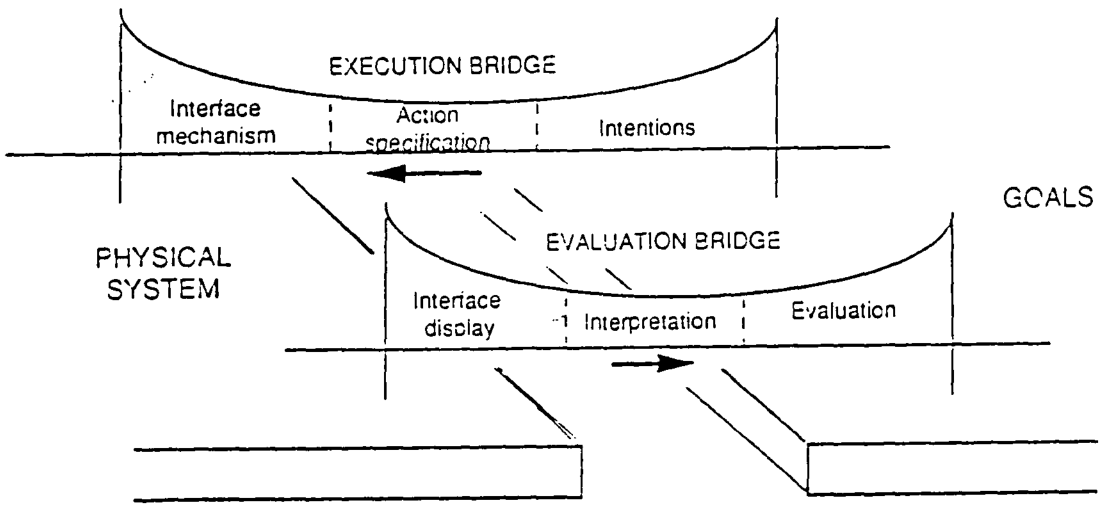


Figure 1.1: The Gulfs of Execution and Evaluation. (Adapted from Norman, D.A. (1986) Cognitive engineering. In Norman, D.A. & Draper, S.W. (eds.), *User Centred System Design,* Lawrence Erlbaum.

Hutchins *et al.* suggest that the feeling of directness is inversely proportional to the amount of cognitive effort it takes to manipulate and evaluate a system and, moreover, that cognitive effort is a direct result of the Gulfs of Execution and Evaluation. The better the system interface is in bridging the gulfs, the less is the cognitive effort needed, and the more direct is the feeling of interaction.

Norman (1986) suggests that it is useful to consider an approximate theory of action, and particularly that it is useful to seperately consider the seven different stages of activity depicted in Figure 1.2 that a person might go through in the performance of a task. The primary, central stage is the establishment of the goal. Then, to carry out an action requires three stages: forming the intention, specifying the action sequence, and executing the action. To assess the effect of the action also requires three stages, each in some sense complimentary to the three stages of carrying out the action: perceiving the system state, interpreting the state, and evaluating the interpreted state with respect to the original goals and intentions. The important point is that the interface must attempt to match the goals of the person. The two gulfs are bridged from two directions; from the system side by the system interface and from the user side by mental structures and interpretation. The bridges from the user's side are aided through the development of appropriate conceptual models; the design model and the user model (see Norman, 1986, for a fuller discussion), which in turn are aided by the system image presented by the interface itself. The more the gulf is spanned by the interface, the less is the distance needed to be bridged by the efforts of the user.

Direct engagement The other aspect of directness concerns the qualitative feeling of engagement, the feeling that one is directly manipulating the object of interest. There are two major metaphores for the nature of HCI; a conversation metaphore and a model world metaphore. In a system built on the conversation metaphore, the interface is a language medium in which the user and system have a conversation about an assumed, but not explicitly represented, world. Here, the interface is an implied intermediary between the user and the world about which things are said. In a system built on the model world metaphore, the interface is itself a world where the user may act, and which changes its state in response to user actions. The world of interest is explicitly represented and there is no intermediary between user and world. Appropriate use of the model world metaphore can create the sensation in the user of acting upon objects of the task domain themselves. Hutchins *et al.* call this aspect *direct engagement.*

This description of the nature of interaction begins to suggest how to make a system easier to use, but it misses the essence of direct manipulation. The analysis of the execution and evaluation process explains why there is difficulty in using a system and it does say something about what is required to minimise the mental effort needed to use a system. However, there is more to it than that. The systems that best exemplify direct manipulation all convey the subjective impression that we are directly engaged with the semantic objects of our goals and intentions. This feeling of first-personness or of direct engagement with the objects that concern us has been discussed by Laurel (1986). If one is analysing data, for instance, then the data themselves should be manipulated.

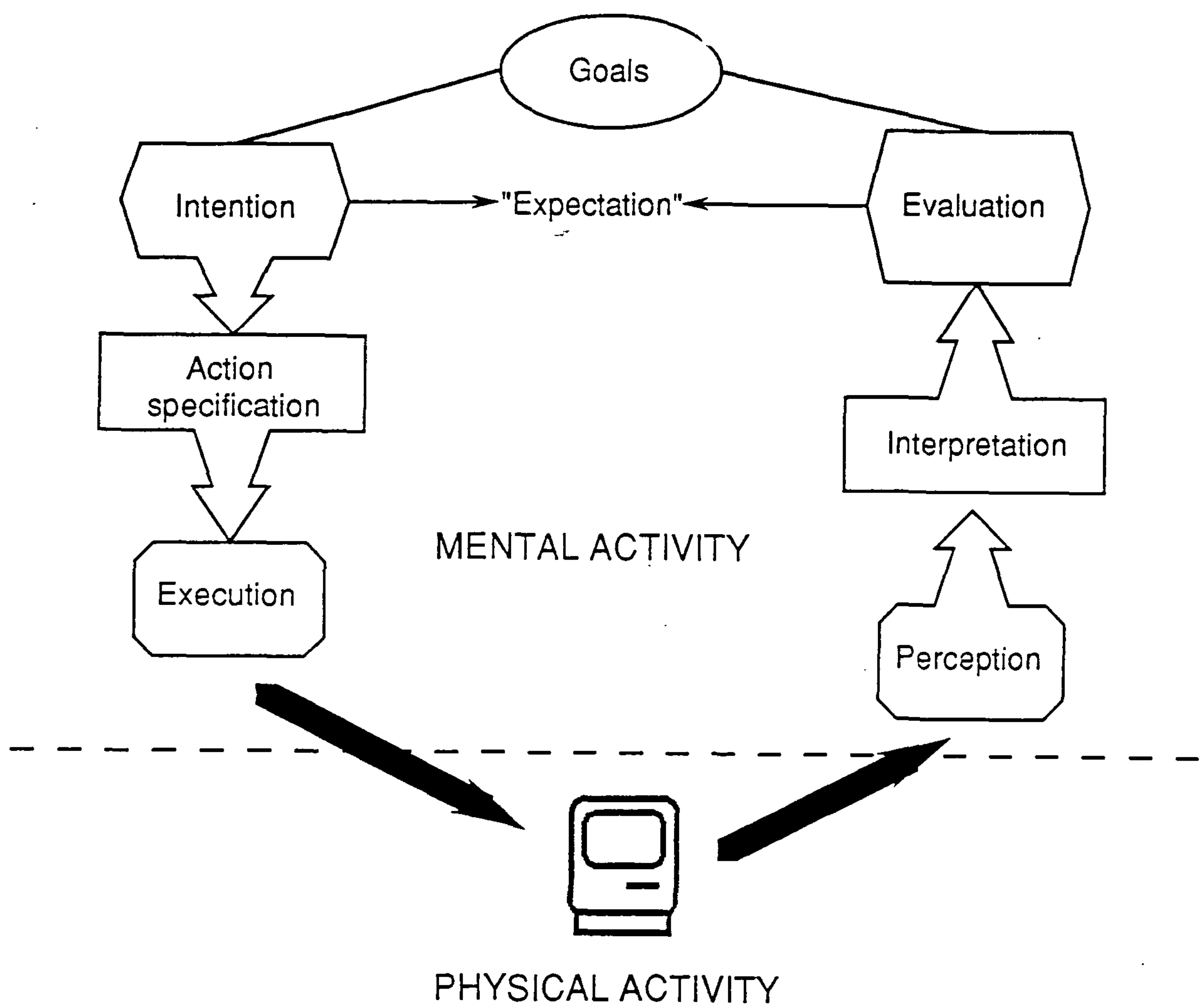


Figure 1.2: The seven stages of user activities involved in the performance of a task. (Adapted from Norman, D.A. (1986) Cognitive engineering. In Norman, D.A. & Draper, S.W. (eds.), *User Centred System Design,* Lawrence Erlbaum.

Interfaces have traditionally been designed around the conversational metaphore. Although there is power in the abstractions that language provides, the implicit role of interface as an intermediary to a hidden world denies the user direct engagement with the objects of interest. Rather, the user is in direct contact with linguistic structures. These structures can be interpreted as referring to the objects of interest, but are not those objects themselves. Making the central metaphore of the interface that of the model world supports the sensation of directness. Instead of describing the actions of interest, the user performs those actions. In the conventional interface, the system describes the results of the actions. In the model world, the system would present directly the actions taken upon the objects (Hutchins *et al.*, 1986).

The acronym WIMPs for the cardinal features of direct manipulation techniques, Windows, Icons, Menus and Pointing devices, has been around for some years although it is difficult to pinpoint its precise etymology. Although it is the interworking of the WIMP features that bestows the interface with its directly manipulable nature, it is valuable to consider each WIMP feature separately, or as separately as possible. Here, the concept of windows is introduced, and which is further discussed in Chapter 3. The role of icons is largely left for Chapter 6, where they are discussed in terms of their ability to improve visual search. The reader will find implications for menu arrangements scattered throughout this thesis. Due to the limited reliance on visual factors, little is said of pointing devices.

1.4 WINDOWS

A window is a region on a display surface whose size, position and display priority relative to other windows may be changed at will by the operator. Also, a window may be regarded as the facility whereby the display screen can be divided into smaller subscreens or windows each of which can display different objects. The object might be a spreadsheet, a textfile, a drawing, etc. Windows allow the user to interact with multiple sources of information at the same time, as if the user had a set of VDUs of different sizes at his work station. It is this multiplicity of contexts, together with the graphical features of the display and the ability to use graphical input devices that is the source of the power of the new interfaces. The basic idea behind windows is straightforward. A window is simply a particular view of some data object in the computer. Windows can depict data structures using different textual and graphical idioms and these can co-exist together on the user's display, popping into and out of existence, moving about and changing shape and size according to the needs of the moment (Card *et al.*, 1985).

In principle, each window can support its own individual human-computer dialogue, thus allowing for the interweaving of multiple concurrent activities at a single workstation. There are usually limits to how many windows can best be displayed, the limits being both technical (has the system enough memory to remember the contents and run the applications in all the windows?) and pragmatic (given a finite screen size, how many windows can feasibly be displayed and still be useful to the user?). The items inside a window may be displayed as various lists, or as icons.

1.4.1 Why have multiple windows?

On the whole, the potential benefit of displaying multiple windows on the same screen is to reduce the load placed on the cognitive abilities of users, particularly—although by no means exclusively—that on short-term memory (STM). Reducing this load would free cognitive resources for other tasks. Basically, this is achieved by the possibility of syncronous presentations of information. Specifically, the following categories of tasks stand to benefit from multiple windows (Bury *et al.*, 1985):

(1) **The display of supplemental information relevant to the user's primary task:** To take the example of on-line assistance or 'Help', in many systems, when Help is invoked the primary screen is cleared and replaced entirely with a panel containing the Help information. Whilst in Help, users are forced to remember the situation requiring assistance (which is no longer displayed). They then must remember any assistance information when they exit Help and return to the primary task. Alternatively, Help can be displayed in a window which overlays a portion of the primary screen. Users may then continue to view and interact with their primary task while viewing the required supplementary information in Help. Help is merely one example. Supplemental information could come from almost any source in the system environment (e.g. addresses from a directory application, a paragraph from a text file).

(2) **Monitoring of changes in a secondary window:** Some systems, for example, allow users to modify a data set in one window and then have the result immediately reflected on a graph displayed by another application in a secondary window. In this case, it is the user who is causing the changes in the secondary window. Changes might also be caused by external events. For instance, monitoring of stock prices could be done whilst performing a secondary task.

(3) **Aid in tasks in which specific locations within multiple applications need to be specified:** Typical examples of this are the Move and Copy functions between windows. Here a user may identify a block of data in one window (by highlighting or some other method), and then select a specific location in another window to which the data are to be moved or copied.

Windows act as visible and substantive place holders for multiple applications on the screen. In other words, they provide users with a three-dimensional (width, height, and depth) physical representation of the task environment. Most terminal systems require users to pass sequentially from one application to another. As each new application is accessed, the old one is erased from the screen, thus requiring users to remember any needed information. The availability of a window manager allows users to organise the presentation of information on the screen to meet their particular needs. The ability to overlap windows has been called the 'messy desk' model of window management (Bury *et al.*, 1985).

1.4.2 Window-management tradeoffs

Although the potential benefits of window management are attractive, There are tradeoffs. Whereas traditional interactive computing has been largely a serial task for users, with a window-management system, users may have many things displayed at once. There are pros and cons to this. Most VDUs are 80 columns wide, and have between 24 and 34 rows. This size has emerged as a compromise between cost and usability, but even the full size of such displays is inadequate for many applications. Each new window added to a display reduces the effective (non-overlapped) space available to the other windows. If the system is such that each window has a wide border which contains icons, or other command/status information, the screen soon becomes crowded with borders with little effective work space showing. The smaller the window becomes, the more scrolling and other processes the user must perform. At some point, any benefits derived from the window manager are swamped by the extra work which is imposed. Some systems have an oversized display which helps reduce the problem. As Bury *et al.* (1985) note, a compounding problem is the fact that most window-management systems in existence today allow only one window to be active at any one time. Having only one windowed application active at any given time requires users to do far more alternating between windows, and consequently expend greater cognitive effort.

1.5 TASK SWITCHING

Most user interfaces are designed to help the user perform particular tasks through to completion but users actually switch back and forth among several concurrent tasks. Task switching can lead to major difficulties unless there is special interface support. Consequently, the state of tasks not on the VDU is difficult to remember and the user may be forced into inefficient activities such as writing information from one task on paper, and then typing it into another task. Conversely, window-oriented systems allow the user to see information from several tasks, yet severe conflicts among tasks for the use of screen space may result in high overheads as users must move, reshape, or scroll windows or shrink and expand icons.

Bannon *et al.* (1983) identified a number of reasons why users switch from one task to another:

(a) digressing to do tasks that users are reminded of while performing another task;
(b) timesharing among concurrent demands;
(c) tasks with long waits;
(d) subtasks; and
(e) 'snags' such as running out of file space.

More could be added. The important point is that interruptions and other sorts of task switching are an important aspect of user activity. Task switching occurs for activities measured over minutes where task switching time and resumption are especially important, and it occurs for activity measured over days, where the memorial aspects of remembering the activities and their state are particularly important.

Bannon *et al.* also suggested six issues that an interface to support task switching should involve:

(a) reducing mental load when switching tasks;
(b) suspending and resuming activities;
(c) maintaining records of activities;
(d) functional grouping of activities;
(e) multiple perspectives on the work environment; and
(f) interdependencies among items in different workspaces.

1.5.1 Reminding

Users engaged in many different activities will be aided by assistance in keeping track of those activities. When part of the cognitive burden of monitoring activities is removed, users are able to concentrate better on actually performing the tasks rather than use mental effort in remembering and scheduling those tasks. Windows (and icons) provide one way to remind users of *interrupted* activities. They can be effective to the point where a jumble of overlapping windowed activities detrimentally clutters the screen. It may then prove more effective to replace the arrangement with a more concise list of titles or descriptors of the activities. 'Reminding' and 'interrupting' are issues returned to in §4.2.

Besides clutter, another way that visual reminders lose their effectiveness is when they are not easily associated with their activities. Activities do not always neatly correspond one-to-one with windows. When there is no visible unit on the screen that is uniquely associated with an activity, there will be no direct way to be reminded of that activity. For example, the display of a calendar programme may not serve as a useful reminder for an activity that the user conceives of primarily as sending a message. Similarly, if a window is being used by three different activities at once, it cannot effectively remind the user of all of those activities. As Cypher (1986) concluded, windows are not the whole solution.

1.5.2 The display as external memory

Analysis of windows also critically depends on an understanding of the user and how he is linked to the display. By necessity, an understanding of the user must amalgamate cognitive considerations of a user's goals and problem solving on the one hand and on the other, the perceptual and visual science issues of how he moves his eyes and what cues from the display affect his ability to locate and discover information needed for cognitive processing.

Card's analysis of the effect of the human processsor on window use begins with the proposition: *A fundamental constraint on user's cognitive performance arises from limitations of working memory.* Ability to perform mental arithmetic is largely limited by difficulties in keeping track of the intermediate products and keeping one's place. In doing financial analyses of a company, debugging a programme or writing a paper, limitations of the number of mental items that can be 'kept track of' produce strict constraints on human cognitive capabilities.

The computer display provides the user with an external memory which is an extension of his own internal memory and with which he is able to mentally operate on more information than without. Notes on paper scattered around the desk only partially serve the same function because these are not as dynamic as something which, for example, can automatically graph in one window the data presented in another. But the computer display is not only an external memory which extends the user's internal memory. It is also (partially) shared with the computer itself, another active agent. Consequently, it is at the same time both memory and communication medium. The full power of the display as external memory and communications medium and also the synergistic interaction between these are only realised when the display supports independent, but related, notes on objects of memory and communication. For this reason the windowing technique with its emphasis on seperate communications contexts, has become the harbinger of dramatic improvements in HCI.

1.6 TYPOGRAPHY AT THE INTERFACE

The basic element of graphic design is *typography.* The craft of typography starts with the design of attractive and legible *letterforms* in a variety of *typefaces* or *fonts,* that may be either *serif* or *sans-serif,* and that are arranged in families encompassing variations in *weight,* such as *regular* and *bold,* and variations in slant, such as *roman* and *italic.* Typography includes the careful arrangement of sequences of letterforms on the 'page' with the goal of enhancing readability. This can be achieved by controlling parameters of individual characters, such as *point size* and *letterspacing,* of words, such as *word spacing,* and of lines, such as *line length, leading,* and *justification.* Typography also encompasses the augmentation of raw text with simple graphic elements such as *rules, leaderlines,* and *logotypes* (Baecker & Buxton, 1987).

Special fonts are effective for emphasis, but not for sustained reading. Eyes would tire if long portions of this thesis were entirely set in bold or italic face. Therefore, roman type accounts for the vast majority of text, whether on a VDU or on hardcopy. Typography has always been relevant to screen design, even in the days of single font fixed-width 24 × 80 character displays, but with the advent of high resolution bit-mapped displays (particularly when coupled to laser printers), its importance is greater.

Bigelow and Day (1983) introduce some of the issues of digital typography, such as fonts and spatial frequencies and the problem of *aliasing,* which is borrowed here and outlined below.

1.6.1 Fonts and spatial frequencies

A digital computer is required to control the on/off pattern of the CRT raster beam, and the type itself is digital because it is comprised of discrete elements. These elements can be line strokes, pixels, colours, shades of grey or any other graphic unit from which a letterform can be constructed. However, the traditional letter is analogue rather than digital. Its final form varies smoothly with the continuous variation of some process used in its creation, such as the pressure of a brush on paper. One of the many effects of computer development is that in the past 20 years typography has been seeing a replacement of analogue text by digital text.

There are substantial advantages with digital typography. For instance, once letterforms are represented as discrete elements they can be efficiently encoded as discrete and distinguishable physical properties in any convenient medium, processed as bits of information by a computer, transmitted over great distance as pulses of current and decoded to reconstitute the letterforms for the individual receiving the message. In fact, once type is digitised it is effectively encoded in the binary language of the computer, and therefore the size, shape and subtler characteristics of letters can be readily modified by a computer programme. Unlike analogue information, digital information is highly resistant to noise or degradation introduced during the transmission of a signal. The digital receiving device need only distinguish between two states of the signal (on or off) in order to decode and recover the information originally transmitted.

> Spatial frequency analysis can be carried out for any two-dimensional image, such as a letter, just as the graph of almost any mathematical function can be approximated by a sum of sines and cosines. A sharp-edged rectangle, for example, can be represented as a straight line parallel to the horizontal axis of the graph; the density of grey or black in the rectangle corresponds to the height of the line above the horizontal axis. In order to approximate the rectangle a unique set of sine and cosine waves can be superposed. The first sine wave in the approximation is shown in [Fig. 1.3]*a*. The spatial wave, or variation from white through shades of gray and back to white, that corresponds to the sine wave is a rather crude approximation to the rectangle. The more sine and cosine components there are in the approximation, however, the better it becomes: the five colored sine waves of various heights and wavelengths are added together along each vertical line to give the smooth red curve in *b*. In order to eliminate high-frequency noise any shape can be electronically filtered to remove components above a certain frequency. In *c* the two highest-frequency components of the red curve in *b* have been eliminated, namely the yellow and [orange] sine waves, and the result is the red curve in *c*. The filtered curve can be sampled by finding its height, or gray-scale value, at evenly spaced intervals along the horizontal axis (*d*). A digital approximation to the filtered curve is made by assigning a discrete shade of gray across an interval that corresponds to each sample (*e*). It can be mathematically proved that the original filtered curve can be completely reconstructed from the samples by interpolating between the sample points, provided the frequency at which the samples are taken is greater than twice that of the highest-frequency component of the filtered curve (*f*). (Bigelow & Day, 1983, p. 113)

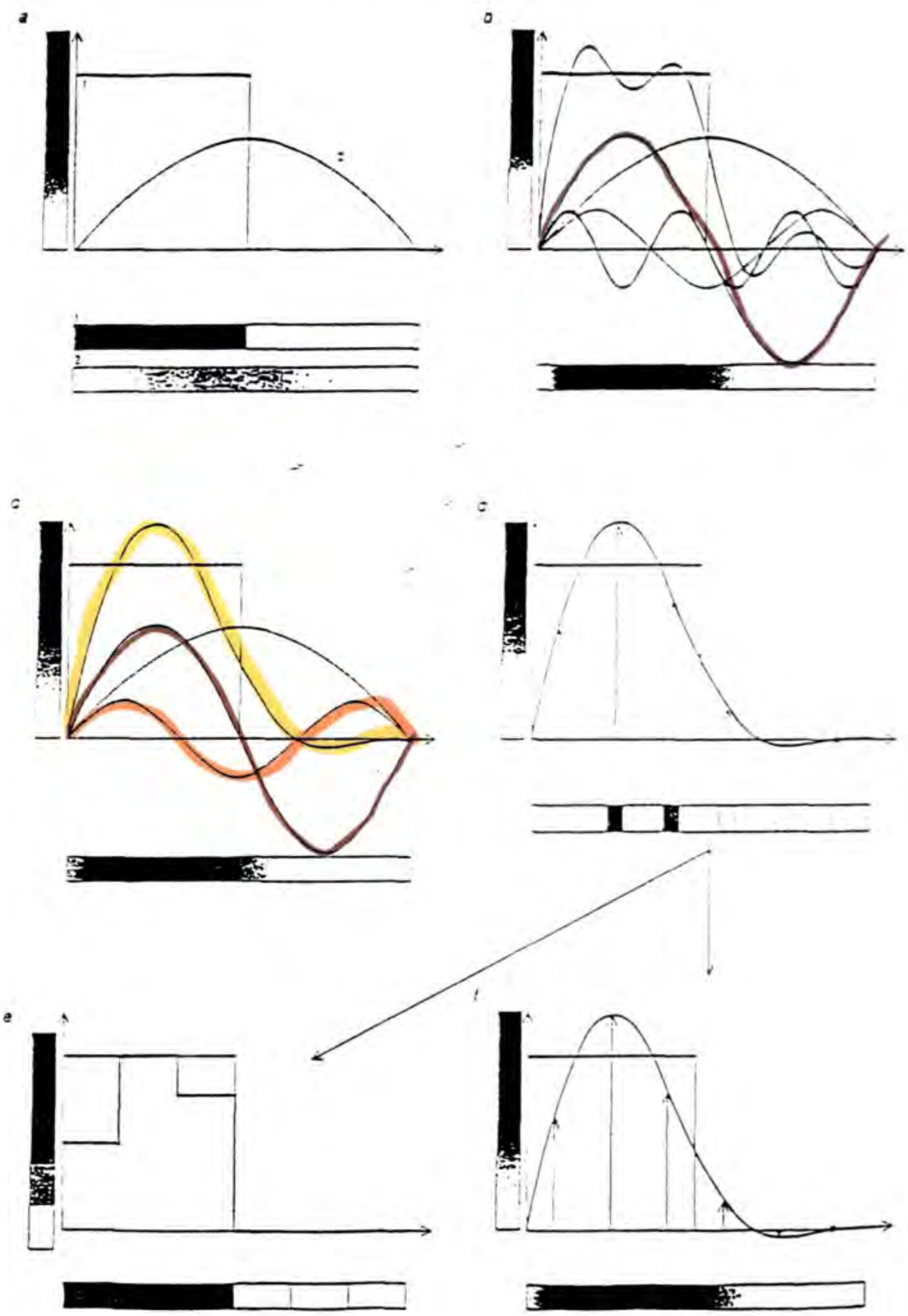


Figure 1.3: A spatial frequency analysis of a two-dimensional image (e.g. a letter).

Bigelow and Day continue that throughout the history of writing, economic pressures have opposed the pressures for more readable and more aesthetically pleasing typography. Need must, of course, be attached to purpose. Today, inexpensive dot-matrix printout is less readable than the costly typography in mass-marketing advertising. The effects on the reader of complex variations in letter design have not yet been quantified, but the response of the visual system to simple spatial variations in light intensity has been studied for more than a quarter of a century. The simplest variation to analyse is a visual analogue of an acoustically pure tone, whose periodic variation of intensity with time can be plotted as a sine or cosine wave. Bigelow and Day give the description that a train of spatial sine or cosine waves can be visualised as a ribbon compressed along its length so that its edge traces a series of ordinary sine or cosine waves. If the top side of the ribbon were shaded in such a way that the crests of the ribbon were black, the troughs remained white and the intermediary sections were various shades of grey, then the blurred parallel bands, or grating, of light and dark that could be seen from above the ribbon would form a spatial sinusoidal wave train.

The amplitude of a spatial sine or cosine wave is the maximum contrast, or deviation from neutral grey, that is found in the lightest or darkest parts of the wave train, and the frequency is the number of variations from light to dark and back again within a given distance. The ability of the visual system to distinguish sinusoidal bands of various contrasts and frequencies from a uniformly grey field have been measured by psychophysicists. It has been found that the sensitivity to spatial variation of light and dark depends on the frequency of the variation; the sensitivity is greatest when the spatial frequency is about three cycles per degree of visual angle, and no contrast, no matter how strong, can be perceived under most conditions when the frequency is greater than 60 cycles per degree.

The importance of these findings for reading are that although the spatial variation from black to white in letterforms is not sinusoidal, fundamental rhythmic patterns in the letterforms are apparent. As an example, as the lowercase letter n is scanned from left to right across its *middle,* the brightness of the image varies rather smoothly from light to the dark stem of the first vertical stroke, then light again in the interior of the letterform, then dark on the second vertical stroke and then finally light again to the right of the letter. Of course, this does not apply to the top of the letter *n;* there is a need for a form of averaging out over the entire letter when talking generally of spatial frequency compositions of alphanumerals. The overall frequency must in some way be taken as the sum of intensities per finite element.

As reading is often carried out under sub-optimal lighting conditions, it might be anticipated that the fundamental frequency of text letterforms has evolved to match the peak contrast sensitivity of the visual system.

> Such a match is almost exactly what is found in the typesetting of English. In close reading the image given the most attention is projected onto the fovea, the most sensitive area of the retina. The fovea subtends an angle of one or two degrees, and at a reading distance of 12 inches the subtended angle corresponds to a linear distance only slightly more than the length of a five-letter word set in 10-point type, the commonest type size. There are on the average about 10 spatial cycles across a five-letter word, and so the spatial frequency of the image received at the fovea is about five to 10 cycles per degree, only slightly higher than the most contrast-sensitive frequency of the visual system. There is good evidence that in fast reading word groups longer than five letters can be read in one eye fixation. The reading image is then partially projected onto the parafovea, the region of the retina surrounding the fovea. (Bigelow & Day, 1983, p. 113–114)

This notion is connected to that of 'visual lobes', which is returned to in §2.3.

So far, it has been implied that the major patterns of black and white that comprise letterforms are sinusoidal spatial waves. But, the sine-wave model can support a more detailed analysis. Almost any form can be analysed as a combination of many spatial sine and cosine waves. To reiterate Fourier's theorem, by superimposing one-dimensional sine and cosine waves of various phases, amplitudes and frequencies the graph of almost any function can be approximated to any desired degree of accuracy. Curves, such as those shown in Figure 1.4 which represent luminance profiles, can be described as the sum of a number of sine waves of varying frequency and amplitude. This figure depicts luminance profiles with the same average luminance of 15 units, indicated by the horizontal lines. Profiles *a*, *b*, and *c* are sine waves. Sine waves can possess high or low spatial frequency depending on the number of cycles of the wave per degree of visual angle. Profile *a* has an amplitude of 10 luminance units and a frequency of 1 cycle per degree. Profiles *b* and *c* possess amplitudes of 3.4 and 2.0 luminance units and frequencies of 3 and 5 cycles per degree respectively.

Summing the two profiles constructs a new profile whose luminance value at each point in space is the sum of the luminance of each component profile at that point. The sum of *a* and *b* is depicted as profile *d*, and if profile *c* is now added to *d*, profile *e* is created. It should be noted that whilst *a* is the shape of a sine wave, *d* has a sharper increase and decrease in luminance, and *e* begins to appear square-wave. Fourier demonstrated that a square wave of a given frequency and amplitude could be constructed by summing an infinite number of sine waves whose frequencies were the odd multiples (1, 3, 5, 7...) of the frequency of the square wave, and whose amplitudes are the odd fractions (1, $\frac{1}{3}$, $\frac{1}{5}$, $\frac{1}{7}$, ...) of the frequency of the square wave. In Figure 1.4, frequencies of 1, 3, and 5 were employed with amplitudes of 10, 3.3, and 2. If the process of adding three or four more sine waves had been continued, the profile would become increasingly more square. Theoretically, to result in a perfectly square profile as seen in *f*, the sum must include all the component profiles up to infinity.

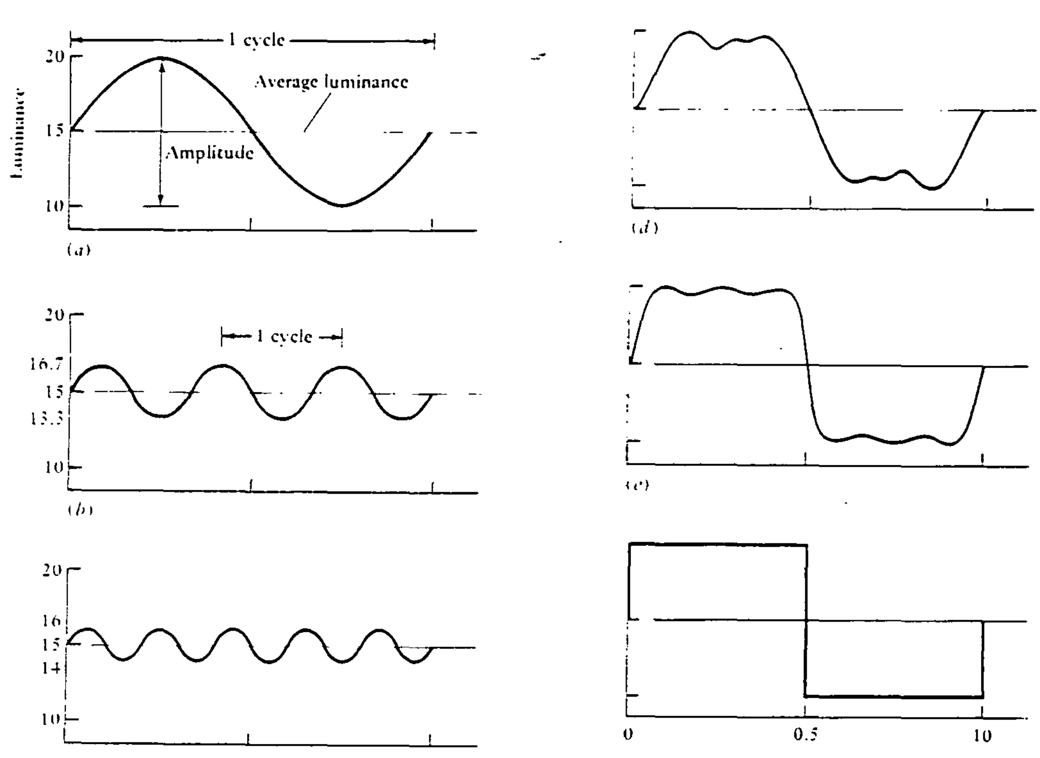


Figure 1.4: Several luminance profiles. From Haber & Hershenson (1980), *The Psychology of Visual Perception.* (Reproduction permission applied for)

> What has been illustrated is how to construct a square profile out of a number of *Fourier component* profiles, each the shape of a sine wave. Mathematically, this is equivalent to the statement that a square wave of a given frequency and amplitude contains within it a number of component sine waves of specifiable frequencies and amplitudes. Thus, we can describe the shape of a square wave by simply listing the frequencies and amplitudes of its component sine waves. That list defines its shape... We can make one final observation from these mathematics: The steeper the luminance profile, measured as rate of change, the more higher-frequency sine components are contained within it. The square wave has an infinite number of higher frequencies, the ramp relatively few higher frequencies, and the gradual profile has only one as it itself is a sine wave. In general, sharp edges have lots of high-frequency Fourier components whereas gradual changes in luminance can be described completely by their low-frequency components. (Haber & Hershenson, 1980, p. 15)

In two dimensions, the spatial sinusoidal components have the additional degree of freedom of their orientation with regards to some fixed direction in the plane. It is possible by Fourier's theorem to to reconstruct any pattern or letterform to any desired accuracy by superposing spatial sine and cosine waves of the appropriate phase, amplitude, frequency and orientation. Generally, the high spatial frequencies of a letter image correspond to its edges and to such details as fine serifes and the tapering of the main strokes of the letter. The low spatial frequencies define the fundamental rhythm of the letter; i.e. its overall pattern of light and dark. The spectrum of spatial frequencies required to represent a letter is termed the 'bandwidth' of the letter. Solely based on the maximum frequency at which the eye can detect contrast, (i.e. 60 cycles per degree of visual angle) it may appear that completely adequate typographic reproduction could be achieved if all the spatial frequencies in a letterform over 60 cycles per degree were removed.

1.6.2 Aliasing and antialiasing

To represent its full spectral bandwidth, a perfectly sharp edge would require an infinite number of sines and cosines. If the bandwidth of a letter is limited in such a way as to include no sinusoidal component above a certain frequency, its edges must be slightly blurred. The superposition of a finite number of sinusoidal components yields a smooth transition from the black to white through shades of grey. Nevertheless, there is a mathematical advantage to the band-limited letter. If one knows the highest-frequency component of a band-limited signal and if samples are taken with equal spacing at any rate greater than twice the highest frequency component, the original band-limited letter can be completely reconstructed from the sample points alone (Nyquist, 1924, 1928; Shannon, 1949). Therefore, a letter whose highest frequency is 60 cycles per degree can be theoretically reconstructed solely from the measured grey values at evenly spaced sample points taken slightly more frequently than 120 times per degree of visual angle (Bigelow & Day, 1983).

The issue of spatial frequencies is returned to within §1.10.1 and Chapter 7 whilst that of antialiased fonts is discussed in depth within Chapter 8.

1.7 AN INTRODUCTION TO COGNITIVE ERGONOMICS

The study of ergonomics has been interwoven with computing science for many years and new areas of research on HCI have recently begun to emerge. These new fields are related to human cognitive activities such as language, problem solving, memory and attention. Many recent studies on the human factors of computing systems have included dependent variables traditionally associated with cognitive psychology such as human response time, the nature of errors, time to learn and self-reports. Additionally, many of the findings and hypotheses from human-interface research are explained in terms of cognitive theories and principles. Among the topics with cognitive implications that have been studied extensively are programming techniques, command languages, database access and editing. There is a cognitive component even in simple keying tasks, although cognitive factors are probably more important for complex tasks.

The discipline of ergonomics is generally said to have emerged shortly before World War II when it became clear that the advanced technology then being developed required improvements in the layout and design of instrumentation. The problem of human use of instrumentation was particularly critical in the area of aviation.

Ergonomics encompasses aspects of psychology, anthropometry and engineering. However, what is now being seen is a meeting of cognitive psychology and ergonomics, now parent subjects themselves, and a spawning of the new offspring—cognitive ergonomics: the application of the cognitive sciences to problems such as human–machine interaction. Cognitive ergonomics contributes to HCI in supporting aspects of the interactions which rely on the knowledge required by the human to use the computer in order to perform work effectively. Although human perceptual abilities, rather than knowledge, form the basis of this thesis, it is worth noting that such cognitive factors are inextricably linked and are far from clear-cut.

1.7.1 Knowledge

To psychologists, 'knowledge' is not straightforward. When learning a new skill, an individual both acquires new knowledge and reorganises existing knowledge. As part of this development, the nature of the knowledge appears to change from an *explicit* verbalisable form to an *implicit* automatically accessed form. Therefore, in contexts where the development of a skill is monitored over time, verbalisations are found to decrease gradually (Chase & Ericsson, 1981; Anderson, 1982). Such findings support the idea that knowledge may be held in memory in at least two forms: *declarative* and *procedural* (Winograd, 1975; Anderson, 1976, 1980). Declarative knowledge represents facts which can be articulated and which are stored as propositions, whereas procedural knowledge represents direct knowledge of actions, either physical or mental.

Anderson's approach to skill acquisition can be summarised as follows. In early stages, the person has available a certain amount of declarative knowledge about the task in hand. External instructions seldom specify the detailed procedures to apply, and so this declarative knowledge must be interpreted by way of pre-existing general procedures for action. This process of interpretation is assumed to make heavy demands on WM and attention. With increasing practice, task-specific procedures are developed, allowing the declarative knowledge to be bypassed. The process by which specific procedures are formed is called *knowledge compilation,* and it is divided into two distinct processes of *composition* and *proceduralisation.* Composition occurs when a sequence of low-level 'productions' is merged into a single higher-level production. Proceduralisation occurs when a production is made more specific in terms of the values or activities that it deals with. Complex tasks have alternative routes to successful performance, and the person must choose between them. A beginner's search of the task space is likely to be an inefficient hit-and-miss affair, but with experience the search becomes more selective (Gardiner & Christie, 1987).

1.8 INFORMATION PROCESSING

In 1958 Broadbent proposed that human information processing is restricted by a limited capacity filter between the large variety of sensations which we have and the attentive stages of input analysis. He claimed that both visual and auditory sensory systems function as parallel information processing channels, and environmental inputs, such as sound and visual stimuli, could be received simultaneously. A 'pre-categorical' analysis was said to be performed in that certain physical features were discernible (e.g. pitch and size) whilst others were not (content, context, or meaning). A filter mechanism allows only one message at a time to pass from sensory memory into the attention system. In an environment within which two human speech tracks run simultaneously, only one can be attended to at a time, Broadbent claimed that about 1.5 s is necessary to switch attention from one sensory modality to another, whilst others have estimated a much shorter span (Broadbent, 1971). Moray (1959), for example, indicates that only 50 ms is needed for very simple auditory stimuli. The degree to which our attention can switch between different inputs is also held as limited. Treisman (1960, 1964) and others considerably modified Broadbent's model to incorporate findings indicating that some content material from the unattended channels do in fact reach active attention. Such models imply that although stimuli in unattended channels are severely attenuated they have not ceased to exist.

Neisser (1967), whose theorising relied more on vision than audition, formulated a central distinction between *preattentive processes* and *focal attention.* He suggested that objects are not identified visually until they are 'segmented'; i.e. until the entire visual field is divided into figures on backgrounds. This segmentation was termed a 'preattentive process' by Neisser, suggesting that it is global and wholistic, whilst subsequent processes focus on specific aspects of figures for further analysis. The preattentive processes are presumably spatially parallel so that the entire visual field is simultaneously processed. These processes are employed principally to direct focal attention, but are also important in directing or guiding immediate bodily motions such as orienting responses and locomotion. As an example, driving a vehicle frequently does not require detailed attention to the surrounding visual world. Rather, one responds to some objects directly, without recognition.

When an important aspect of the environment is segmented, 'focal attention' occurs either through a change in the area of view or by a narrowing of attention. Focal attention acts on the important aspects of the field segregated by the preattentive processes. Consequently, analyses may be performed, or figures may be constructed, using the information extracted previously. The constructions may result in identification (e.g. 'It is a circle'), in description of the attributes ('It is a green and edible'), or perhaps in negative statements ('It is not an apple').

Preattentive processes locate and isolate objects in the visual field. No further specification of the properties of such objects is denoted other than that given by the visual features themselves. One region of the field might thus be segregated according to colour, but not according to shape. The discrimination between a square and a circle, for example, requires further processing beyond what the preattentive mechanisms alone can provide. This general issue is returned to within Chapters Two and mentioned in the context of icons versus words within Chapter Six. Neisser's distinction between preattentive and focal attention is a useful heuristic device. It allows one to be more explicit in seperating some of the control function of eye movements from the identification components of target search (Haber & Hershenson, 1980).

'Divided attention' refers to the situation where simultaneous processing of multiple sensory inputs is required. The serial controlled mechanisms of attention produced divided attention limitations, but one group (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977) have shown that these limitations can be by-passed when automatic parallel processing is used. Automatic processing takes place in long-term memory (LTM), is triggered by specific inputs and operates largely independently of the subject's control. When automatic attention processing is activated, it will not necessarily affect ongoing controlled processes. Thus there is an obvious similarity between the distinction Neisser (1967) makes between preattentive and attentive processing and the distinction Shiffrin and Schneider make between automatic parallel and serial controlled processing. However, there are also a number of dissimilarities. For instance, contrary to Neisser's view, parallel processing from multiple external stimuli to meaningful content and context is possible within the approach of Shiffrin and Schneider in so far as the automatic parallel processes for the appropriate multiple inputs are well-learned.

1.8.1 The tripartite model

It may be argued that a cognitive ergonomics approach ultimately involves an understanding of user models, user motivations and user reasoning. It should provide information for prediction, planning, and for putting the user in control. However, this 'higher level' approach is not the central theme of this thesis. Here, a more peripheral or perhaps truly interactive approach is taken up.

Like all human behaviour, interacting with a computer involves three types of basic human processes; perception, cognition and motor activity. A system designer's job is to minimise the work required by these processes, both individually and in combination, and to achieve this he needs some key concepts from psychology, physiology and ergonomic studies. Ergonomics bears most on the perceptual and motor processes, concerned as it is with the application of knowledge to human performance. The field of cognitive psychology provides insight into our memory and learning processes. The tripartite model is akin to the 'Model Human Processor' of Card *et al.* (1983). Similarly, this model can be described by (a) a set of memories and processors together with (b) a set of principles, called the 'principles of operation'.

The perceptual process Perception is the process whereby unintelligible physical stimuli (in this case generated by the computer) are received by the receptor organs, transmitted to the brain, and are there recognised by a process theorised to be akin to pattern recognition. The dominant stimuli in computer use are visual, although audio stimuli are usually present to some degree (keyboard clicks, disc access arms moving, etc.), and are increasingly used as an adjunct to or replacement of visual stimuli, e.g. tones to catch the user's attention. Tactile stimuli from the keys and so forth are also present.

The cognitive system Ergonomics has traditionally not been concerned with cognition but concentrated instead on designing equipment for efficiency of manipulation whilst optimising physiological limitations. However, tasks involving interactive computers, almost by definition, have non-trivial cognitive components even at this lowest, lexical, level. The study of cognition provides insights into ways to structure hierarchical menus, the number of choices to present, the types of words to use, and ways to abbreviate commands. When one learns how to use an interaction technique, one acquires and organises information concerning its use. If the information fits into categories or concepts already understood, then the learning can proceed rapidly; if not, learning is slower. A menu of twenty choices is easier to comprehend if the choices are grouped in several logically related subsets (Foley *et al.*, 1984).

The motor process The motor process comes into play when the user, having received, recognised and decided how to respond to the stimuli, performs a response in physical actions. This might involve picking up a stylus, moving it to the tablet, and then causing the cursor to move to a particular point on the screen. The process almost always depends on continuous perception and cognition to close the feedback loop. Perception informs the user of the locations of the mouse or stylus, and of the cursor and target, respectively, and cognition continuously decides whether or not these locations have converged. In typing a command in response to a prompt, however, a touch-typist would not depend on visual perception of the keys for feedback but instead would rely on kinaesthetic, tactile and auditory perception. The typist would require negligible cognition to know whether she had hit the right key. A novice, 'hunt-and-peck' typist would behave quite differently, however.

The design and selection of interaction techniques for a task must take into account the perceptual, cognitive and motor processes involved, even if they should appear trivial. Generally, the design goal is to minimise the time taken by each process, although this cannot always be measured in the case of cognition. Learning time must also be considered by the designer, especially for infrequent users of a system. *Task analysis* (which may very loosely be regarded as identification of the processes involved) is the important first step in analysing and designing a technique (Foley *et al.*, 1984). Task analysis is one of the major methodological approaches available to HCI specialists. Our task of describing this approach in any great detail is obviated by the fact that what is perhaps the definitive text on the subject is available (Diaper, 1989). The Keystroke Level model and the GOMS model of Card *et al.* (1983) integrate classical time and motion study concepts with certain aspects of cognitive psychology, providing useful engineering models of user performance.

1.9 VISION ERGONOMICS AND VISUAL PSYCHOPHYSICS

1.9.1 Aspects of vision relevant to display design

This thesis takes the line that most interaction techniques commence with visual perception. The user finds a menu selection, an item to be detected, or the cursor and recognises an object or character. An important consideration is therefore how to display information so that the user can quickly locate the items needed. This can involve methods such as colour coding, spatial coding, blinking, brightening, movement and reverse video to call attention to the specific parts of the display. If a task might involve any of the entities being displayed, there is no point in displaying them all. Often, however, particular subsets of displayed information are most germane to the application at specific points in time (Foley *et al.*, 1984).

The retina is sensitive to light and records its intensity, wavelength and spatial distribution. Although the eye takes in the visual scene over a wide angle of about a half-hemisphere, detail is obtained only over the narrow region (approximately two degrees across) of the fovea. The remainder of the retina provides peripheral vision mainly for saccades, each taking in the order of 30 ms to jump to a new point of focus and resting there 60–700 ms for a total duration of

$$\text{Eye movement duration} = \text{travel} + \text{fixation time} = 230\ (70 \text{ to } 700)\ \text{ms} \quad (1.1)$$

In this expression, the rate 230 ms represents a typical value and the numbers in brackets indicates that values may range from 70 ms to 700 ms depending on conditions of measurement, task variables, or subject variables (Russo, 1978). When the target is greater than some 30 degrees away from the fovea, head movements occur to reduce the angular distance. These four components—central vision, peripheral vision, eye movements and head movements—form a unified system to provide a continual representation of the visual world to the observer (see also §5.2.1).

Although no one,as yet, has defined or precisely described the domain of vision ergonomics, a major component within vision ergonomics must be visual psychophysics. In this section, some relevant aspects of this field are overviewed.

Electronic displays are ubiquitous within HCI and by far the most important sense used to interpret and analyse information provided by such displays is vision. This area has a vast and distinguished history which features names as eminent as Da Vinci, Descartes, Newton, Helmholtz and Young. Here, more recent developments are concentrated on and the review of Travis (1990) is followed of those areas relevant to users of electronic displays.

It is a fundamental premise within vision ergonomics that the quality of the interface can be improved by first acknowledging the perceptual characteristics of users and, second, by aiming to produce displays that optimise visual performance. This will ensure that images, whether icons (see Chapter 6), graphics (see Chapter 8), are maximally legible. Three seperate components of visual ability can be considered:
(a) spatial,
(b) temporal, and
(c) chromatic factors.

1.9.2 Space

As the retina is traversed outwards from the fovea, the density of cone cells decreases markedly. At just 2 mm from the fovea the density of cone cells is reduced to one percent of their maximum value (Perry & Cowey, 1985), and hence the human eye is most sensitive to objects presented within the central 2° of visual angle. Under daylight viewing conditions, spatial and colour vision drop dramatically as a function of eccentricity from the fovea. This area of interest is discussed more fully within Chapter 2.

Visual acuity may be measured in many ways other than the typically encountered optician's eye chart (Westheimer, 1979). An individual's threshold for recognising a letter (about 30 seconds to 1 minute of arc) is substantially higher than the threshold for detecting the discontinuity between two lines (about 2 seconds of arc). Both of these are higher than the threshold for detecting the mere presence of an object (about 0.5 seconds of arc). A more important limitation with these measures is that they assess vision only for small, high contrast objects. In real life, it is often necessary to distinguish large objects of low contrast (e.g. an approaching car on a foggy road).

The most complete description of the spatial abilities of the visual system is provided by the *contrast sensitivity function* (CSF). This is a function relating the sensitivity of a spatially repetitive pattern as a function of *spatial frequency.* The pattern most often used by vision scientists is the sine wave grating (e.g. Campbell & Robson, 1968). Such a pattern consists of a series of light and dark bars sinusoidally modulated around a mean luminance (see Fig. 1.5). Those depicted in Figure 1.5 are examples of *square wave* gratings. Continuous sine wave gratings (where there are no distinct transitions from light to dark) are also typically found in the research literature.

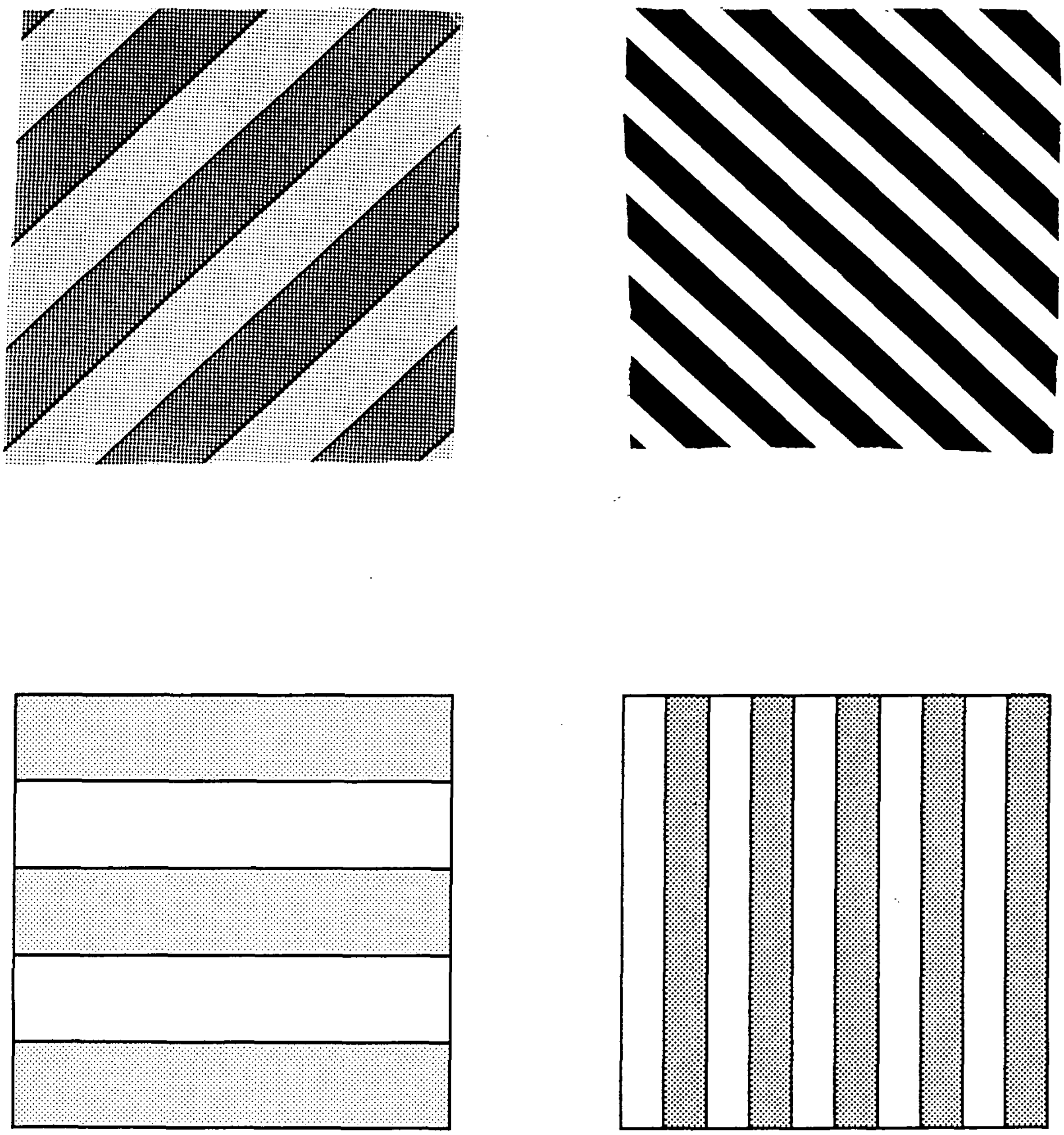


Figure 1.5: Four examples of square wave gratings.

Contrast sensitivities offer one measure of the resolving power of the visual system by providing a determination of the minimal contrast needed to perceive a test pattern. For instance, rather than reducing the width of black/white grating bars, it is possible to reduce the difference in luminance between the black and white stripes. The maximum contrast possible with a printed pattern of black on white is about 20 to 1, but if a positive transparency slide is constructed and light is projected through it, the range of contrasts possible increases to 100 to 1 or more. It would then be possible to construct a 'frequency of seeing' function for a grating of any spatial frequency by altering the contrast. The difference between the two luminances at which the grid is resolved 50 percent of the time is considered the contrast threshold (Haber & Hershenson, 1980).

It is generally agreed that the visual system appears to perform a Fourier analysis of the visual scene. Essentially, this means that the visual system decomposes objects into a set of sinusoids of different frequencies and amplitudes (although see also Marr, 1982). As *any* object can be created from these building blocks, the CSF is obviously a more relevant measure of visual performance than classical tests of acuity. In fact, it has been shown, for instance by Ginsburg *et al.* (1982) studying pilots' flying performance, that the CSF is a more accurate predictor of visual performance than visual acuity. An introduction to the field of sine wave and Fourier analysis is provided by Sekuler and Blake (1985). The topic of sine waves is discussed at length in Chapter 7.

Visual angle is the appropriate metric for specifying the size of objects on a display. It is defined as the angle subtended on the retina by the two extreme points of a visual stimulus. For example, if an object of size h is a distance d from the retina, then the visual angle subtended, ϕ, is:

$$\phi = \arc \tan \frac{h}{d} \tag{1.2}$$

At a normal reading distance, an uppercase letter on this page will subtend a visual angle of 25 minutes. A Macintosh terminal screen viewed from the same distance would subtend a visual angle of about 25°. Figure 1.6 depicts the technique for computing visual angle.

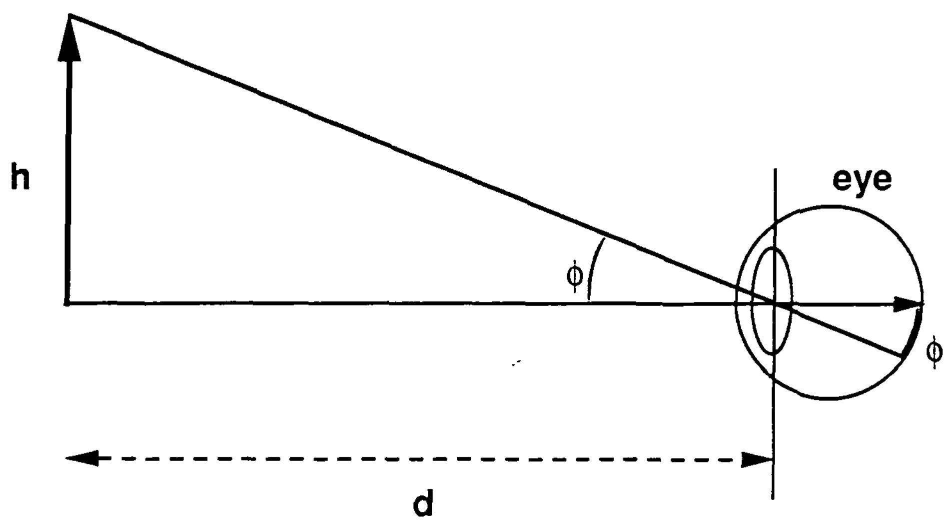


Figure 1.6: An illustration of how the visual angle is computed.

1.9.3 Time

Flicker Considerable literature exists on VDU flicker (e.g. Farrell *et al.*, 1988). Figure 1.7 shows sensitivity to flicker as a function of luminance. Data are shown for two luminance levels; the high luminance level was 850 td (trolands; see §1.8.4), and the low luminance level was 7.1 td. The stimulus was a 65° white field modulated sinusoidally in time. The two plots show that the flicker rate to which a user is most sensitive changes with retinal illumination. That function which best approximates the average user of an average VDU is the upper curve labelled 'high luminance'. Under these conditions, users are most sensitive to flicker at about 16 to 25 Hz.

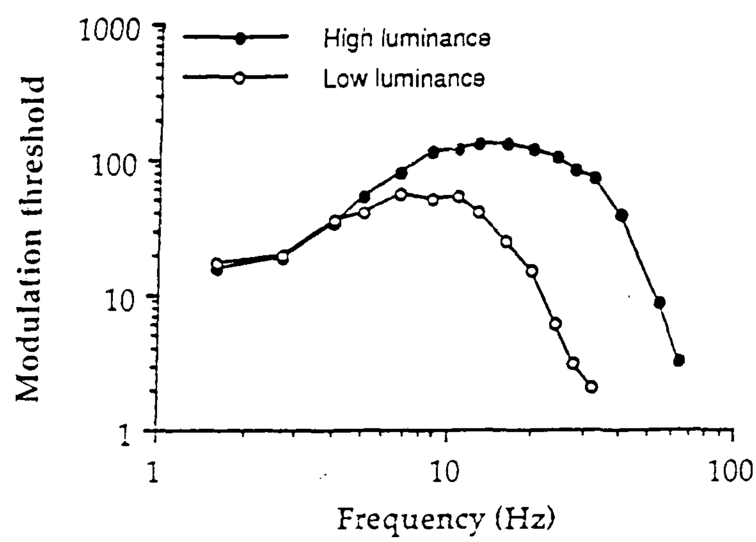


Figure 1.7: Threshold as a function of flicker frequency. The data were obtained by Kelly (1961), are tabulated in Wyszecki & Stiles (1982), and are graphed in Travis (1990). From Travis (1990), Visual psychophysics and user interface design, *Behaviour and Information Technology*, **9**, p. 431. Reproduced with permission.

At luminance levels typically found with VDUs, users are sensitive to flicker above 50 Hz, which is the refresh rate of televisions in the UK. Indeed, for very bright screens users are sensitive to flicker at levels well in excess of this. Therefore, even at quite moderate brightness levels, display flicker may be a problem. For instance, the VGA graphics card, a 'PC' standard, refreshes monitors at 70 Hz. One way to reduce display flicker is to reduce luminance. Whilst the designer may not be able to reduce the VDU luminance in the work-place, he may still partly govern this by producing light-on-dark, positive contrast displays. This would reduce the overall luminance, but it should be noted that negative contrast displays are generally to be preferred (see §1.9.4).

1.9.4 Quantification of visual patterns

For visual perceptions to be properly quantified, a number of terms require definition.

Luminance is often incorrectly considered to be synonomous with 'brightness'. Brightness is a subjective term, whilst luminance is a photometric term. Photometry is the branch of physics concerned with characterising lights according to their visual effectiveness. Thus, luminance is a measure of the light reflected or emitted from a surface and passed through a filter that is intended to represent the luminance sensitivity of the eye. The SI units of luminance are cd/m^2. However, it is not possible from this to state how much light is actually entering the eye as this will depend on pupil size and this may vary from individual to individual as well as with varying illumination levels. However, when the pupil size of the observer is known, *retinal illumination* may be specified in 'trolands' (*td*). As one's perception of brightness does not depend only on the photons absorbed in the retina, lights of equal luminance may differ in brightness and vice versa (Burns *et al.*, 1982).

Contrast is commonly defined as the ratio of two luminances, but at least three precise definitions are available:

$$\text{Michelson contrast} = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} \tag{1.3}$$

$$\text{Contrast ratio} = \frac{L_{max}}{L_{min}} \tag{1.4}$$

$$\text{CIE contrast} = \frac{L_{max} - L_{min}}{L_{min}} \tag{1.5a}$$

where L_{max} represents the maximum luminance and L_{min} represents the minimum luminance. The Michelson contrast is sometimes referred to as *contrast modulation* and its use is best restricted to conditions where the mean luminance does not change. Character or object contrast on a cathode ray tube (CRT) is best computed across conditions where mean luminance may vary by use of equations (4) or (5a). A variation of equation (5a) is sometimes used and this allows a distinction between negative contrast screens and positive contrast screens.

$$\text{Contrast} = \frac{L_{object} - L_{bg}}{L_{bg}} \tag{1.5b}$$

where L_{object} is the luminance of the object or character and L_{bg} is the luminance of the background.

Figure 1.8 shows two seperate CSFs; one function being obtained at high luminance, the other function at low luminance. It might be noticed that the sensitivity is highest for medium sized objects and decreases for both larger and smaller objects.

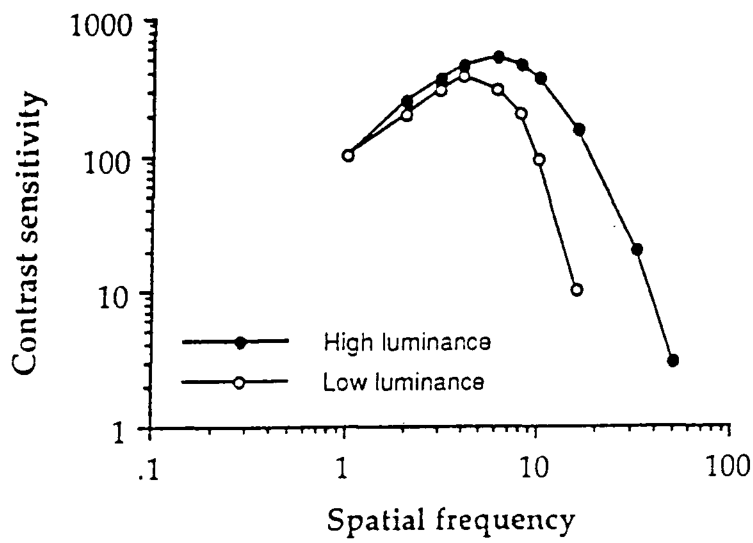


Figure 1.8: Contrast sensitivity functions for gratings at a high mean luminance (filled symbols) and a low mean luminance (open symbols). The x-axis plots spatial frequency: large objects (sine waves that consist of fat bars) fall at the left and small objects (thin bars) fall at the right. The y-axis plots contrast sensitivity (the reciprocal of Michelson contrast at threshold). From Travis (1990), Visual psychophysics and user interface design, *Behaviour and Information Technology,* 9 p. 428. Reproduced with permission.

From Figure 1.8, it can be seen that sensitivity decreases over virtually all spatial frequencies as luminance is decreased. Conventional measures of acuity can be approximated from these curves by extrapolating the high spatial frequency parts of the curve until they cross the abscissa (Ginsburg, 1979). The effects of luminance on sensitivity are considerable up to around 900 td (van Nes & Bouman, 1967). For an average observer this is about 110 cd/m^2. Increasing the luminance still further has only minor effects on improving sensitivity.

This change in sensitivity with luminance is of direct relevance to the designer of VDUs. A common decision is whether to use negative or positive polarity screens. Positive polarity screens result in a higher space averaged luminance and are said to possess substantial advantages (Bauer & Cavonius, 1980) because, as Travis (1990) notes:

(1) Visual sensitivity increases with luminance (van Nes & Bouman, 1967); see Fig. 1.7).
(2) At high luminance levels the pupil diameter is relatively small. This results in greater depth of field (the user needs to refocus less often) and fewer optical distortions (Millodot, 1982).
(3) They are less subject to glare (Bauer *et al.*, 1981). It has been shown (Findlay & Wilkinson, 1984) that glare reduces sensitivity to all spatial frequencies approximately equally. Whilst glare can be reduced or eliminated by appropriate placing of the VDU relative to the light source (e.g. perpendicular to windows), it is often the case that the display designer has no control over the siting of hardware. In such situations, positive polarity screens are clearly preferable.
(4) They cause less contrast adaptation in the user (Pelli, 1988; see below).

A disadvantage with positive polarity screens is that the human sensitivity to flicker increases with luminance, too (Kelly, 1961). Whilst users are therefore more likely to perceive flicker on positive polarity VDUs, this problem can be overcome by using screens with high refresh rates (see §1.8.3).

Size Whether data displayed are in the form of alphanumerics, graphics or iconic images, for a VDU to be usable the data presented on it must be discriminable. In order to make a decision about size, the designer needs to take account of the abilities of the user as well as the conditions under which the display will be used. For example, an icon that needs to be identified at a viewing distance of a few metres (e.g. in a process control plant) should be larger than one that requires identification at 50 cm (e.g. in an office environment). The abilities of the specific users of the system also need consideration. For example, the ability of the eye to focus or accomodate substantially decreases with age. Older users will therefore have different visual requirements than younger ones. This is largely an intractable problem as designers will rarely have control over who uses the display.

1.9.5 Colour

Colour is a multidimensional concept varying in:

(a) hue (what is commonly regarded as colour; e.g. red, blue or green),
(b) saturation (the amount of white in a colour, or its purity), and
(c) brightness (how bright or dark the colour is).

For most experimental purposes, hue and saturation can be considered as a uniform construct, *chromaticity*. It is often commented that colour is one of the most misused terms by designers and ergonomists alike.

From psychophysical and physiological measurements, a model of colour vision has emerged. The first stage of colour vision is termed *trichromatic* since it is based on light absorption in three types of cone cell. The three classes of cone (loosely termed red, green and blue) have broad and overlapping sensitivities to wavelength (Smith & Pokorny, 1975; Baylor *et al.*, 1987). Effectively, each class is 'colour-blind' in that a class is unable to distinguish a hue change from a brightness change (Naka & Rushton, 1966). In order to extract information from these signals, outputs from the different classes of cones must be compared. This occurs during a second stage where the signals are antagonistically compared and differenced. Two classes of these 'opponent pathways' have been isolated. One differences the output of red and green cones, and one that differences the output blue cones with some combination of red and green cones (Krauskopf *et al.*, 1982; Derrington *et al.*, 1984). At this stage, brightness information is not carried in a seperate physiological channel. Rather, it is multiplexed with the chromatic information carried in the 'red/green' opponent pathway (Lennie & D'Zmura, 1988). Boynton (1979) and Lennie and D'Zmura (1988) contain more complete information on this model. Colour is discussed again as a variable within Experiment 6.

Colour coding The most useful feature of colour coding is possibly that it adds an extra dimension of coding to the display. The London Underground map is one example and Prestel and Teletext are examples where this sort of coding is used in screen displays. A further illustration is 'colour thresholding' where the value of some attribute of a device (e.g. the temperature of an engine) is converted to a colour. With this technique, the engine could be drawn as an icon on the screen and the colour of the icon might change according to a particular scheme. A good, if somewhat now dated, review of colour coding was provided by Christ (1975) who concluded that 'colour may be a very effective performance factor under some conditions, but... it can be detrimental under others'. There seems little doubt that judicial use of colour coding can help objects 'pop out' of their background (Treisman & Gormican, 1988; see §2.5).

Colour coding contains three dimensions which may be considered seperately; brightness, saturation and hue. A simple example of brightness coding would be the use of a grey scale, although a single hue varying in brightness could also be employed. Saturation coding might be used where some table or neutral state is indicated by white and as the state becomes more critical the colour becomes a more saturated red. Both of these types of coding are probably preferable to hue coding, where different conditions are assigned to different hues. Individuals seem to possess a cognitive metric for brightness and saturation, but not for hue. One can measure brightness and saturation relative to some norm; blackness in the first instance, whiteness in the second. There is no stable state for hue, however. It is impossible to quantify how different red is from green as opposed to from yellow. It is often believed that psychological 'distance' can be approximated by the spectral ordering of hues, but this assumes that all users know the sequence of colours in the spectrum. Even when this is so, accessing this information is not effortless, and to make colour coding effective it should be as effortless as possible (Travis, 1990).

1.10 EYE MOVEMENTS

Eye movements are classified into various types, some of which are more important for perceptual processes than are others. *Saccadic movements* are very rapid, jerky movements that move the eye in the socket to a new orientation. This is the principal movement used in visual search, in reading, in examining a scene, or in noticing movement on the periphery of vision. *Smooth pursuit movements* are used to track an object moving across the field of view, as in watching a moving vehicle. These movements are smooth, rather than jerky, much slower than saccadic movements and require a moving object for them to occur. The third type of large eye movements are *vestibulo-occular movements,* in which the eyes rotate in the head to compensate for head and body movements in order to maintain the same direction of vision. Saccadic, pursuit, and vestibulo-occular movements each produce *conjugate movement,* in which the two eyes move in the same direction and (with slight exceptions) by the same amount. *Vergence movements* produce changes in the direction of the two eyes in opposition to each other that occur when we shift our gaze from a near to a far object (divergence) or vice versa (convergence), or when we move our body away from or towards an object upon which we maintain our gaze. While all of these movements can be either large or small, there are three types of movements that are always small—*drifts, micro-saccades* (often called flicks), and *physiological nystagmus* or *tremors* (Haber & Hershenson, 1980).

Fixations are pauses between saccades. For a few hundreds of milliseconds, gaze is virtually static with respect to the head. Positional drifts and microsaccades may occur during longer fixations, but their amplitude does not exceed ten minutes of arc (e.g. Steinman, 1965) and their functional significance has been questioned (see Ditchburn, 1980). However, fixations are the most important aspect of visual search, or 'the basic unit of encoding' (Loftus, 1972). Much of what we can expect to understand about visual search comes from an analysis of how long these fixations are and where they occur.

At least three processes may be assumed to occur within the 300 ms or so of a typical visual search fixation: (a) the analysis of the visual stimulus in the foveal field, (b) the sampling of the peripheral field, and (c) the corresponding planning of the next saccade. The minimum duration of the stimulus-processing has been estimated at about 100–150 ms (e.g. Eriksen & Eriksen, 1971), although further processing may occur after correct identification. Therefore, minimum times necessary to guide further movement or to memorise stimuli may be underestimates. It is likely that these three processes are performed concurrently, with several nervous structures acting in parallel, although the amount of overlap among them is largely unknown (e.g. Potter, 1983). Indeed, the encoding of the visual stimulus may exceed the duration of the fixation (Intraub, 1980). However, what is known is that the time spent for each process, which can be construed as a reaction time, is affected by several contextual factors (Viviani, 1990).

1.11 VISUAL ATTENTION (WITHOUT EYE MOVEMENTS)

The term 'attention' is used to describe a number of phenomena which are not necessarily related. The volume edited by Parasuraman and Davies (1984) illustrates the range of meanings. In his review of the orienting of attention, Umiltà (1975) notes that attention can mean 'a selective process, whereby some information is perceived consciously, whereas other information is either analyzed unconsciously or is filtered out. It can mean a laborious process, whereby processing resources are voluntarily allocated to a particular task or activity and a sustaining process, whereby receptivity to input information is regulated so that vigilance is heightened for a short period of time or maintained more or less constant for a longer period' (p. 175). Umiltà also notes that even in the restricted sense of a *selective* process, attention refers to different phenomena (see, e.g., Johnston & Dark, 1986). An individual may selectively attend: (a) to information presented in a particular modality (e.g. Posner *et al.*, 1976; Broadbent, 1982; Shapiro *et al.*, 1984); (b) to information originating from a particular position in space (e.g. Moray, 1975; Posner, 1980; Shapiro & Johnston, 1987); (c) to stimuli possessing a particular colour or shape (e.g. Francolini & Egeth, 1980; Lambert & Hockey, 1986); or (d) to items belonging to a particular class or category (Posner & Snyder, 1975; Neely, 1977). Whilst these are disperate phenomena, all of them contain the common feature that information to which attention is selectively allocated is processed more efficiently than non-attended information. In this thesis, our concern is primarily with spatial attention; i.e. with the selective allocation of attention to a particular position within a display.

1.11.1 The attentional focus

The typical analogy for the deployment of visual attention regards the focus of attention as like a *spotlight* beam. Such a beam has three important properties: (a) it moves from one location to another; (b) it moves in analogue fashion rather than jumping instantaneously from one location to another; and (c) it is characterised by a specific size. It therefore illuminates at different times different locations and the extent of the illuminated area at a given time is restricted.

Shifts of attention are normally accompanied by eye movements and an observer, besides directing attention to a position in space, also foveates on it. This creates a problem because a performance (information processing) improvement at that position could be due to retinal rather than attentional factors. As there is ample evidence that attention can be shifted from one locus in the visual space to another without shifting eye position (e.g. Posner, 1980), the confounding between retinal and attentional factors can be overcome. As was first indicated by Jonides (1983), there appear to be two mechanisms, one visual and one attentional. Foveation is the visual mechanism and this allows an individual to overcome the limited resolving power of the peripheral retina via fixation shifts. Independently, the attentional mechanism permits a shift in the internal allocation of processing capacity from one position to another without a shift in gaze. Umiltà's review considers studies in which an improved performance at a given location could be solely attributable to attentional factors.

Much of what is known about the size of the attentional focus comes from visual search tasks. When visual displays are presented so as to exclude the effects of retinal factors, accuracy and/or time for identifying targets can be held to reflect attentional processes. From this work, it would appear that under certain conditions, the resources of the attentional system are distributed evenly over the entire display, with parallel processing of the display items, whereas under other conditions a focused or serial scanning occurs. A number of two-stage models of attentional selection arose from these findings (e.g. Shiffrin, 1977; Hoffman, 1978, 1979; Duncan, 1980; Bergen & Julesz, 1983), within which that proposed by Jonides (1983) has been particularly influential. Within this model there are two modes of attending to a visual display: attentional resources can be allocated evenly across the entire display or they can be concentrated on one display location only. In the first mode all the display elements are processed in parallel at a uniform, and relatively slow, rate. In the second mode processing of the precued element is facilitated, whereas processing of the other elements is inhibited.

Eriksen and colleagues (e.g. Eriksen & Yeh, 1985; Eriksen & St. James, 1986) have subsequently proposed a modified version of the two-stage model in which they suggest that attention is not restricted only to two distinct modes of operation; i.e. a distributed and a focused mode.

> In their opinion, the 2 processing modes described by Jonides (1983) are merely poles on a continuum of attentional distribution in the visual field. The distribution of attentional resources may vary from a state in which they are

uniformly spread over the entire visual field to a state in which they are highly concentrated or focused on an areas as small as a fraction of a degree of visual angle. They also believe a zoom, or variable power, lens is a more apt analogy than a spotlight. In essence this is a different hypothesis because the zoom lens analogy explicitly predicts an inverse relationship between the size of the area of focused attention and the efficiency of processing obtained in that area, whereas the spotlight analogy is uncommitted in that respect. (Umiltà, 1988, p. 177).

1.12 ATTENTION AND EYE MOVEMENTS

As noted by Shepherd and co-workers (1986), the relationship between eye movements and orienting of attention can manifest itself in three forms:

(1) According to the *identity hypothesis*, the mechanisms involved in the generation of eye movements are identical to those that produce attention shifts.

(2) Diametrically opposed to this is the *independence hypothesis* which maintains that there are two sets of mechanisms, one for eye movements and one for attention shifts, which are not functionally related.

(3) The third alternative, the *interdependence hypothesis*, suggests that the two sets of mechanisms are neither identical nor completely independent, so that the functioning of one can be facilitated or inhibited by the other.

The identity hypothesis seems to have been disproven by several studies (e.g. Jonides, 1981; Posner & Cohen, 1984; Maylor & Hockey, 1985, 1987; Briand & Klein, 1987; Tassinari *et al.*, 1987) which show that attention can be allocated, either automatically or voluntarily, to different points in the visual field without eye movements. The evidence concerning the other two hypotheses is less clear cut. For instance, whilst papers by Klein (1980) and Posner (1980) have supported the independence hypothesis, both Klein's findings (Shepherd *et al.*, 1986; Rizzolatti *et al.*, 1987) and those of Posner (Umiltà, 1988) have been disputed. Perhaps the available evidence (e.g. Remington, 1980; Shepherd *et al.*, 1986) favours the interdependence hypothesis. Umiltà (1988) provides a detailed discussion of this debate.

1.12.1 Fixation durations during free scanning

Several studies (overviewed by Viviani, 1990) have indicated that the statistical properties of fixation durations during free scanning of a complex scene are not stationary. For instance, Antes (1974) demonstrated that the average duration over successive tenths of a 20 s episode of free scanning increase from about 220 ms at the beginning of the exploration to about 300 ms at the end. The first moments of the observation appear to be spent concentrating on the salient features of the scene with relatively short fixations. Later, other features are examined for longer times. In addition, Nodine (1982) recorded subjects' eye movements whilst they compared original and altered versions of famous paintings and noted a similar concentration of early fixations on the most salient features. This trend of increasing fixation durations with viewing time was confirmed by Friedman and Liebelt (1981), but they demonstrated that this increase was not a consequence of shifting the *attention* from more salient to less salient features. Essentially, it was argued that fixation durations depended on both the degree of contextual likelihood and their serial order. There are implications for the serial versus parallel processing debate and Viviani argues for the former in this sort of task. This is an interesting on-going debate within visual attention theorising, but not of central significance to the search tasks used within this thesis. Here, the subject needs to locate and (usually) actually process the information within the displays for understanding or recognition.

Viviani also comments that when the selection of targets is imposed by the experimenters, only the timing, size and duration of the saccades need be considered. With complex, naturalistic scenes, an internal organisation should also emerge in the actual sequence in which targets are selected. Noton and Stark have reported that subjects viewing a naturalistic picture tend to look at a particular idiosyncratic sequence of features termed a 'scanpath' (Noton, 1970; Noton & Stark, 1971*a,b*). However, because exploratory/search movements are intrinsically variable, scanpaths are difficult to identify; i.e no specific sequence can be expected, either repeatedly or interspersed with other sequences, with exactly the same parameters. As no quantitative criterion was proposed to discriminate scanpaths from other saccadic sequences, the original formulation of 'Scanpath Theory' proved of limited value. Stark and Ellis (1981) then described a probabalistic approach for defining such a criterion (see also Ellis & Smith, 1985).

> The approach is based on a mathematical tool, the *transition probability matrix,* which is routinely used for the description of discrete Markov processes (Bharucha-Reid, 1960). Specifically, let us suppose that, by a preliminary analysis of the scanning movements, we have been able to identify some regions of the scene where fixations tend to cluster. Let $R_1, R_2, \ldots, R_n$ be these regions (states, in the Markov parlance). Then, a sequential analysis of the visual scan permits us to estimate a $(n \times n)$ square matrix, P_{ik}, whose entries represent the probability that a fixation on R_i is followed by a fixation on R_k [Fig. 1.9]. The peculiar merit of this formalism is to expose directly and operationally the transition between the static description of its dynamic properties. In particular, if only a few entries in each row dominate all the others (the sum of

the entries in a row is always 1), scanning movements should contain cyclic sequences [Fig. 1.10]. Everytime the movement enters one of these cycles, it has a high probability of going through the entire sequence. Subcycles of a main sequence can also be defined by particular combinations of entries in the transition matrix, and their probability of occurrence can be calculated. (Viviani, 1990, p. 382)

Figures 1.9 and 1.10 depict a Markovian analysis of scanpaths within in a hypothetical example in which the visual scene has been divided into 25 regular regions $R_1, R_2, \ldots R_{25}$. Figure 1.9 shows a portion of the 25 × 25 square probability matrix that describes the sequential regularities of an experimental scanpath (asterisks indicate unspecified values). Each entry P_{ij} of this matrix represents the conditional probability that a fixation in the region R_i of the scene (starting position) is followed by a fixation in the region R_j (target position). These probabilities must be estimated as the ratio between the number of transitions $i \rightarrow j$ and the total number of saccades during the period of exploration. Once the gaze has left a position it must land somewhere. 'Thus, the probabilities in a row sum to one. The basic feature of the Markovian model is that the probability with which the gaze will move from the current location to other regions is independent of the previous path followed to reach that location. A transition probability matrix can always be calculated from actual data. However, simple inspection of the matrix will not reveal whether or not the scanpath is Markovian. This can only be ascertained through the analysis of the n-steps transitions' (Viviani, 1990, p. 381). In the present example, it is assumed that from each indicated starting position the gaze could only reach two targets, although one of those two is highly preferred.

Target position

	1	2	3	*	*	*	10	11	*	13	*	19	*	22	*	*	25
1	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
2	0	0	0				9	0		0		0		0			1
3																	
*																	
*																	
*																	
10	1	0	0				0	0		0		9		0			0
11	0	9	0				0	1		0		0		0			0
*																	
13	0	0	2				0	8		0		0		0			0
*																	
19	0	0	0				0	0		0		2		8			0
*																	
22	1	0	0				0	0		9		0		0			0
*																	
25	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Figure 1.9: A Markovian analysis of scanpaths (from Viviani, 1990; reproduced with permission).

Figure 1.10 shows the scanpath that would result if, starting from one of the regions, the gaze always selected the preferred target. It should be noted that despite the high transition probabilities associated with the arrows, the predicted probability of observing the complete cycle is only 0.41 and drops to 0.16 for two successive cycles. Therefore, in order for repeated cycles to actually be present in the scanpath, each transition in the cycle must be extremely probable.

An early use was made of this type of modelling by Bozkov *et al.* (1977) in their analysis of target point selection during scanning eye movements. They employed irregular polygons as stimuli and noted that the majority of fixation points were at the angles. The method proposed within this thesis is complimentary rather than in opposition to those described by Stark and Ellis or Viviani and is more concerned with sequential saccade *direction* rather than saccade *position*.

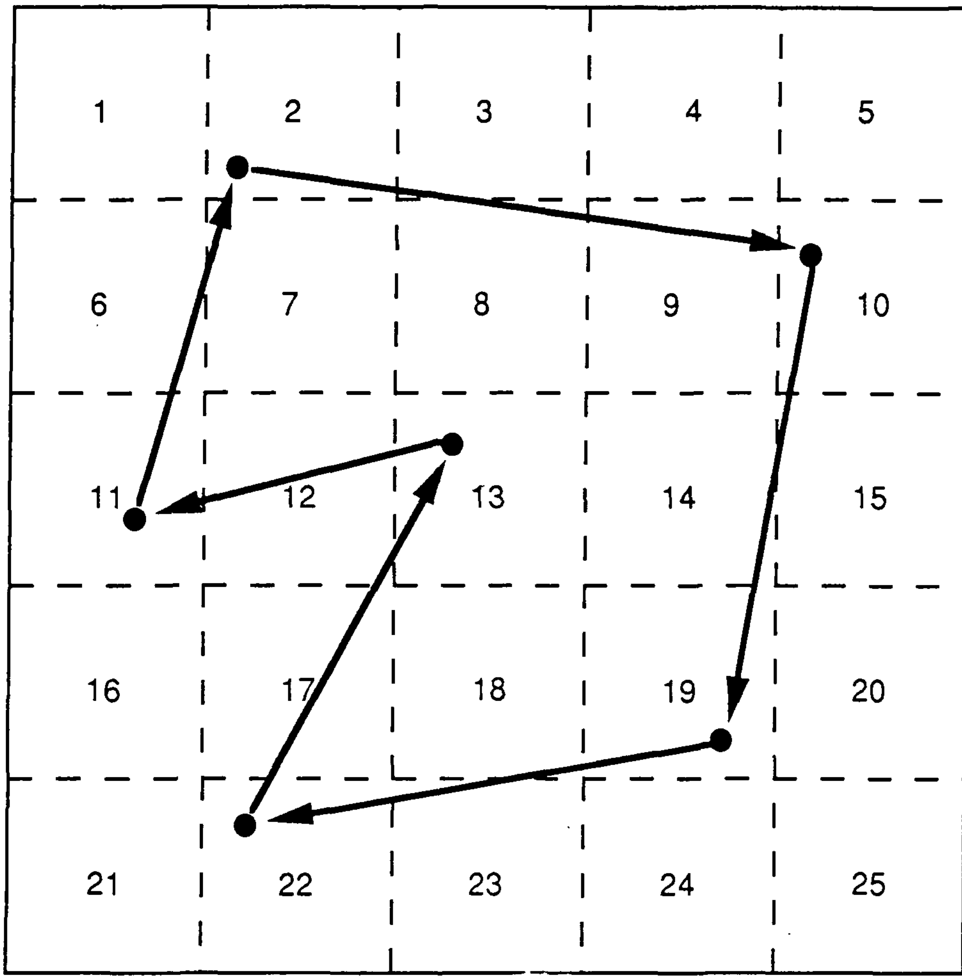


Figure 1.10: An example of a scanpath (from Viviani, 1990; reproduced with permission).

1.13 MEASURING EYE MOVEMENTS

1.13.1 General considerations

As all areas investigated within this thesis contain eye movement analysis studies, it would be appropriate to introduce the important field of eye movement monitoring techniques at this point. Eye movements are angular rotations of the eyes that orient the eyes for different directions and distances of space. Any translational movement (i.e. movement of the eye laterally as a result of a head movement) should not be mis-measured as angular rotation. According to Hallett (1986), an ideal measuring instrument should satisfy the (comprehensive and strict) following requirements:

(1) Offer an unobstructed field of view with good access to the face and head.
(2) Make no contact with the subject.
(3) Meet the practical challenge of being capable of artificially stabilising the retinal image if necessary.
(4) Possess an *accuracy* (relative position) of at least one percent or a few minutes of arc; e.g. not give a 10° reading when truly 9°. Accuracy is limited by the cumulative effects of nonlinearity, distortion, noise, lag and other sources of error.
(5) Offer a *resolution* (related to velocity) of 1 minute of arc per second, and thus be capable of detecting the smallest changes in eye position; resolution is limited only by instrumental noise.
(6) Offer a wide dynamic range of one minute to 45° (= 3000-fold) for eye position and one minute of arc per second to 800° per second (= 50,000-fold) for eye velocity.
(7) Offer good temporal dynamics and speed of response (e.g. good gain and small phase shift to 100 Hz, or a good step response).
(8) Possess a real-time response (to allow physiological manoeuvers).
(9) Measure all three degrees of angular rotation and be insensitive to ocular translation.
(10) Be easily extended to binocular recording.
(11) Be compatible with head and body recordings.
(12) Be easy to use on a variety of subjects.

It is useful to note that in considering the different methods available, that 10° is about a hand's breadth at arm's length. The subtense of the (rod-free) fovea is usually taken to be approximately 1° (0.291 mm of the retina), which also corresponds to the angular height of an average man at 100 mm. The cone spacing in the fovea may be taken as about 0.6 minutes of arc (= 2.5 μ at the retina). Therefore, a translational movement of 0.1 mm would be appreciably inaccurate if registered as rotation.

Angular resolution and temporal bandwidth are two parameters typically used to describe performance. Hallett feels that many authors specify 'resolution', which is easily established from the instrument's noise level, but few quote 'accuracy', which includes all sources of error and is difficult to estimate without comparison with an ideal (and *non-existent*) instrument.

The review of eye monitoring methods by Young and Sheena (1975) also describes the different methods. ASL, who produce eye monitoring systems, publish a document entitled 'How to evaluate an eye tracking system' (Borah, 1987). Carr (1988) describes some technical details of several available devices as well as giving further references of evaluation reports.

Eye movement monitoring offers considerable potential for studying user performance at the VDU and for optimising screen display design.

1.13.2 Limbus tracking

The clear boundary between the white sclera and the dark iris (the 'limbus') of the eye is a fairly conspicuous edge that can be optically detected and tracked in a range of ways. There are considerable limitations to direct observation methods of assessing gaze position. For instance, a human observer looking at the eye can be surprisingly inaccurate in determining where a VDU user is looking. With the aid of video recording (from a level parallel with the eye) and playback, it is possible to detect with 80–90 percent accuracy on a 4×4 matrix of a 12 inch VDU which cell a subject is gazing at. This means an accuracy of within only about $7°$ of visual arc (3 inches at 57 cm $= 7.5°$).

Almost all limbus tracking systems use infrared (IR) illumination, and all measure, albeit often indirectly, the position of the limbus relative to the photodetectors. For head-fixed photodetectors and illuminators, free head movement is possible (although contaminating), and the measurement is of the eye relative to the head. Fine IR beams, usually from an ophthalmist's spectacle frame system, are projected onto the eye. As the white sclera reflects more light than the coloured iris, a pattern of reflection is produced which changes as the eye moves horizontally. These changes are detected by similarly housed photodiodes and are converted into analogue or digital signals. Vertical eye movements can be detected through the changing reflections caused by eyelid movement, but recording up-down movement is problematic with this system. It is probably fair to regard limbus tracking as suitable for precise horizontal tracking only. Head movement artifacts are extremely difficult to avoid, particularly in a setting where the subject is being monitored for more than just a few minutes. Various ingenious devices such as bite-bars and dental wax impressions have been used to restrain head movement.

Reflection methods based upon absorption of visible or IR light by sclera and iris have some advantages but do not satisfy requirements (d), (e) and (g), while the recording of vertical eye movements is difficult and that of torsion is impossible.

1.13.3. Pupil tracking

The eye is again illuminated by an IR source to enable an IR imager to obtain an image of the pupil. By using various algorithms, the edges of the pupil can be automatically extracted and the centre of the pupil calculated. The axis through this centre is then related to the direction of gaze (Slater & Findlay, 1975), and careful calibration is required to know where it is directed. One advantage offered by the pupil over the limbus is that it is smaller and therefore unobscured by the eyelid for a much greater range of eye motion. In addition, the edge of the pupil is usually a sharper boundary than that between the sclera and the iris, resulting in higher resolution measurement. On the other hand, when viewed under normal illumination, the pupil appears black and therefore presents a lower contrast with its surrounding iris than the iris does with respect to the sclera. This results in it being a more difficult problem to automatically discriminate the pupil.

Many implementations of these two techniques yield good accuracies for a reasonable range. The output is an electrical record that can have a good frequency response. The frequency response is however limited by the video framework; one cannot sample movements less than 16–20 mm apart.

1.13.4. Corneal reflection

The corneal bulge produces a virtual image of bright lights in the visual field and region. Because the radius of curvature of the cornea is less than that of the eye, the corneal reflex moves in the direction of eye movement, relative to the head. As it only moves about half as far as the eye, it is displaced opposite to the eye movement relative to the optic axis or the centre of the pupil. Lateral head movement may contribute a large error when the head moves relative to the light. Therefore, many of the fixed head versions and the head-mounted devices need accurate stabilisation of the light source and recording device with a bitebar or head strap.

Two basic types of corneal reflex methods exist with differing locations of the light source; head-mounted or head-fixed corneal reflex techniques, and corneal reflex point-of-regard techniques. With the earliest method, the light source is fixed with respect to the subject's head. Therefore, relating eye position requires either a fixed head system, a method of recording head position (linear and angular), or a technique for recording the field of view relative to the head at every sample. The second method fixes the light source in the target field rather than to the head. With the light source in the target field, movements of the corneal reflex relative to the pupil indicate the point of regard of the eye in the field rather than relative to the head. Such techniques are much less sensitive to head position.

With a corneal reflex camera, accuracy of $\pm 0.5°$ is possible, but strict head fixation is required. Because it is subject to greater error from the relative movement of the light source with respect to the eye associated with head movements, the accuracy is poorer with a head-mounted corneal reflex camera, although $\pm 2°$ is possible.

The uncorrected linear range of all corneal reflex systems which employ a *single* light source for the reflex is limited to eye excursions of $\pm 12°$–$15°$, vertical or horizontal. Larger movements put the reflex in the non-spherical and higher peripheral portion of the cornea and require a complex (usually computer-generated) calibration and linearisation technique. Ultimately, the reflex range is limited by the size of the cornea and its partial obscuring from the lids. In addition to head movements, other factors that limit the accuracy of these methods to 0.5°–1.0° are variations of cornea shape and the production of other reflections by eye glasses. The output from these systems is typically graphic and as such requires measurement and/or recording. Any of those systems which produce a single bright spot on film or directly on a video signal, however, can be used to provide conversion to the x–y coordinates of that spot.

Crane and Steele (1984) note that corneal and limbus eyetrackers can measure very small eye movements, but their absolute accuracy is poor. This inaccuracy arises from eye translation movements, which, in these instruments, are indistinguishable from eye rotation movements. For example, in a corneal-reflection or a limbus eyetracker, 0.1 mm of eye translation causes approximately a 1° artifactual signal in the eye-rotation record.

1.13.5. Electro-oculography (EOG)

It was found, many decades ago, that the position of the eye could be measured by placing skin electrodes around the eye and recording potential differences (Mowrer, 1936). The source of the electrical energy is the corneoretinal potential or electrostatic field that rotates with the eye. The cornea remains 0.4–1.0 mV positive with respect to the retina. This is attributable to the higher metabolic rate at the retina. As the eye rotates, the electrical potential systematically alters. Skin electrodes placed on the outer canthi measure conjugate horizontal eye position. By reference to any third electrode over the bridge of the nose, a measure of horizontal eye vergence can also be detected. The recorded potentials are only in the range of 15–200 μV with nominal sensitivities of the order of 4 μV per degree of eye movement. The signals may be difficult to detect in the presence of large muscle-action-artifacts, which are also registered as potential differences by the skin electrodes (Shackel, 1967). Unless care is taken to shield the system, the presence of external electrical interference is troublesome.

EOG has the largest range of any of the objective methods practical for human studies as it does not require visualisation of the eye. The method is suitable for eye movements up to ±70°. Typical accuracy with surface electrodes is from ±1.5° to 2°. This method satisfies only requirements (d), (e) and (h) of the above. The chief sources of error are muscle artifacts, eyelid interferences, basic nonlinearity in the technique, and variation in the corneoretinal potential attributable to light adaptation, diurnal variations, and the state of alertness.

1.13.6. Purkinje image tracking

As light passes through the eye, reflections are produced at the various interface surfaces. At the surface of the cornea, there is corneal reflection or the first Purkinje image. A second occurs at the rear surface of the cornea, a third at the front surface of the lens, and a fourth at the rear surface of the lens where it interfaces with the vitreous humour. Two of these images (the first and fourth) are tracked and their displacements are used to calculate the eye movement.

The dual-Purkinje-image (DPI) method of eyetracking eliminates the translational artifact from the eye-rotation measurement. It is based on the use of a pair of reflections from optical surfaces in the front portion of the eye. These reflections move by the same amount with eye translation but differentially with eye rotation. By monitoring the spatial seperation of these two images, eye rotation can be measured accurately without being confused with eye rotation. A DPI eyetracker can measure eye movements with high frequency response and with an accuracy on the order of 1 minute of arc (Crane & Steele, 1984).

A different method for distinguishing between the translational and rotational components of eye movement is based on measuring the position of the corneal reflection with respect to the eye pupil (Merchant *et al.*, 1974). The advantage of this approach is that the instrument can be located relatively far from the subject. However, the pupil is not a stable reference, and accuracy is limited to approximately 30 minutes of arc. Because the system is based on video scanning, the temporal response is limited to about 40 ms.

In 1978, Crane and Steele described a three-dimensional eyetracker that comprised of an early model DPI which measured horizontal and vertical eye movements, and an optometer, which measured the refractive power of the eye. This has since undergone many improvements.

1.13.7. Contact lens/search coil methods

The most precise measurements of eye movements are made with one of several techniques employing some device attached to the eye with a contact lens. One contact lens system is the 'optical lever', in which one or more plane mirror surfaces ground on the lens reflects light from a light source to a photographic plate or photocell or quadrant detector. When a plane mirror normal to the visual axis is used, only rotation of the eye and the mirror will produce deflections of the projected image. Nevertheless, due to the high inherent accuracy of the system, very careful head stabilisation relative to the recording device is usually used. Hallett (1986) remarks that within the limits of their recording systems, contact lens methods provide the most technically satisfactory means of recording ocular rotations uncontaminated by translational artifacts.

The main non-optical contact lens measuring method is the search coil technique introduced (for use with primates) by Robinson (1963). An induction coil was mounted on a scleral contact lens. Around the head, a high frequency a.c. magnetic field was generated. Due to the 'transformer effect', an a.c. potential was induced in the ocular coil, with a magnitude and phase related in a simple manner to angular eye position. By appropriate amplification and phase-sensitive detection the coil voltage was changed into an analog voltage, proportional to the sine of the angular eye position. By generating a horizontal and vertical magnetic field in phase-quadrature or at different frequencies and using seperate detection stages, simultaneous measurements were accomplished. A special configuration of the scleral coil even allowed the recording of torsion. Resolution levels better than 1 minute were achieved.

The only drawback to this early technique was the need for the subject to wear a scleral contact lens held to the eye by suction. If not individually fitted, such lenses caused some bulging of the cornea and were uncomfortable. Slipping was hard to control and was aggravated by the forces exerted by the lids on the suction tube.

In order to apply Robinson's method to untrained human volunteers, Collewijn *et al.* (1975) developed a special carrier. This is a flexible, silicone rubber ring which fits on the limbic area, concentric with the cornea. It is concave at the ocular side with a radius of curvature (6.0 mm) smaller than that of the globe. The induction coil (9 windings of magnet wire of 0.05 mm diameter) is embedded within the material of the ring. The ends of the wire are connected to the detection device. The weight of the device is 0.1 g.

The subject is seated within horizontal and vertical electromagnetic coils (perpendicular pairs of hoops), and thus weak magnetic fields (perhaps 54 KHz) are generated around the subject. Movements of a smaller electromagnetic coil situated within this field would therefore generate an induced potential. Due to this principle, if that small coil happened to be situated within a contact lens, movements of the coil would correspond to horizontal and vertical movements of the eye. The system needs to be calibrated for offset (i.e., what changes in signal equals 'straight ahead'?) and gain (i.e., what changes in signal corresponds to a $\pm 10°$ swing?). The current changes equivalent to movements can then be led off via the fine wire from the search coil and then amplified. They may also be stored on magnetic media and computer analysed.

For application the surface of the eye is anaesthetised with a few drops of a local anaesthetic. The ring is wetted with Ringer's solution, placed on the limbus with the leads at the side of the outer or inner canthus, and pressed upon the eye. In this way fluid between eye and ring is evacuated and the elasticity of the ring causes a slight underpressure which holds the ring firmly in place without the need for further attention. The mounting procedure is facilitated by handling the ring with a suction implement. The device then generally only removable by lifting its edge; e.g. with a blunt forceps. Since the ring is void in its centre, it does not interfere with normal vision. Subjects may wear their own spectacles, if necessary. Wearing the device for thirty minutes or somewhat longer typically produces no problems.

1.14 PREVIOUS STUDIES OF EYE MOVEMENTS AND VDUS

In this section, a sample of previous studies are offered which have applied eye movement monitoring techniques to the investigation of visual search and the VDU. Other previous studies which concern visual search, VDUs and eye movements are described within the relevant chapters of this thesis.

1.14.1 General eye scanning studies

Zwahler and Adams (1987) describe two studies conducted to investigate the eye scanning behaviour of users under different tasks and screen conditions. Both studies used an original hardcopy for half of the work sessions and a split-screen presentation (original information/workspace) for the other half. Two different screen polarities were also used. Results showed that the operators spent more than twice the amount of time looking at the screen during the split-screen presentation when compared to the hard copy presentation for both data entry and file maintenance tasks. Transition matrices showed the number of times the subjects looked from one screen area to another. Not surprisingly, the greatest amount of transitions were found to occur between the screen and the document, and vice versa. It was concluded that there exists a considerable amount of intra- and inter-subject variability and that apparently many visual strategies can be used (e.g. more fixations of a shorter duration versus less fixation of a longer duration) to maintain a fairly constant level of operator performance. Furthermore, light characters on a dark screen or dark characters on a light (green) background do not appear to affect any of the eye scanning measures in a statistically or practically significant manner.

1.14.2 Geometrical characteristics of displays

Using ten differently structured display formats containing the same information, Graf *et al.* (1987) investigated the relationship between 'geometrical' characteristics of a display, search time, subjective ratings and eye movements. The four basic geometrical characteristics which were suggested by Tullis (1983) which may affect how well users are able to extract information from a display are:

(1) **Overall density.** The number of characters displayed, expressed as a percent-

age of the total spaces available.

(2) **Local density.** The average number of characters in a 5° visual angle around each character.

(3) **Grouping.** The extent to which characters on the display form well defined perceptual groups.

(4) **Layout complexity.** The extent to which the arrangement of items on the display follows a predictable visual scheme. It was found that there was a high correlation between the geometrical characteristics of a display and the search time for an item on the display. Therefore, it was possible to predict search times according to geometrical characteristics. Furthermore, it was shown that the eye movement parameters (i.e. saccade amplitude and fixation time) correlated partly with the geometrical parameters.

1.14.3 Layout and cognitive load

The effect of the layout of alphanumeric VDUs as well as the effect of different cognitive loads on eye movement parameters was investigated by Graf and Krueger (1989). Subjects were required to answer questions concerning the information on a VDU. Thus, the task may be regarded as consisting of both reading and searching components. The layout of the displays and the difficulty of question (cognitive load) was varied. The results showed that 'the investigation of a man-machine-interface using eye-movement data (for example during the prototyping phase) offers a measure of performance which allows an evaluation of the layout of a user-interface as well as an evaluation of the cognitive load of users' (p. 659).

1.14.4 Readability

Kolers *et al.* (1981) recorded eye movements as people read texts presented on a VDU in two different spacings, two different character densities, and at five different scrolling rates. Differences in efficiency of reading single- and double-spacing were statistically significant, although of little practical significance. Character densities of 35 characters or 70 characters per line favoured the smaller size character with respect to efficiency of reading. A comparison of scrolling rates suggested that the static page was processed more efficiently than was the page scrolled at the subject's preferred rate or at a rate 10 percent slower than that, and pages presented faster than the preferred rate were read more efficiently. Kolers *et al.* suggest that scrolling has certain facilitative effects on reading, above and beyond the 'pleasantness' that Oléron and Tardieu (1978) have ascribed to it.

> Scrolling can induce people to read more efficiently (i.e., with fewer and shorter fixations), thus taking less time to read and with no apparent loss in compre-

hension. The rate at which this facilitation occurs deserves extended consideration. If scrolling occurs at less than the reader's preferred rate, the reader makes more and longer fixations. As the rate gradually increases, performance increases correspondingly, with best performance found at a rate 20% faster than the preferred scrolling rate studied in this experiment. (Kolers *et al.*, 1981, p. 526)

Results also showed little, if any, change in preferred rate occurring as a function of practice with 16 pages of text. Kolers *et al.* concluded that: Systems in which, by program control, text was presented at rates ten or twenty percent faster than the preferred scrolling rate should lead to more efficient performance, but might create some problems of user acceptance.

1.14.5 Menus

Lee and MacGregor (1985) presented a model, stemming from eye movement work, which assumed that when users examine menu pages, they start with the first item and work sequentially through the list. They assumed that this sequential search might take two forms: (a) the user might examine all **a** options before making a decision (exhaustive search), or (b) he might stop on reaching an item which is judged to be correct (self-terminating search).

In a later extention of this approach, MacGregor *et al.* (1986) proposed that menu search is governed by a criterion-based decision process, in which a menu item is selected if it exceeds a subjective level of probability/confidence. 'In this view, the number of menu items examined prior to a selection can vary on a continuum from 1 to some multiple of **a**. At the lower end of the continuum lie "direct searches", where a highly practiced user goes immediately to the desired item. At the upper ranges of the continuum, users may examine items repeatedly, resulting in "redundant searches". However, as in the original model, this more recent approach assumes that the search process is essentially systematic' (MacGregor & Lee, 1987, p. 627).

However, several papers have indicated that menu search is done randomly rather than by systematic eye movements (e.g. Parton *et al.*, 1985). A further paper (Giroux & Belleau, 1986) proposed that random searches may occur with command menus, whilst sequential searches occur with database menus. To help clarify this situation, the MacGregor *et al.* (1986) data were re-analysed by MacGregor and Lee (1987). In the earlier study, subjects had viewed items in strict sequential, rather than simultaneous, order. The question was then whether, when the cumulative probability of selecting a target was plotted as a function of time, an exponential (indicating a random search) or a linear function (indicating a systematic search) was obtained. The obtained relationship was closely approximated by an exponential function, which accounted for 99 percent of the variance. The authors concluded that their results demonstrated that 'strict sequential search not only can, but in this case did, produce an exponential function of the type thought to be associated with random search' (1987, p. 629).

It seems likely that the form which a search takes will be influenced by the nature of the task. Specifically, factors which reduce uncertainty about which is the correct item, and where the item is located upon the page, are likely to affect how users examine items. The following statement serves to emphasise that little is yet known about user search processes in menu retrieval.

> Such factors might include the simplicity/complexity of the items themselves, how they are organized on a page, and how often users have seen the page. For example, frequent users of software command menus might learn the layout of pages sufficiently to be able to conduct "direct searches", where they fixate immediately a desired item. Such searches would be essentially systematic, with a purposeful direction of gaze towards the approximate position of the target item. Some component of random error might be involved in the process, requiring a random search within the immediate area of the target. But the process itself would be intentional, with a small random component, rather than random. (MacGregor & Lee, 1987, p. 630)

The issues of random versus systematic searches, as well as many other notions relating to the vast area of visual search, are discussed within the next chapter.

1.15 POTENTIAL STUDIES OF EYE MOVEMENTS AND VDUS

The study of any aspect of HCI contains numerous potential variables, at both the machine side and the human side. The aim of this thesis is to explore aspects of visual search in relation to VDUs. However, to do justice to the research area it was felt inappropriate to pursue (as is typically so within theses) one unitary research line. A valid examination of optimising visual search at the VDU must include parameters such as the overall array presentation (e.g. screen shape; see Chapter 3), how information may be optimally presented either in central/foveal vision (e.g. cursor presentation; see Chapter 4) or in peripheral vision (e.g. status bar information; see Chapter 5), when and why icons may be preferred to verbal labels (see Chapters 6 and 7), and how 'newer' font designs can enhance locatability of items and perhaps VDU-reading generally (e.g. antialiased fonts; see Chapter 8). The general experimental approach chosen is that of employing both relatively 'pure' search tasks (e.g. Expts 1 & 10) and relatively 'real' search tasks (e.g. Expts 4 & 6), and supplementing these with a fine analysis of the eye movements involved in order to provide an understanding of the underlying procedures which would otherwise have remained unknown or at best only inferred. The following chapter therefore provides an in-depth discussion of visual search.

Chapter Two

A Review of Visual Search

This chapter concentrates on general and theoretical aspects of visual search. Topics include attention, eye movements and search modelling. The chapter starts by looking at some of the early work on visual search.

2.1 EARLY WORK ON VISUAL SEARCH

2.1.1 The work of Neisser

In this sort of visual processing task, or 'continuous visual search task', an observer is instructed to look for a particular letter in a large list of letters and to press a button as soon as it is located. Such a requirement may also be regarded as a matching task, where a representation of a sought-for letter is matched with the visual representation of the letters within the list. Neisser and collaborators have reported several experiments of this nature (e.g. Neisser, 1963, 1964; Neisser *et al.*, 1963; Neisser & Lazar, 1964; Neisser & Beller, 1965; Neisser & Stoper, 1965). In a typical situation, 50-line lists containing a target in a random position are presented and as soon as the list appears the subject begins to search down the list as rapidly as possible. Figure 2.1 (*a*) shows an example where the letter K is the target and the *critical item* is the row which contains it ($ZHFK$). The processing time per row is calculated by dividing the time required to find the letter by the number of letters in the list. Figure 2.2 shows typical results. The slope depicts a linear increase in search time with target position. The function of this line may be taken to represent the average processing time per row, and this parameter should reflect the complexity of the information-extraction process.

a	b	c	d
EHYP	ZVMLBQ	ODUGQR	IVMXEW
SWIQ	HSQJMF	QCDUGO	EWVMIX
UFCJ	ZTJVQR	CQOGRD	EXWMVI
WBYH	RDQTFM	URDGQO	VXWEMI
OGTX	TQVRSX	GRUQDO	MXVEWI
GWVX	MSVRQX	DUZGRO	XVWMEI
TWLN	ZHQBTL	UCGROD	MWXVIE
XJBU	ZJTQXL	DQRCGU	VIMEXW
UDXI	LHQVXM	QDOCGU	EXVWIM
HSFP	FVQHMS	CGUROQ	VWMIEX
XSCQ	MTSDQL	OCDURQ	VMWIEX
SDJU	TZDFQB	UOCGQD	XVWMEI
PODC	QLHBMZ	RGQCOU	WXVEMI
ZVBP	QMXBJD	GRUDQO	XMEWIV
PEVZ	RVZHSQ	GODUCQ	MXIVEW
SLRA	STFMQZ	QCURDO	VEWMIX
JCEN	RVXSQM	DUCOQG	EMVXWI
ZLRD	MQBJFT	CGRDQU	IVWMEX
XBOD	MVZXLQ	UDRCOQ	IEVMWX
PHMU	RTBXQH	GQCORU	WVZMXE
ZHFK	BLQSZX	GOQUCD	XEMIWV
PNJW	QSVFDJ	GDQUOC	WXIMEV
CQXT	FLDVZT	URDCGO	EMWIVX
GHNR	BQHMDX	GODRQC	IVEMXW

Figure 2.1: Portions of lists for visual search. From 'Visual search', U. Neisser.

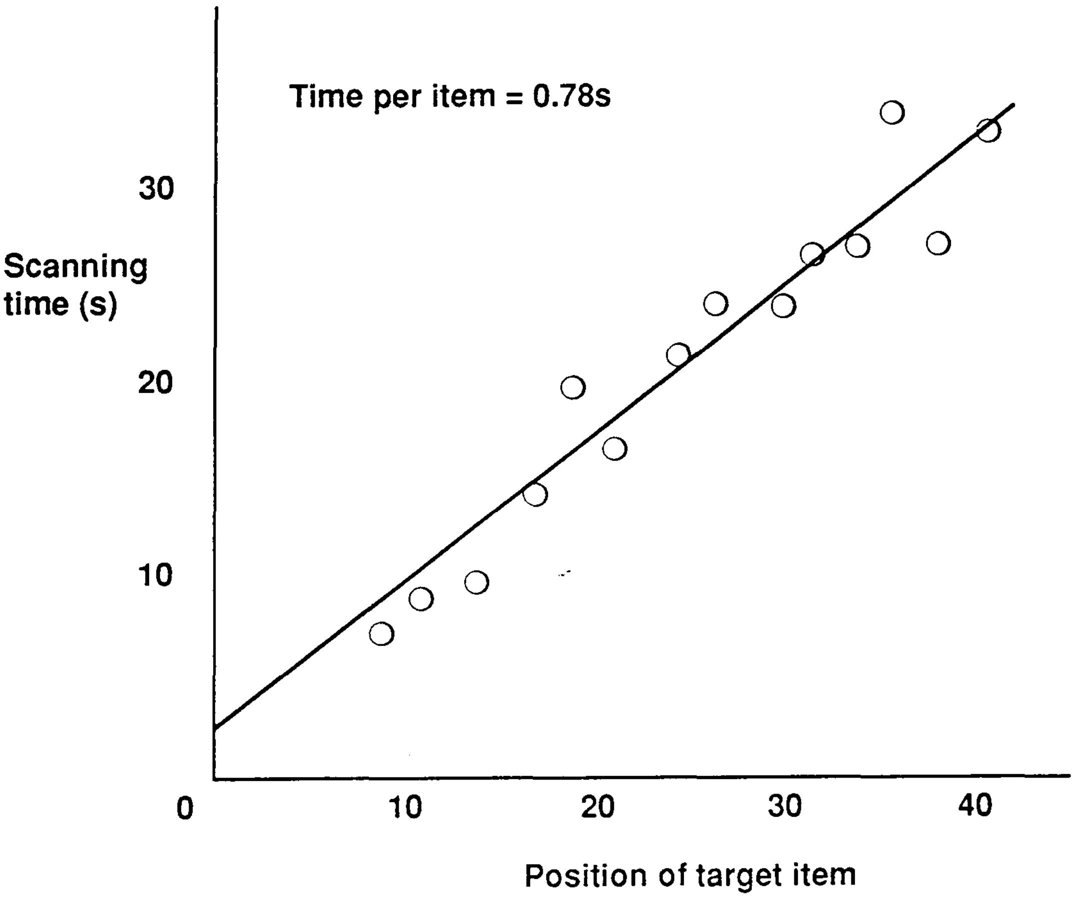


Figure 2.2: Typical results from a visual search experiment. The data used here were for a search by an inexperienced subject whose time per item was 0.78 s. Reproduced with permission from Neisser (1963).

Many parameters of this task have been varied. For example, it was found that searching for the presence of a letter (e.g. the letter *K* in list *a*) was faster than looking for the absence of a letter that was present in all but one row (e.g. searching for the absence of a *Q* in list *b*). Furthermore, it is easier to find a *Q* within a background of angular letters (see Fig. 2.1*c*) than within round letters; and conversely, a *Z* amongst round letters than angular letters (see Fig. 2.1*d*). If the features that discriminate targets from non-targets are easy to differentiate, search is quicker than if the targets and non-targets require more complex identification. Such results suggest that searching can be performed on the basis of visual-feature characteristics, and therefore that information at this level must have been available (Haber & Hershenson, 1980).

Neisser (1963) also investigated what happens in the search task when the observer is required to search for more than one target simultaneously. For instance, rather than looking only for a K in list a of Figure 2.1, he may need to search for a Q or a K, and to respond when an instance of either is found. Without practice, observers take longer per row when they have to search for more letters (although not in a linear-like three times longer for three letters). After many days of practice, however, observers were able to search down the list at the same speed regardless of how many targets they were searching for. One of the many explanations offered for this improvement with practice in multiple-target search tasks is a change from a serial feature-by-feature processing, in which each feature of each target is examined one at a time, to one in which several targets are processed at the same time. 'This might be due to a change in the strategy of the perceiver or to a change in the familiarity of the target items that permits them to be grouped or chunked together into a more meaningful unit, thereby reducing the number of seperate items to be processed' (Haber & Hershenson, 1980, p. 357).

2.1.2 The work of Bloomfield

Neisser and his followers were more interested in the (hypothesised) mental processes underlying search than with visual processes. A different tradition, and one more relevant to this thesis, was developed by Bloomfield (1970, 1975). For most of the experimental work reported by Bloomfield, a competition task was used, with the target having to be located when other confusing, non-target stimuli were also present in the display. An adaptable apparatus was devised. This allowed a series of extensive studies to be undertaken with well practiced observers, with the following results:

(1) As the difference in target and non-target stimuli decreased search times increased.
(2) More time was required to find a particular target in a homogeneous display than in a heterogeneous one made up of stimuli all larger or all smaller than the target. However, if some non-targets are bigger and some smaller, much more time was required for the heterogeneous displays.
(3) As the density (and number) of non-targets increased search times increased.
(4) An irregular arrangement of non-targets took longer to search than a regular one.
(5) An incentive payment training schedule produced very large improvements in search performance.
(6) As the amount of target information available was decreased the time to find difficult targets was increased with no change for easy targets for some observers, while the reverse occurred for others.

(7) Varying the shape of the display area made some differences to search time.

To take the *shape* of the search display as an example, Bloomfield described three experiments in which the shape was varied. In all three the displays varied in their horizontal and vertical extents. This issue is returned to in depth in Chapter 3.

2.1.3 A systematic description of search variables

Gottsdanker (1960) listed some of the many factors involved in search. He was mainly concerned with stimulus determinants and he did not include all other variables which might affect search behaviour. His terms, with the equivalent terms used by Bloomfield in parentheses are given below. Firstly, there are *search determinants*: interposition (obstruction), smallness (relative size of target), weakness (threshold), distortion (distortion), imbeddedness (camouflage) and competition (competition). Secondly, there are *search goals* (target specification) and *aids* (search area). Gottsdanker also cites strategies, although clearly an observer's search strategy is dependent upon almost all the variables listed here. In trying to understand search we are trying to understand search strategies.

2.1.4 Search determinants

Search is necessary when, for some reason, a target cannot be immediately located. Bloomfield (1970) cites several reasons why search may be necessary and, in a particular search situation, one or more of them will be operating.

1: Competition Here the target is readily distinguishable from its immediate background and is difficult to detect because it is confusable with non-target stimuli also present in the search area.

2: Camouflage In this case, the target fails to emerge perceptually from its immediate background. This failure is caused by the patterning of the background and the target combining to obscure the target.

3: Threshold As in the camouflage case, the target is not easily detectable as it is embedded in the immediate background. There is a low contrast between target and background with the target being near threshold in a plain or, perhaps, a noise background.

4: Relative size of target As with competition, the target is clearly distinguishable here from the immediate background, although it remains difficult to detect. This is so in this case because it is very small relative to the total search area.

5: Orientation Detection may be slow because the target appears different or unusual due to being viewed in an unusual orientation or from an unusual angle. This is a relatively unimportant determinant in many search situations, since the difficulty will be reduced as the observer becomes increasingly familiar with the viewing angle or orientation.

6: Distortion The target may be distorted by the optical system or the atmosphere through which it is viewed. The effect of distortion is again likely to be reduced with increasing familiarity.

7: Obstruction The target may be partially or even completely obscured. Visual search alone may not be sufficient, though it may enable the observer to deduce the most likely target location. For a positive detection, visual search may have to be augmented either by bodily or manual search; i.e. the observer may need to move himself or the obstruction. As in the case of orientation and distortion, the effect on search of obstruction is likely to be reduced with practice.

2.1.5 The observer

Many observer variables can be identified. Those most relevant to search include the following.

1: Visual acuity Eyesight may be tested in several ways including 'minimum visible' (i.e., detecting a single line or dot) and 'minimum seperable' (i.e. the resolution of interspaces between contours), but the most familiar measure of visual acuity is that obtained with a Snellen chart. However, search deals also with peripheral vision and glimpses which will involve aspects not normally tested as eyesight. For a fundamental understanding of search it is necessary to learn exactly what can be seen in a single glimpse of fixation, and how glimpses are related to each other to produce the visual complex experienced by the observer.

2: Experience The effects of past experience on present behaviour are difficult to control with any precision. However, it is possible to investigate the effects of varying the observer's immediate past experience, and to look at specific training, practice and expectancies which might have been built up in previous search trials.

3: Motivation It is expected that altering the level of motivation will have a considerable effect on search behaviour. The general level of motivation will depend on factors such as the observer's personality, their general state of well-being, and the observer/experimenter relationship; factors not usually investigated in search experiments.

4: Age The main effects on search of the age of the observer will be mediated through the variables already mentioned: eyesight and experience. In children, increases with age are likely to be associated with increases in experience of searching activity. Thus, at this age level, the older the observer is the better he should be. In the elderly, however, increases in age will probably be associated with failing eyesight. At this age level, the older the observer the worse he is likely to be in search tasks.

2.1.6 Task variables

1: Movement The search situation may be entirely static, or there may be movement of the observer, the target, the background or some combination of the three. In addition, the movement may be in various directions and of varying speed and acceleration. Changes in movement are likely to lead observers to use a wide variety of strategies for search.

2: Optical environment The optical system and atmosphere through which the search area is viewed must be clearly stated, since great changes in performance may result from changes in the optical environment.

3: Nature of search trials Many search tasks consist of discrete trials in which there are a known number of targets. Much experimental work has been carried out using a single target. In real life, the search task is often continuous and may, in fact, be a vigilance task also as far as the observer is concerned. There is an extensive literature on vigilance which should not be ignored in continuous search tasks, particularly where the amount of stimulation given to the observer is limited.

4: Target specification The observer's approach to a search task will vary with his knowledge of the situation and, in turn, his knowledge will depend on a combination of past experience (above) and briefing. The amount of information given in briefing varies in several ways. It can be given visually or verbally. It may be vague, e.g. 'The target is different from other stimuli', or precise, e.g. 'The target is a bright red *X* in a background of dull brown *O*s'. It also may relate to the total area in which the observer will have to search.

5: Search area and aids The degree to which search area is defined, the method by which the boundaries are marked, and its size and shape all may vary. In situations where search aids such as cues or sector markings are found to be of assistance they probably reduce the effective search area.

6: Time limits A time limit is imposed in some search tasks. This may occur in laboratory and natural search situations and may affect performance.

7: Search measures Performance in search tasks can be viewed in two ways. It may be measured in terms of success in finding targets or, alternatively, the pattern of search can be examined. The best measure of the former is 'search time', which is the length of time between the moment that the observer begins to search to the moment that he indicates that he has located the target. Other 'success' measures may be used: e.g. the number of exposures, of a fixed short interval, of the search area needed before the target is found; or, in tasks where only a limited search time is given, the proportion of successful trials. The pattern of search is best investigated by recording the observer's eye movements during search, and then extracting the spatial relationships of consecutive fixations, the length of fixations, and/or a classification of the type of area fixated. When no eye movement apparatus is available or if considered invasive, some indication of the observer's search strategy may be obtained by comparing search time with target position, or by simply eliciting a verbal protocol.

2.1.7 Maximum search rates

In experimental investigations of visual search in computer-driven visual displays, Sperling *et al.* (1971) studied visual search with many different presentation rates in addition to those that most closely approximated natural search. They discovered that the most rapid visual search actually occurred when new arrays were presented every 40 ms, five times faster than the fastest possible saccade rate. At these artificially high presentation rates, search proceeded at a rate of one background character per 10 ms, which is about twice as fast as Neisser's maximum rate and twice as fast as in the 240-ms presentation rate that simulated Neisser's conditions. There was only a small difference in detection accuracy between inter-array times of 120 and 240 per second, which suggests that in some natural searches the motor control of the eye is actually the limiting factor. In Neisser and colleagues' search task, if their subjects' eyes had executed saccades every 120 ms, search rates may have doubled with little impairment in accuracy. Data suggest that the second half of many fixation pauses may have been wasted waiting for the eyes to move.

In contrast to Neisser's lists, the computer-generated arrays of different sizes are searched at similar rates (characters per second). There is also a considerable trade-off possible between scanning characters in one array or in several. Therefore, almost as many background characters can be scanned in one array presented for 120 ms (12) as in three arrays, each presented for 40 ms (4 per array).

2.2 BASIC SEARCH THEORY

For an adequate theory of search, Bloomfield (1970) regarded several characteristics as being necessary:

(1) A specification of the area in which the target can be detected in a single fixation for various target/background combinations.
(2) A description of the relation between consecutive fixations; knowledge of the extent to which fixation coverage is exhaustive and to which it is efficient.
(3) Knowledge of how fixation patterns are affected by experience, briefing and cues.
(4) An ordered description of target/background complexes.

A model of search, derived from Bloomfield's 'basic search theory', was found to handle much of the data for competition and non-competition search situations. The model gave equations predicting that search time is a function of fixation time, total search area, and the area within which the target can be seen in a single fixation. The latter term could be related to measurable characteristics of the target and background stimuli. Each term in the equation was separately determined, both for competition search and for threshold search, where a near threshold target had to be found in a uniform, unstructured display.

Finally, the model assumed that observers search in a methodological way, and that the degree of efficiency with which they do this depends upon the amount of overlap between successive fixations. The earlier search equations for exhaustive efficient search were amended, although those based on an independent fixation strategy were abandoned. A cumulative probability equation accounting for response time factors, fixation overlap and the characteristics of targets and non-targets was given.

Bloomfield (1970), in his 'macro-analysis', concluded that much more theoretical thinking is needed before a complete understanding of visual search is achieved and this must be closely related to continuing empirical work.

2.3 PERIPHERAL VISION AND VISUAL LOBES

Over the last decade, there has been considerable interest in studying performance differences between central and peripheral vision in various visual tasks, and relating these differences to the anatomy and physiology of the visual system. Two types of tasks have been extensively investigated at different eccentricities: (a) *grating detection tasks* (e.g. Hilz & Cavonius, 1974; Koenderink *et al.*, 1978; Virsu & Rovamo, 1979; Virsu *et al.*, 1982; also see Chapter 7), and (b) tasks requiring not only the detection of a stimulus pattern but also the processing of positional relationships between pattern elements (e.g. Westheimer, 1982; Levi *et al.*, 1985; Rentschler & Treutwein, 1985; Levi & Klein, 1986; Virsu *et al.*, 1987).

Peripheral vision plays a role in guiding subsequent eye movements to informative regions of a scene (e.g. Antes, 1974). Generally speaking, information is absorbed during a fixation when the control of length and direction of the next saccade is also determined. Much research has been devoted to the details of visual perception during this immobile period. In this section, some background knowledge to the notion of visual lobes is presented. These following two sub-sections rely largely on the chapter by Widdel (1983).

2.3.1 Visual lobes

The peripheral area around the central fixation point from which specific information can be extracted and processed is the 'visual lobe area' or 'useful field of view'. The search field is approximately circular, centred slightly above fixation. However, locations with fewer neighbours or with adjacent blank spaces are easier to search (e.g. Harris *et al.*, 1979) so that the measured search field is distorted by boundaries of the stimulus. The search field depends on the stimuli used to measure it, with extremely rapid presentations or extremely small size characters shrinking the search field. However, these parameter variations do not necessarily alter the *shape of the search field* which suggests that the search field is an invariant property of the visual system. Obviously, the search field pattern in part reflects perceptual limitations. Nevertheless, the spatial distribution of attention can be voluntarily altered (e.g. Lambert *et al.*, 1988), so that the search field also reflects voluntary, cognitive factors (Sperling, 1989).

Among others things, the size of this area depends on characteristics of the visual target and the surrounding context. When the target is embedded in a complex background or surrounded by irregularly positioned non-target items, it is relatively small (Brown & Monk, 1975), as it also is when the density of non-target items and the visual lobe increase (Mackworth, 1976) and when the similarity of target and non-target items is high (Bartz, 1976).

Acuity decreases from the fovea across the visual field to the periphery (e.g. Anstis, 1974). Most studies have measured the decrement either of detection or of discrimination capacity in static situations in which the observer maintains his sight upon a central fixation point and a stimulus is presented in the peripheral field of vision. In contrast, interest in other studies lies in their assessment of visual material input collected during the fixation while the eyes are scanning the displayed material. For instance, Widdel (1983) recorded the spontaneous pattern of ocular scanning of scattered elements displayed over a wide area, amongst which a single target has to be located. Of particular interest is the final part of the scanning pattern in which a saccade carries the gaze from the current position to the target. Data reveal that the target is reached from a greater distance when the density of the display is lower, which Widdel interprets as a correlate or a measure of mental work load. As Lévy-Schoen (1983) points out, it could also be the effect of lateral interference reducing the discriminability of the target in peripheral view, which might result secondarily in a trade-off between density and area, keeping constant the amount of material successfully analysed.

This useful field of view within a dynamic scanning task, termed 'visual lobe area' by those such as Widdel has also been investigated by others such as Prinz (e.g. 1983) under the name of 'control area'. Prinz provides evidence showing a vertical asymmetry in favour of the lower part of the field where targets have a higher probability of detection at a larger distance than in the lines above the line being scanned. Several reasons have been proposed for such an asymmetry. A so-called 'structural' source of asymmetry could be augmented sensory sensitivity in the lower field. It seems more reasonable to consider the asymmetry of the scanning task as a powerful incentive to emphasise downward rather than upward perceptual efficiency. Prinz calls this the 'functional source of asymmetry', as opposed to the 'structural' one. It is postulated that this functional asymmetry depends on attentionally controlled operations only effective in central vision. Of particular interest is the possibility of obtaining a change in the capacity for target detection around each fixation point by changing the subject's scanning strategy; e.g. using a bottom-up sequence rather than the normal top-down sequence. The experiment shows some effects that can be attributed to the way peripheral information is used while scanning rather than to its input capacity (Lévy-Schoen, 1983).

2.3.2 Measuring the visual lobe area

The visual lobe area is usually measured with tachistoscopical techniques, and this includes the work of Mackworth (1976) who measured the visual acuity limit in a peripheral discrimination task, and that of Engel (1977) who analysed what he called the conspicuity area. This relatively simple method provides a favourable setting for the analysis of the shape of the visual lobe area because it is not influenced by eye movement strategies and dynamic processes dependent on stimulus material. As indicated by Enoch as long ago as 1959, the shape has a longer horizontal than vertical extension. However, as Bellamy and Courtney (1981) subsequently discovered, the shape also depends on target characteristics. One possible disadvantage of this technique is that measurement is based on static fixation behaviour and not on eye movement activities. A similar problem exists with the modified technique used by Mackworth (1976), in which stimulus material moving through two windows was presented. The distance between the windows was altered and subjects has to fixate on one window to detect targets which could appear in either window.

A second approach for measuring the visual lobe area involves the analysis of the interfixational distances, and takes into account the dynamic behaviour of eye movements during a visual search task. This method has been used by several investigators, e.g. Snyder and Taylor (1976) and Megaw and Richardson (1979). It is useful for specific research questions, and whilst not a direct measure of the visual lobe, it does allow a deduction of its average size. A disadvantage of this method is that it is illogical to measure the visual lobe size for specific targets or target characteristics. A special measurement technique of analysing the visual field, as used to estimate mental work load, has been evaluated by Voss (1981). Voss analysed performance in peripheral vision by presenting peripheral light stimuli for detection. Lights were presented with a spectacle frame fixed on the head to eliminate the effects of head movement. The calculation of the detection rate of the light stimuli during car driving shows that the functional visual field is sensitive to external load. This method represents a valid new indicator of mental work load, although visual field sensitivity was a secondary task and it is not regarded as appropriate for analysing specific aspects of visual performance.

A novel method which combines advantages of the techniques described for measuring visual workload is also presented by Widdel (1983). The philosophy is that during a 'successful' fixation in a search task, the target item can be detected in the peripheral area. The following saccade is determined by this peripheral stimulus and the next fixation will attain the target. In these terms, a successful fixation is one which is immediately followed by the fixation of the detected target. During a visual search task a subject will be executing n fixations, and generally $n > 1$. When the detected target is fixated the fixation n is identical with the target fixation and the fixation $n - 1$ is termed a successful fixation because the target was detected peripherally (see Fig. 2.3).

In a series of search trials the distances of the fixations $n - 1$ to the targets can be measured and the distributions of the distances can be computed. Those fixations in the vicinity of the target have to be analysed during the time in which the target is not being detected peripherally. When no target fixation follows, or a target fixation follows without detection of the target, the fixation is considered to be 'unsuccessful'. During an unsuccessful fixation, the target is not detected and the length and direction of the following saccade is not determined by the target location. It is possible that hitting the target may be a random event, and it cannot be prevented that such a fixation would be falsely judged as successful (although by definition an unsuccessful one) because of lack of peripheral detection of the target. Such random hits will occur in all experimental situations and measurement procedures, and therefore will be balanced in comparing different visual lobe areas.

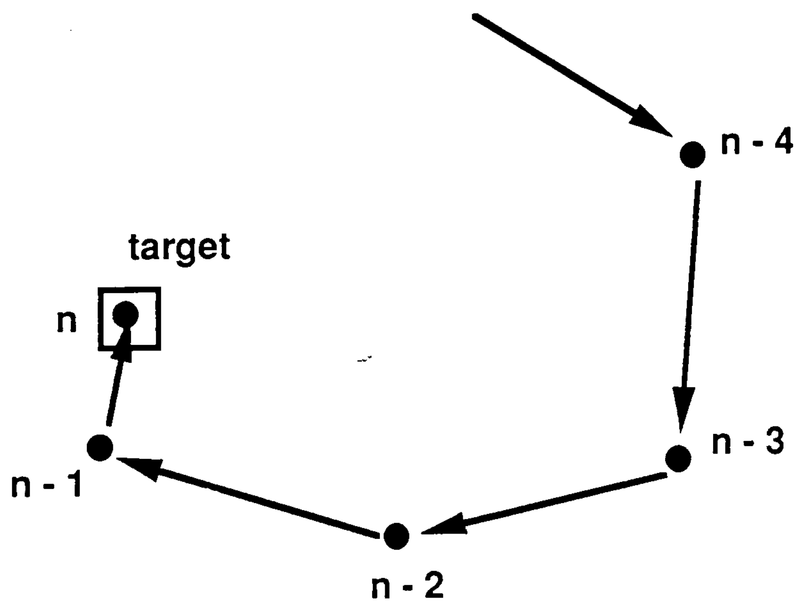


Figure 2.3: Last five fixations of a search path. Fixation n is the target fixation and fixation $n - 1$ the successful fixation. During fixation $n - 1$ the target is detected in peripheral vision and the saccade to fixation n is determined. (Adapted from Widdel, 1983.)

One technique of quantifying the visual lobe area is in terms of

$$\frac{S}{(S+F)} \times 100 \tag{2.1}$$

where the number of successful fixations (S) is divided by all fixations ($S + F$), which include the unsuccessful fixations (F), for each unit of distance from target. Measurement of the visual lobe area must consider the search area near the target. The outer limit of this area has to be established empirically, is determined by the fact that the frequency of successful fixations approaches zero, and is expected to have an approximately circular shape.

2.3.3 Control areas in visual search

Experiments reported by Prinz (1983) explore the size and shape of functional visual fields; i.e. of control areas in a continuous visual search task. The results of these experiments generally suggest that: (a) control areas can be asymmetrical on the vertical dimension; and (b) that control area shape is determined by both functional (attentional) and structural (sensory) factors, which can be complementary. These studies have used a visual search task of the Neisser type in which the subject must locate a target letter in a random display. Subjects are instructed to scan the list row by row, as in reading. In this sort of experiment, subjects frequently report that they detect the target prior to scanning the row which contains it. For example, the subject may detect the target at a certain location in row 9 while scanning through row 6 and fixating some location in that row. In such a case, the target would be detected at a (vertical) distance of three rows. Some of the data observed can be summarised as follows:

> (1) Vertical detection distances range between about 0–6 degrees of angular distance. A detection distance of 0 implies that the target is not detected prior to scanning the target-containing row itself. A distance of 6 degrees implies that the target is detected about 6–10 rows prior to scanning the target-containing row (the exact value depending on list size and viewing conditions, of course).
> (2) Detection distances are clearly related to target-non-target discriminability. For example, an angular letter (such as *A*) has a larger detection distance if it is embedded in a 'round' non-target context (composed of, e.g. *O*, *C*, *B*, *G*) as compared to an angular context (composed of, e.g., letters such as *Z*, *N*, *T*, *X*).
> (3) When several targets are to be searched for simultaneously, the targets clearly differ in their mean detection distances. For example, when the subject has to search for *A* and *Q* simultaneously in an angular non-target context, *Q* will be detected at much larger vertical distance than *A*.
> (4) There are substantial individual differences in variation of detection distances. In a given simultaneous search task, there are always some subjects who show a random fluctuation about the zero distance (even for 'easy' targets), and some others who virtually show clear indication of parafoveal detection (even for difficult targets). (Prinz, 1983, p. 83–84)

From these results, Prinz suggests that subjects are able to exert some control over the detection of a target not only in the row just fixated, but also over some larger area around the fixation point, depending on the type of non-target context. According to this view, the mean detection distance for a given target is an estimate of the vertical extent of the subject's control area for that target. Because this method does not permit recording detection distances above the fixation point, however, this estimate is restricted to the area below the row just fixated, and there is no indication of the extent of the control areas above the fixation point.

Prinz also suggests that one of two views may be adopted in dealing with this problem. The first view assumes that control areas are basically concentric in shape, i.e. approximately symmetrical in their vertical extent. Within this view, the upper halves of the control areas virtually possess the same target detecting capabilities as the lower halves. By contrast, the second view assumes that the control areas are symmetrical under the conditions of this task. 'As subjects are never required to watch for a target above the fixation point it is reasonable to assume that they either learn to concentrate their attention—and thereby their target detecting capabilities—on the zone below the fixation point (leading to an attentional neglect of the upper zone) or that they learn to adapt the sensitivity of various sections of the visual field according to the demands imposed on them by the actual task (leading to a decrease of sensitivity in the upper zone) (Prinz, 1983, p. 87).

2.3.4 Visual conspicuity

In everyday life few visible objects around us are actually observed. Whilst some objects do strike our attention, most objects are overlooked unless our attention is directed towards them. Thus, attention performs the process of information selection. The factors influencing this selection process can be divided into object factors and subject factors, or external and internal determinants of attention. Visual conspicuity may be considered an objective factor—at least in relation to background. For example, a red ball surrounded by similar red balls is inconspicuous, whereas in other surroundings it may be conspicuous. Visual conspicuity may be operationally defined as that combination of properties of a visible object in its background by which it attracts attention via the visual system, and is seen in consequence. The relative relationship to the background is paramount.

Various methods for measuring properties related to visual conspicuity have been reported. For instance, in early experiments, Engel (1971) determined the mean number of tachistoscopic exposures required for a correct localisation of a test object in several combinations of test object and background. By measuring the eye fixation points of the subject during the exposures, it was found that in contrast to conspicuous objects, inconspicuous ones had to be fixated very closely in order to be reported correctly. More or less random search behaviour would consequently be circumvented by determining the retinal locus within which the object to be searched was noticed in 15 ms exposure. This area was termed the 'conspicuity area'. As the conspicuity area expresses the chance of seeing the object during search, it was suggested that its size can be used as a measure of the visual conspicuity of the relevant objects in its background. Engel regarded it as likely that the conspicuity area concept can also be applied as a measure of more conspicuous combinations, and for other visual conspicuity factors such as size, colour contrast, density of background elements, and so forth.

The concept of conspicuity area comes close to the *field of short term visual search* proposed by Chaikin *et al.* (1962). The conspicuity area may be useful in understanding search behaviour. The problem of search then becomes the problem of finding a strategy of directing one's eyes so that the object to be searched for falls within the area of conspicuity as soon as possible. The optimum saccade size should depend on the diameter of the conspicuity area. However, the mean distance between subsequent fixation points in tachistoscopic search increased only slightly. Statistical calculation may be used to derive search time from the size of the conspicuity area as done by Williams (1966) in the reverse direction.

In conclusion, it can be said that peripheral vision and the attendant concept of visual lobes play an important role in visual search. Techniques to measure the lobe area have been devised and, consequently, knowledge relating to visual lobes is now accumulating.

2.4 FEATURE-INTEGRATION THEORY

An influential line of work, independent of that described above, has been developed by Treisman (e.g. 1986).

2.4.1 Feature analysis in early vision

Vision provides an organised representation of the world around us, and this includes objects and organisms located or moving on a structured ground. The seemingly effortless ability to perceive meaningful wholes in the visual world depends on complex processes to which a person has no conscious access. However, some simple generalisations about visual information processing are beginning to emerge. One of these is a distinction between two levels of processing. Certain aspects of visual processing seem to be accomplished simultaneously or in *parallel* (i.e. for the entire visual field at once) and automatically (i.e. without attention being focused on any one part of the visual field). Other aspects of visual processing appear to depend on focused attention and are done *serially* (i.e. one at a time), as if a mental spotlight were being moved from one location to another (Treisman, 1986). However, it is likely that both serial *and* parallel processsing operate to varying extents within many tasks.

For a long time, visual information processing studies have made a distinction between *preattentional* and *focal attentional* processes. Preattentional processes operate rapidly and in parallel across the visual field, whilst focal attentional processes operate serially over more limited spatial areas. A feature-integration theory (FIT) of visual processing has been proposed by Treisman and her colleagues (e.g. Treisman & Gelade, 1980) that maintains that preattentional vision makes available information about discrete single features, but it does not make information available about how these features are spatially arranged or how they combine. To combine features, both in the same domain (e.g. two form features) and across domains (e.g. form and colour), the features must be focally attended (Müller *et al.*, 1990). Within the theory, attention must be focused on a given spatial location before single features can be combined into a coherent object. The theory proposes that prior to focused attention, isolated features are free-moving and may combine to produce *illusory conjunctions:* combinations of features that belong to different objects (e.g. a green *O* composed of the 'green' feature of a green *I* and the 'shape' feature of a red *O*).

Treisman's phrase *features and objects* is therefore a three-word characterisation of the emerging hypothesis about the early stages of vision. Treisman believes that there are many reasons to agree that vision does apply specialised analysers to decompose stimuli into parts and properties, and that extra operations are needed to specify their recombination into their correct wholes. The evidence is, in part, physiological and anatomical. Specifically, the effort to trace what happens to sensory data suggests that the data are processed in different areas of considerable specialisation. One area is concerned mainly with the orientation of lines and edges, one with colour and another with directions of movement, and only after processing in these areas do data reach areas which appear to discriminate between complex natural objects. Other evidence is behavioural. It appears, for instance, that visual adaptation (the visual system's tendency to become unresponsive to a sustained stimulus) occurs seperately for different properties of a scene.

In recent years, this model has received extensive experimental support. For instance, it has been shown that single features, such as the colour green, can be detected in parallel and can yield texture segregation. Experiments have determined which featural dimensions lead to such 'pop-out' effects. These dimensions include colour, size, shape, orientation, and curvature. Consistent with Treisman and Gelade's model, conjunctions of these elementary features cannot be detected in parallel, but require, serial, attentive search.

In several experiments (e.g. Treisman & Gelade, 1980), shape has been shown to yield a pop-out effect, and may therefore be proposed as a feature distinguishing chairs from other objects. Such pop-out effects occurred, however, for conditions in which shape was confounded with more elementary parameters such as curvature or closure. To assume that there are preattentive detectors for the shape of a chair transfers the burden of shape recognition to a lower level without solving the problem. Furthermore, such an example contradicts Treisman's views that *object perception requires attention.*

2.4.2 Preattentive processing

One strategy suggested by Treisman for subjecting the preattentive aspect of visual processing to laboratory examination concerns the fact that in the real world parts which belong to the same object tend to share properties. They have the same colour and texture, their boundaries show a continuity of lines or curves, they move together, and they are approximately the same distance from the eye. The investigator can thus ask subjects to locate the boundaries between regions in various visual displays and may learn what properties make a boundary immediately salient, or make it pop-out of a scene. Such properties are likely to be those which the visual system normally employs in the initial task of segregating figure from ground. It transpires that boundaries are salient between elements that differ in simple properties such as colour, brightness and line orientation but not between elements which differ in how their properties are combined or arranged. In other words, analysis of properties and parts preceeded their synthesis. If parts or properties are identified before they are conjoined with objects, they must have some independent psychological existence.

It follows that errors of synthesis should sometimes occur; subjects should sometimes see illusory conjunctions of parts or properties drawn from different areas of the visual field, and this is indeed what is found. Another way to make the early, pre-attentive level of visual processing the subject of laboratory investigation is to assign visual-search tasks. It is assumed that if the preattentive processing occurs automatically and across the visual field, a target that is distinct from its neighbours in its preattentive representation in the brain should 'pop-out' of the display.

The difference between a search for simple features and a search for conjunctions of features could have implications in industrial settings. For example, quality control inspectors might take more time to check manufactured items if the possible errors in manufacture are characterised by faulty combinations of properties than they do if the errors always result in a salient change in a single property. Likewise, the symbols representing, for example, the destinations for baggage handled at airline terminals should be characterised by a unique combination of properties (Treisman, 1986).

Treisman's team have thus exploited visual search tasks to test a wide range of candidate features which might pop out of displays and reveal themselves as primitives; basic elements in the language of early vision. These candidates fell into a number of categories such as length or number; properties of single lines such as orientation or curvature; properties of line arrangements; topological and relational properties such as the connectedness of lines or the ratio of the height to the width of a shape.

In summary, it would appear that only a small number of features are extracted early in visual processing. These include colour, size, contrast, tilt, curvature, and line ends. Other research shows that movement and differences in stereoscopic depth are also extracted automatically in early vision. Generally speaking, the building blocks of vision seem to be simple properties that characterise local elements, such as points or lines, but not the relations amongst them. The most complex property that pops out preattentively appears to be closure. Finally, Treisman's findings suggest that several preattentive properties are coded as values of deviation from a null, or reference value.

2.4.3 Later stages

Treisman (1986) then turns to evidence that focused attention is required for conjoining the features at a given location in a scene and for establishing structured representations of objects and their relations. One line of evidence suggesting that conjunctions require attention emerges from studies in which subjects were asked to identify a target in a display and say where it was positioned.

Interestingly, subjects sometimes identified the target correctly, e.g. a target distinguished merely by colour, but gave it the wrong location. Conjunction targets were different. The correct identification was completely dependent on the correct localisation. Indeed, attention does seem to require focusing on a location in order to combine the features which it contains.

How objects are perceived once attention has been focused on them and the correct set of properties has been selected is the next question which Treisman raises. Particularly, how does one generate and maintain an object's perceptual unity even when objects move and change? Kahneman (e.g. 1973) and Treisman suggest that object perception is mediated not only by recognition, or matching to a stored label or description, but also by the construction of a temporary representation that is specific to the object's current appearance and is constantly updated as the object changes. The perceptual continuity of an object would then depend on its current manifestation being allocated to the same file as its earlier appearances. Such allocation is possible if the object remains stationary or if it changes location within constraints which allow the perceptual system to keep track of which file it should belong to.

2.4.4 Criticisms of FIT

So, in real-world scenes, objects such as chairs cannot be found through a single-feature search. Instead, several features must be locally combined before the object can be recognised. It would follow that search must be serial through all locations in a scene. Yet, at a rate of perhaps 50–100 ms per location, such a search would take an enormous amount of time, particularly if totally irrelevant but nevertheless salient locations (e.g. a crack in the wall) have to be scanned. Such a scenario appears to contradict the extreme speeds found in the perception of real-world scenes (e.g. Biederman *et al.*, 1974).

As Dehaene (1989) notes, the feature-integration theory encounters some theoretical problems when applied to the perception of real-world scenes. Considering the task of looking at a chair in a room, humans are generally highly successful and are able to find a seat within a few hundred milliseconds. However, like most objects, chairs are not defined by a single feature that would otherwise be absent in the room. Although there may, at times, in a given environment, be a feature which does suffice to seperate chairs from non-chairs, there is no feature which systematically serves as a foundation for this distinction in all environments. A chair may be the only red object in the room, but unless we know *a priori* that we are searching for a red chair, a colour-based search strategy may result in us spending much time examining red curtains in a room where all chairs happen to be blue.

In essence, continues Dehaene, the difficulty is that attention should be (and appears to be) directed only at potentially interesting objects. Determining what and where these objects are, however, requires focused attention. This is a problem of practical efficiency similar to Dennett's (1978) 'frame problem'. If attention scans all salient locations of an image, considerable time may be lost examining things that are irrelevant to the present context.

Treisman and Gelade proposed an additional mechanism that may accelerate scene perception. The proposal was that default values for features are assigned to objects in the absence of focused attention, so that 'even when attention is directed elsewhere, we are unlikely to see a blue sun in a yellow sky' (p. 98). Prior to complete scanning, the viewer would possess a fairly correct idea of the attributes of several objects in a scene. Nevertheless, before applying this *a priori* knowledge, a brief but attentive analysis of the scene would be necessary, to ensure, for example that part of the scene is in fact the sky.

Dehaene mentions several other proposals which can be made to enrich the initial theory. Firstly, as certain objects acquire special relevance, the observer may become aware of new features that distinguish these objects from others more efficiently (Gibson & Gibson, 1955) and allow for an automatisation of perceptual processing

(Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Secondly, it is possible that the visual system scans the world in a hierarchical fashion. The heuristics of segmenting a visual scene first into large blocks, and then into smaller details, may reduce identification times. For instance, in searching for a chair, one could avoid attending to the titles of books lying on the table. Finally, one may take advantage of the many featural dimensions in a visual scene. Whilst objects in a scene are not singled out on the basis of only one feature, most of them, at least those that belong to basic-level categories (Rosch *et al.*, 1976) are generally very distinct from each other in terms of the number of shared features. Presumably, a chair shares some, but very few, features with a bookshelf or a crack in the wall. Although a single feature rarely suffices to seperate an object from the rest of a scene, if attention were first directed to locations that are globally highly discriminable from their background, an observer would be more likely to observe the relevant objects rapidly.

Dehaene concludes that the model most compatible with the data is a two-stage model along the lines of Hoffman (1978, 1979). The first stage consists of a parallel pre-selection of candidates with a single-feature item. All items selected enter a second stage of serial, self-terminating examination. During this latter stage, each item is compared to the search target. The comparison latency increases as a function of the number of features shared with the target but might also be modulated by the total number of dimensions used in the display.

This sort of model is also compatible with Treisman and Gelade's (1980) theory as it postulates that only single-feature searches can be carried out in parallel. This model also seems plausible in view of Julesz's experiments. For instance, Sagi and Julesz (1985) showed that the position and number of features that pop-out of a display can be determined in parallel, but that the identity of these features can not be reported with presentations of very short duration. The number and position of the candidates has to be computed in parallel in the preattentive stage of Hoffman's (1979) model in order to drive the next stage of focused attention.

Hoffman's (1978, 1979) two-stage model has recently received further support from experiments by Pashler. Pashler (1987) showed that search times increased when target confusable distractors were added, although the total number of distractors remained constant. This finding also confirms the results of Egeth *et al.* (1984). In one of Pashler's experiments, a two-target search task, RTs to one target were slowed to a similar degree when distractors similar to that (present) target, or to the other target (not itself present) were added to the display. It was concluded by Pashler that decision noise, specifically the difficulty of reducing false alarms when target-confusable distractors are present, was responsible for the effect of target-distractor similarity. This finding is again consistent with the two-stage model, which predicts that more candidates will have to be serially scanned when confusability increases.

Furthermore, the two-stage model is also compatible with recent reports of fast, quasi-parallel, search for conjunction targets (McLeod *et al.*, 1988; Nakayama & Silverman, 1986; Steinman, 1987). This appears to happen only with particular featural dimensions, which allow for a perceptual segregation of all items that share one of the target's features (Treisman, 1988). Rapid search for conjunctions is compatible with Treisman's model if it is accepted that search can sometimes be totally restricted to a subset of items (Treisman, 1988). Finally, a two-stage model partially resolves the difficulty mentioned earlier. Can we reconcile the slow processing predicted by serial scanning with the extreme speed observed in processing real-world scenes? Dehaene's data demonstrate that Treisman's views are basically correct. When a target for search is defined by a conjunction of features, some serial scanning is necessary, even if dimensionality and target/distractor discriminability are large. 'However, the rate of this search, inherently slow, can be considerably improved by a parallel selection of candidates. Although still serial, the search then becomes functionally equivalent to a parallel pop-out, which enables subjects to search through a display of 36 items in less than 700 ms' (Dehaene, 1989, p. 80).

2.5 EMERGENT PROPERTIES

Theories based on the parallel computation local primitives in the visual array has often been frustrated by the sensitivity of early human vision to so-called emergent properties. For instance, as Weisstein and Harris (1974) have noted, the detection of a line segment is heavily influenced by the three-dimensionality of its embedded context. Enns and co-workers note that even with dimensionality controlled, line detection is influenced by emergent features such as corners and closure (Enns & Prinzmetal, 1984; Enns & Gilani, 1988). Also not immune from troublesome emergent effect is texture segregation. 'Early models based on luminance differences (Julesz, 1975) were overturned by demonstrations that certain form differences alone were sufficient for grouping (Julesz & Bergen, 1983). Later models based on form were weakened by demonstrations that the proposed elements were neither necessary nor sufficient to produce reliable grouping (Enns, 1986)' (Enns, 1990, p. 37).

Considerable attention has been devoted to the visual features that pop-out because it may be that these are the basic elements on which to construct a comprehensive theory of visual attention. Research by Julesz and by Treisman have catalogued the visual features that permit preattentive detection. One obvious conclusion from such a table is that the primitive features of shape/form involve a simple contrast in the two-dimensional visual array. For instance, a line that differs in orientation from other lines, or a blob that differs in hue from other blobs, tends to pop-out. This area of work has now been extended by Enns (1990) to the third dimension of vision. At least two reasons exist to ask whether the human visual system is tuned to respond directly to three-dimensional properties. One reason is best articulated by Gibson (1966): even though all visual information passes through the two-dimensional retina before the brain, visual systems have evolved to assist organisms in their interaction with objects in a three-dimensional world. It is therefore conceivable that natural selection has equipped vision to extract some aspects of three-dimensional shape. A second reason is because object descriptions based on such properties currently hold considerable promise in the computational approach to object recognition. Enns remarks that several researchers argue that three-dimensional units termed generalised cylinders (e.g. cones, blocks, cylinders and spheres) provide the best balance between the combinatorial power needed in a generic set of visual units and the complexity of the information needed to represent those units (Marr, 1982; Pentland, 1986; Biederman, 1987).

The series of experiments reported by Enns (1990) asked whether parallel visual search could be based on emergent properties that define aspects of the three-dimensional world. Target and distractors were equated for primitive features (number and nature of lines and blobs), but differed in emergent features (perceived dimensionality, object type, object orientation). The results indicated that although some emergent features required serial search (e.g. rotation about the picture plane), others could be detected in parallel (e.g. slant, direction of lighting). These data imply that preattentive vision is not nearly as simple-minded as present theories might assume.

2.5.1 A hypothetical model

Much of what we see is recognised and labelled although this is not essential to vision. Unless basic cues (e.g. solidity) are completely misleading, people can manouevre in an unfamiliar environment. In perceptual (rather than developmental) terms, it is becoming increasingly popular to talk in terms of early versus late visual processing. The goal of *early vision*—to form a description of the three dimensional surfaces around us—may be distinguished from that of *later vision*—to identify and recognise objects and their settings (Marr, 1982). It is agreed amongst most theorists that the early description derives from spatial grouping of a small set of simple primitives that are registered in parallel across the visual field. These primitives, or functional features, need not correspond to simple physical dimensions such as wavelength or intensity, but rather their function should be to provide an 'alphabet soup of descriptive chunks that are almost certain to have some fairly direct semantic interpretation' (Witkin & Tenenbaum, 1983, p. 509). Examples might be coherent regions, edges, symmetries, smooth gradients, and flow patterns. Thus we expect the visual system to be sensitive to simple regularities in elements of different reflectance, colour and texture (Treisman & Gormican, 1988).

In their recent article, Treisman and Gormican (1988) review some new evidence relating to early visual processing and propose an explanatory framework. A series of search experiments tested detection of targets distinguished from the distractors by differences on a single dimension. Their aim was to use the pattern of search latencies to infer which features are coded automatically in early vision. For each of 12 different dimensions, one or more pairs of contrasting stimuli were tested. Each member of a pair played the role of target in one condition and the role of distractor in the other condition. Many pairs gave rise to a marked asymmetry in search latencies, such that one stimulus in the pair was detected either through parallel processing or with small increases in latency as display size increased, whereas the other gave search functions that increased much more steeply. Targets defined by larger values on the quantitative dimensions of length, number, and contrast, by line curvature, by misaligned orientation, and by values that departed from a standard or prototypical colour or shape were detected easily, whereas targets defined by smaller values on the quantitative dimensions, by straightness, by frame-aligned orientation, and by prototypical colours or shapes required slow and apparently serial search. These values appear to be code by default, as the absence of the contrasting values. They found no feature of line arrangements that allowed automatic, preattentive detection; nor did connectedness or containment—the two examples of topological features that were tested. Treisman and Gormican interpreted these results as evidence that focused attention to single items or to groups is required to reduce background activity when the Weber fraction distinguishing the pooled feature activity with displays containing a target and with displays containing only distractors is too small to allow reliable discrimination.

Two kinds of decomposition into more primitive elements are possible: analysis into properties and analysis into parts. The visual system may respond separately to values on different dimensions of a single stimulus; e.g. the stimulus colour, size, orientation, or direction of movement; or it may respond seperately to different component parts, e.g. a vertical line or an intersecting curve in a two-dimensional shape, or a flat surface or cylindrical legs in a three-dimensional object. Treisman and Gormican concern themselves with dimensional analysis; with properties rather than parts.

2.6 EYE MOVEMENTS IN VISUAL SEARCH

There has been considerable recent interest in two distinct patterns of performance in visual search. Under some conditions, search latencies seem to be relatively independent of the number of distractors. Within other conditions, search latencies increase linearly with the number of distractors in the display, and the slope of this function on target-present trails is half that on target-absent trials. The former pattern is usually referred to as indicative of parallel search and the latter pattern as indicative of serial self-terminating search. It should be noted that each of the patterns can be produced by both serial and parallel mechanisms (Townsend & Ashby, 1983). Broadbent (1987) and others (e.g. Ward & McClelland, 1989), for example, have adapted to visual search a parallel activation model of the sort used by Ratcliff (1978) to account for the performance of subjects in a memory-search task (Sternberg, 1967).

Within a range of frameworks, including Treisman's FIT (Treisman, 1986; Treisman & Gelade, 1980), the serial self-terminating search pattern is attributed to the sequential inspection of each item by an internal attentional system, a process that terminates when the target is located. In a typical search study, subjects are free to move their eyes. In addition, at least within a serial search, they have time to make several saccades with the larger set sizes. This raises the possibility that it is not the number of covert shifts of an internal attentional 'beam' that produces the linear RT increase in serial search, but rather the average number of saccades required to scan the display. In Klein and Farrell's (1989) study, two manipulations were used to eliminate the possibility of scanning of the array with saccadic eye movements. They hypothesised that if the serial pattern is correctly attributed to attentional shifts, then these manipulations should have little effect, and the typical patterns should be observed. On the other hand, if eliminating eye movements produces atypical patterns of search performance, then we may have to reexamine the assumed role of covert shifts of attention of attention in producing serial search patterns.

Visual search performance (with sets chosen to elicit both serial and parallel search patterns) under two conditions that precluded saccades was compared to the typical situation in which visual inspection of the array is possible. In one condition, the display duration was so brief that any saccades that were executed would be too late to bring the targeted portion of the array into the fovea. In the other, the display was present until the subject's response, but eye position was monitored and trials with shifts in fixation were excluded from analysis. The latter condition produced search latencies that were nearly identical to those with free inspection. Brief exposure, by contrast, did not produce the pattern typical of serial search, presumably because of strategies induced to deal with the rapid decay of the visual array. It is concluded that saccadic eye movements play a role in the patterns of performance used to infer serial and parallel search, and that brief exposure is not a satisfactory technique for exploring the role of saccadic eye movements in visual search.

To investigate the possibility that eye movements were a *limiting factor* in classical visual search, a computer-driven display was devised to allow a visual search, using a sequential search procedure, to proceed without eye movement (Budiansky & Sperling, 1969). In the sequential search procedure, a sequence of briefly flashed letter arrays is presented on a VDU, with each new array falling on top of its predecessor. A critical array containing a lone numeral target is embedded somewhere in the middle of the sequence. The target's spatial location (within the array) and its identity are chosen randomly on each trail. The task of the subject is to detect the location and to identify the target.

In rapid, natural visual search through simple material, the eyes make about four saccadic eye movements per second, each movement lasting a few tens of ms (depending on the distance traversed) with the eyes relatively motionless between saccades. To approximate this natural search mode, the computer-generated arrays are exposed for durations of 200 ms with brief 40 ms blank periods between arrays. The subjects are instructed to maintain stable eye fixation on the centre of the display. The successive arrays displayed to the stationary eye approximate the stimulus sequence ordinarily produced by saccadic eye movements. The exposure parameters are not critical. For example, data obtained with 200-ms exposures followed by 40-ms blank periods are not different from data obtained with 10-ms exposures and 230-ms blank periods (e.g. Sperling & Melchner, 1989*a*).

The computer-generated sequence has many information processing advantages over the natural sequence. For instance, in natural search, when the eyes do not move quite far enough between fixations, some of the same material falls within the eyes' search area in successive fixations and is searched twice, which is wasteful. Even if redundant material on the retina is ignored, the redundant material still usurps space within the search area that could have been occupied by new material. If the eyes move too far between fixations, they leave unsearched lacunae in the stimulus.

In natural search, there are two unknown factors: (a) the eye movement strategy and (b) the attentional field around the eye fixations. Eye movement strategy must be known to determine the attentional factors. In the computer-generated sequence, eye movements are effectively eliminated, so that the attentional field around fixation can be determined (Sperling, 1989).

Perception is inherently selective, and selective processing of information from the visual environment reflects an ongoing interaction between perception and action. Global and detailed processing of visual arrays frequently guide subsequent information gathering activities that, respectively, provide further material for processing by the visual perceptual sytem. Selective sampling of available stimuli is an adaptive mechanism which allows one to focus upon the most important aspects of the environment.

In order to be selective, it is necessary to process and integrate information from various parts of the visual environment. The 2° region of the visual field projecting onto the fovea represents the area of greatest acuity and provides the most detailed processing. The peripheral retina, however, receives stimulation from a much larger region of the visual field (approximately 200° horizontally and 130° vertically) and is well designed for selective sampling of a visual array due to its sensitivity to levels of illumination and to movement. Although information from the visual periphery is relatively crude and indistinct, it appears to provide an important basis for the direction of focused attention. Eye movements may be considered as an attempt to orient objects detected in the periphery onto the fovea for closer examination.

The pattern of successive eye fixations is regarded as an observable manifestation of selective processing, particularly in visual scanning tasks.

The interaction of foveal and peripheral processing in conjuction with cognitive strategies has been frequently demonstrated in the visual search literature. Williams (1966), as one example, showed that colour cues could be used to speed and by inference, direct visual search to target. Such findings support the notion that scanning is not a random process and that the less detailed processing of peripherally-located information can be a major determinant of eye fixation patterns during visual search. It follows that if searchers are given some prior knowledge regarding the relevance of particular stimulus information in a search task, they can selectively locate attention and/or accompanying eye fixations to specific information which is most critical for task resolution.

By way of concluding this section, it has been said that this picture of the searcher as a selective processor and integrator of multiple stimulus features is consistent with that given by many visual perception theorists, whether their primary concern has been information processing (Neisser, 1967), selective attention (Kahneman, 1973), visual scanning (Hochberg, 1970), or perceptual development (Gibson, 1969). All these theoretical positions place a burden on the perceiver both to appreciate figure-ground seperations, invariant relations, or distinctive features in a stimulus array and to integrate these elements in the visual environment with strategies, expectations, or intentions. Such selective, integrative activity presumably underlies changes in the locus of gaze and results in efficient visual exploration with regard to stimulus properties and the task-in-hand (Cohen, 1981).

One underlying theme implicit throughout this section has been the overt/covert vision distinction. The discussion relating to this distinction is now continued in connection with directed attention.

2.7 DIRECTED ATTENTION

Engel's attention area experiments indicate that a difference may exist between the spot where our eyes are fixated on, the fixation point, and the location where our attention is directed to the attention point. The results of Kaufman and Richard (1969) also emphasise this distinction. They compared the points where a naive observer believes his eyes are directed and where in fact the eyes are oriented to; points which do not coincide.

Due to the difference obtained between the conspicuity area and the attention area, directed attention selectively emphasises peripheral vision in a particular retinal area. While looking at 'complicated' stereoscopic pictures illuminated instantaneously by electric sparks, Helmholtz (1925) remarked:

> It is a curious fact that the observer may be gazing steadily at the fixation mark, and yet at the same time he can concentrate his attention on any part of the dark field he likes, so that when the spark comes he will get an impression about objects in that particular region only. In this experiment the attention is entirely independent of the position and accommodation of the eyes or, indeed, of any known variations in or on the organ of vision. Thus it is possible, simply by a conscious and voluntary effort, to focus the attention on some definite spot in an absolutely dark and featureless field. In the development of a theory of attention, this is one of the most striking experiments that can be made.

Prior knowledge concerning stimulus location decreased the peripheral intensity threshold (Lie, 1969), while the number of corrected judgements for very brief exposures near time threshold increased (Grindley & Townsend, 1968; Keeley, 1969). Engel's experiments demonstrated the influence of directed attention to retinal locus. Naturally, an isolated object is very conspicuous, so conspicuity and visibility area are likely to be equal. This is consistent with the results of Grindley and Townsend (1968, 1970) who concluded that in peripheral vision attention acts selectively only when there is a complex pattern of stimulation. Interestingly, only the direction of attention with respect to the fixation point and not the distance from the fixation point seemed to be significant in Engel's (1971) experiments.

The attentional mechanism reduces the incoming information, which to some extent can be 'tuned'. Apart from the additional extension, the attention area equalled the corresponding conspicuity area, so that visual conspicuity and expectation concerning test object location must be considered to be independent determiners of attention. Visual conspicuity may be considered as an external determiner of attention, whilst expectation concerning the test object location is understood to be an internal determiner of attention. Over a century ago, Fechner (1860) described these as voluntary and involuntary determiners of attention.

Engel supposed visibility to be determined at the sensory level, where information reduction is already effected. As at the attention level, here selection is influenced both by internal determiners, such as eye movement, and by external determiners such as receptor adaptation and inhibition. As the extension of the attention area was limited by the visibility area, information reduction by visibility and by attention seem a serial process. Disagreement remains about the level at which selection by the attention mechanism may be successfully reduced at the different levels of the sensory processes. The visual conspicuity of an object in its visual background may be postulated as related to a relatively low level of processing. Other selection factors like novelty obviously concern memory and learning within the subject and must be located at higher processing levels.

2.8 ATTENTIONAL FOCUS

Despite the vast literature on visual search, few studies have addressed the question posed by Eriksen and Webb (1989): If a subset of stimuli must be attended in a complex large visual display, does it matter which way this subset is distributed throughout the display and intermixed with distractor stimuli? To preclude the role of saccadic eye movements, they dealt only with displays presented for duration too short for changes in eye fixation. There follows below a précis of Eriksen and Webb's rationale, methodology and experimentation described in their 1989 report.

Two previous studies produced conflicting conclusions (Shaw & Shaw, 1977; Shaw, 1978; Podgorny & Shepard, 1983). The VDU used in Podgorny and Shepard's (1983) study was a 3×3 matrix of squares. Before a trial, subjects were instructed to distribute their attention over a subset of these three squares. Squares to be attended were lightly shaded. When the subject felt that their attention was deployed over the relevant area, they initiated the trail. At this point a dot appeared in one of the nine squares, the subject made a discriminated response to whether the dot occurred on a shaded or a non-shaded square, and the latency of the response was recorded.

Latencies were shorter when the target fell on attended squares. In addition, the latencies were reduced when the sub-region of shaded squares (attended) constituted a compact region on the display. Latencies increased as attended squares were distributed over the display with non-attended squares intervening between attended ones.

In a re-analysis of Podgorny and Shepard's (1983) data, Crassini (1986) concluded that a benefit in responding occurred only when the attended squares were not seperated or interspersed with squares that were not to be attended. These results suggest that attention can be distributed over sub-areas of a visual display and, in addition, that difficulty in this distribution of attention occurs if the to-be-attended area is not compact, or perhaps unitary. However, one problem was that as the experiments provided no control for eye fixation, possible attentional effects are confounded with retinal acuity since it is likely that the subjects varied their eye fixations to correspond with the area of the display that was being attended.

At variance with the findings of Podgorny and Shepard are the results of the Shaw and Shaw (1977) and Shaw (1978) studies. A two-choice RT task with visual displays containing one target letter and five distractor letters was used in the Shaw (1978) study. Although Shaw did not specifically address the question of distribution of to-be-attended areas throughout the display, the data from her experiment do bear directly on this question. Before each trial, different display locations were assigned different probabilities of target occurrence. Variation of these locations in the display ranged from compact to well distributed. The assignment of different probabilities of target occurrence to different display locations was a manipulation designed to control attentional deployment in the display. According to Shaw's model, visual attention is regarded as consisting of a fixed amount of resources that can be allocated to different display locations on a trial in quantities corresponding to the probability of a target occurring in that location. Both the latency (Shaw, 1978) and the accuracy (Shaw & Shaw, 1977) of the discriminated response were found to correspond with the *a priori* probability of the target occurrence. There was no indication, however, that the results were dependent upon the particular locations or proximities of high-probability target locations throughout the display. Although Shaw's model does not specifically address the issue of the effect of compactness of the to-be-attended display areas, the model does suggest that compactness or unitariness of the to-be-attended areas would not be a factor in determining the effectiveness of attentional allocation. Attention may be allocated to discrete areas of the display, and processing of stimuli in these locations would proceed in parallel, but at different rates if the allocation of resources to the different positions was unequal. However, note Eriksen and Webb, as with Podgorny and Shepard's (1983) experiments, the data are less than conclusive as to whether or not attention was actually deployed on several seperated locations on each trial.

Late-selection models of attention (e.g. Hoffman, 1979; Duncan, 1980) do not seem to offer a basis for an expectation that the distribution of high-probability target locations would affect performance. What distinguishes these models is the proposition of a large-capacity parallel-processing preattentive stage in which stimuli are processed to the point of identification (Duncan, 1980). On the basis of the discriminations and identifications occurring in the preattentive stage, stimuli are selected for serial entry into STM, consciousness, and access to responses. 'Applied to our problem of the distribution of designated locations in a complex display, late-selection theories would not expect distribution to make a difference since all locations are parallel-processed. There might be a gain by restricting the target to subregions of a display in that these regions or locations might constitute a discriminative cue that would therefore expedite the processing of the stimuli in these locations. Again, however, since this subset of locations would be processed in parallel, there is no reason to expect their adjacency or seperation to affect the efficiency with which they are processed' (Eriksen & Webb, 1989, p. 176).

A zoom lens model of the visual attentional fields has been proposed by Eriksen (Eriksen & Yeh, 1985; Eriksen & St. James, 1986) in which the attentional field can vary in size from approximately one degree of angle to a size encompassing the entire effective visual field. Analogous to a zoom lens, as the size of the field varies, there is a reciprocal variation in the resolving power for detail. When attention is directed to a specific stimulus, the attentional field effectively zooms in on the stimulus, concurrently increasing resource density and decreasing field size, until the resource concentration is sufficient to make the required discrimination or extraction of detail and information.

Whilst the model is unspecific concerning the shape that the attentional field can take on, the initial assumption is that attentional resources are uniformly distributed over the field. If two stimuli located six degrees of angle apart are to be simultaneously attended, the attentional field must be enlarged to at least six degrees to encompass both stimuli. Here, there will be fewer resources for processing all of the stimuli than if the stimuli had been only three degrees apart. In the six degrees case, the density of resources on a given location in the attentional field will be lower. Although lower density may be adequate to make the necessary discrimination if the discrimination requirement is not high (e.g. between X and O), it is assumed that the discrimination will be slower than if the resource density had been higher. Instances will occur where the resource density in a large attentional field is not sufficient to extract the necessary information from the stimuli in order to make a discrimination. In such a case, the attentional field is presumed to zoom in on one stimulus and then the other, alternately processing each.

Eriksen and St James (1986) employed circular displays of eight letters and cued or precued one, two or three adjacent locations. A two-choice RT paradigm was used, and a target letter occurred in only one of the cued locations. It was found that mean RT increased as the number of cued locations increased. The supplementary analyses and experimental operations strongly indicated that the cued locations were processed or searched in parallel rather than serially, and this was consistent with the conclusion that the attentional focus or field enlarged or contracted with variation in the number of cued locations. Data strongly indicated that when the number of cued locations approached half of the display area, subjects tended to abandon processing the cued area alone and instead enlarged the attentional field to include the entire display. This strategy was effective since the discrimination involved in the experiments was between *S* and *C*, with distractor letters that were composed of straight lines and angles, such as *N* and *W*. In these experiments, the cued locations were always contiguous, and so the attentional field could remain unitary even when enlarged to include three locations. 'Inasmuch as RT increased with increases in cued area, interspersing cued locations with noncued positions containing noise or distractor stimuli would be presumed to decrease the density of processing resources to the point that they were insufficient to perform the task. This would be particularly likely if the discrimination task was made difficult. At some point, we would anticipate that the attentional field would zoom in to focus on only one position at a time, and thus processing would become serial' (Eriksen & Webb, 1989, pp. 176–177).

Although the above models are not exhaustive in terms of different ways of conceptualising visual attention, they are representative of three classes of models and embody most of the distinguishing characteristics relevant to the present issue. The model of Shaw (1978) provides for a dividing of attention in the visual field by allocating different amounts of resources to the separate locations. It would, incidentally, be interesting for future researchers to specifically relate this model to the windows environment. Similar to Shaw's model, late-selection models with parallel processing of all the stimuli in the preattentive stage would not predict a differential effect of location distribution on processing efficiency. Both models, however, would predict that the greater the number of these to-be-attended locations, the poorer the performance. With Shaw's model, this would be attributable to the dividing of a limited number of resources among the different locations. On the other hand, late-selection theories would attribute a decrease in performance with an increase in number of to-be-attended locations to an increased probability of confusion errors and the necessity for a higher response criterion.

By way of conclusion, it has been noted further by Eriksen and Webb that from the zoom lens model (and also from the variable-size spotlight model of LaBerge (1983) a prediction can be made that performance would be impaired if the to-be-attended areas are distributed throughout the display and interspersed with distractor stimuli. In order for the attentional field to enlarge to include all to-be-attended locations, resources would have to be distributed over non-designated locations also, so as to keep the attentional field unitary. Processing resources would thus be thinly distributed and wasted on non-designated locations. Eriksen and Webb also remark that both the zoom lens and variable-size spotlight models provide for a serial search approach to the multi-locations task. If the discrimination is not able to be efficiently handled by a distributed attentional field, the spotlight can narrow or the attentional field can zoom in to concentrate resources on one location at a time. Naturally, the time required for serial search will increase with increases in the number of designated locations. The effect of distance between these locations would be dependent upon whether attention moves in an analogous fashion across the visual display. Whilst certain studies conclude that attention does move in this analogous manner (Shulman *et al.*, 1979; Tsal, 1983), the adequacy of their evidence has been questioned (Eriksen & Murphy, 1987). More recent studies (Sagi & Julesz, 1985; Murphy & Eriksen, 1987) conclude that attention indexes locations in the visual field at a speed independent of the distance of the initial starting point (Eriksen & Webb, 1989).

2.9 ATTENTION IN VISUAL SPACE

Oculomotor behaviour serves to place specific images on the fovea for detailed visual analysis, as visuomotor tasks obviously require an efficient flow of relevant visual information. Monitoring of the appropriate region of visual space by the visual attentional system is equally important for facilitating the input of appropriate visual information to the oculomotor system. The components of the visual attentional system which provide such distinct monitoring are the foveal and peripheral visual systems (e.g. Shapiro *et al.*, 1984).

Foveation during object movement allows important information about object velocity and shape to be acquired. In dynamic displays, it is necessary for the two systems to work together to produce a scanpath. 'Thus, a scanpath can be defined as a series of saccadic eye movements alternated with short fixations (of stationary objects) or foveations (of smoothly moving objects) which allow the viewer to sample the visual environment discretely. Cognitive control is an important dimension determining the scan path and permits an observer to choose various strategies for assessing the information in a display' (Shapiro, 1990, p. 170).

In dynamic viewing situations where objects may appear and disappear unexpectedly, there might be disadvantages to frequent sampling. For instance, a saccade to an inappropriate viewing area may decrease detectability of objects appearing at another, more important, location on the display. Because saccadic eye movements are extremely fast, visual detection is reduced during the eye movement (Matin, 1974). More importantly, the initiation and execution of a saccade may deplete attentional or cognitive resources. It follows that efficient use of the oculomotor system should enhance performance in certain dynamic tasks by controlling the flow of visual and cognitive information. Learning to perform a task optimally may therefore involve the acquisition of efficient, rather than inefficient, eye movement or scanpath strategies (Shapiro, 1990).

Whilst foveation for detailed processing is the primary function of the oculomotor system, images falling on peripheral areas of the visual field also undergo substantial visual processing. Several workers (e.g. Eriksen & Hoffman, 1972; Posner, 1980) have shown that a viewer can dissociate the locus of attention from the point of fixation. As such, when an object appears in the periphery (e.g. as in Expt. 6 of this thesis), a VDU user may utilise the (albeit reduced) visual processing capacities of the peripheral fields by directing attention to that area and eliminating the costly effort of making an eye movement to fixate the object. Should the object be sufficiently large to be effectively analysed by the peripheral areas of the visual field, then shifting or widening attention to peripheral areas may be an important contributor to an efficient oculomotor strategy, assuming that attentional shifts are less costly and faster than oculomotor shifts.

One's ability to successfully monitor the visual periphery without a significant cost to the concomitantly attended foveal region received support in an experiment by Shapiro and Raymond (1989). They used a simple oculomotor task in which subjects were required to visually track a small target which required foveal vision. To ensure that foveal vision was employed, a small probe stimulus requiring a RT response appeared at random intervals close to the target. Concurrent with the foveal RT task, an additional task was for the subject to detect the appearance of another small target which could appear anywhere in the periphery (up to 20° of visual angle away from the foveal target) of the display. The appearance of this peripheral target also required an RT. Subjects were instructed to attempt to bring their RTs to increasingly faster levels on both tasks by whatever possible means. Although subjects adopted different strategies, all were able to reduce their RTs on both tasks, suggesting that they were able to monitor foveal and peripheral regions concurrently.

Shapiro's (1990) study investigated three hypotheses: (a) that efficient oculomotor behaviour is a crucial component of complex visuomotor tasks; (b) that this behaviour can be acquired through practice on a series of simple, unrelated tasks; and (c) that such behaviour can be transferred subsequently to the complex visuomotor task. A video game was used as the complex task from which performance was evaluated. Two groups of subjects were exposed to two different sets of simple tasks, or drills. One group (the efficient eye movement experimental group) received training designed to minimise eye movements, whereas a second group of subjects (the inefficient eye movement experimental group) received training designed to increase the frequency of eye movements. Oculomotor training was interspersed with practice on the game. Performance of these two experimental groups on various measures was compared with an untrained control group. The group receiving efficient oculomotor training exhibited significantly superior performance on the video game, demonstrating fewer foveations than either the inefficient or control groups (which did not differ from each other).

The results indicated that it is not just the oculomotor training *per se* which facilitated the performance improvement seen in the group receiving the efficient training regime. It would appear that it was the interaction of the differential monitoring of the foveal and peripheral regions of visual space which allowed the effectiveness of the oculomotor training to be accomplished. Merely keeping one's foveal vision on the designated target would not be sufficient to accomplish the task confronted on the screen, as well as a host of other visuomotor tasks. In other words, it is the concurrent monitoring of the peripheral visual field which enabled the task to be effectively accomplished.

> The mechanism by which the peripheral visual field is monitored is the subject of some debate. Given the task confronted by subjects in the present investigation, it is likely that they expanded the field of visual space being monitored, rather than divided (or alternated) attention between the foveal and peripheral channels. This conclusion is drawn on the basis of the nature of the tasks, given that the target tracking task, necessitating the foveal visual channel, required nearly constant monitoring. As Posner (1980) and others have suggested, monitoring a wider region of visual space is accomplished with a cost proportionate to the increase in attentional area. Thus the metaphore of visual attention as a searchlight implies that widening the beam means that the same amount of light must now be spread over a larger area. However, since the peripheral task (mine detection) required little by way of focused attention, it is likely that such a strategy would be effective. In other visuomotor tasks, where the peripheral task required more than mere detection of the presence of an object, such a distribution of attention might be insufficient to accomplish the task.

(Shapiro, 1990, p. 175)

Shapiro concludes that there are two active components of the visual perceptual system which can be brought under voluntary control by incorporating each into an oculomotor part-task training regimen: (a) the oculomotor system, which acts to pace both moving and stationary objects on the fovea, and (b) the visual attentional system which acts to maximise the use of foveal and peripheral visual sub-systems.

Voluntary and involuntary shifts of attention In a recent study by Flowers (1991), displays consisting of a single target letter, located 2° on opposite sides of fixation, or a target letter and distractor located 2° on opposite sides of fixation were preceded by bar markers adjacent to possible target letter locations. The location of these markers was either *unrelated* to target position or identity (Expt. 1), 80 percent predictive of target location (Expt. 2), or 80 percent predictive that the target would appear on the opposite side (Expt. 3). Large performance decrements occurred when the markers occurred adjacent to a noise letter in Experiments 1 and 2, while Experiment 3 demonstrated a limited ability to overcome involuntary attentional capture when the bar marker signalled a need to shift attention away from its location. Nevertheless, the potency and time course of involuntary capture by onset of marker cues, irrespective of their overall predictive value, merits consideration in theoretical accounts of visual attention and in the *design of optimal information displays.*

2.10 SEARCH MODELLING

2.10.1 Exploring search strategy

Humans have little control over their visual lobe size, and little control over fixation durations. Therefore, the only parameters of the search task amenable to manipulation by the subject are their search strategy and the total search duration before a search field is classified as containing no targets. Search strategy has been studied using eye movement recordings in industrial inspection, X-ray interpretation, driving and flying. There is little agreement, however, on the search strategy employed. Search is neither totally systematic or completely random. In many instances, it does not even cover the complete search field. Nevertheless, the strategy does change with practice and subjects do learn to improve their performance in search tasks. Exactly what is learned is still unclear (Drury, 1991).

Theories of visual search are not particularly helpful in elucidating search strategy. For mathematical convenience, Drury and others have derived theories of random search and systematic search, which perhaps serve best as lower and upper bounds on actual subject performance. Some intermediate theories have been proposed (e.g. search such as a memory-limited process) but there is an obvious need for guidance to the modellers before more effort is expended in exploring the mathematical structure of unrealistic models.

Drury (1991) presents three recent approaches to search strategy, pointing out that none have yet produced a definitive search model. Firstly, the movements of a restricted FOV across a search field was studied (Bhatnager, 1987). The context was inspection of micro-circuit chips under an optical microscope. FOV movements could be classified into a small number of consistent patterns, and the patterns related to overall inspection accuracy. Secondly, a study of circuit board inspection (Chi, 1989) used post-experiment questionnaires as a means of understanding search strategy. Subjects could respond in some detail with their strategy, but the validity of these responses was untested. Finally, a restricted FOV movement task using circuit boards was performed by Kleiner (1989) while taking continuous verbal protocols and comparing these to post-experimental protocols. The picture emerging from these analyses is one of considerable cognitive complexity. Drury does conclude, however, that all three studies suggest that it may be possible to eventually interpret and model human visual search strategies.

2.10.2 Target detection in continuous visual search

Nattkemper (1991) addressed the issue of how to conceptualise the operations underlying the detection of targets in continuous search. In the view that target detection is accomplished by two mechanisms that may be conceptualised as localisation (where-processing) and identification (what-processing), target detection would be regarded as a two-stage process. Specifically, the processing of the identity of the to-be-searched object is preceded by processing of the position where the object is located.

To obtain some empirical evidence about the operations underlying the detection of targets Nattkemper traced back the 'prehistory' of detecting a target and analysed the spatial and temporal parameters of saccades when the eye approached the position of the target.

Results showed that the duration of fixations next to the target (F_n) was longer than average fixation duration. Also, the duration of this fixation was a function of the distance to the target. The larger was the distance between the fixation location and the position of the target, the longer was fixation duration. Thirdly, the duration of the preceding fixation (F_{n-1}) located in more remote positions relative to the target was shorter than average duration. Finally, the amplitude of the saccade between these two fixations depended on the distance between the location of fixation F_{n-1} and the position of the target. The larger the distance, the larger was the saccade amplitude.

Overall, the results seem to suggest that target detection is accomplished by two mechanisms. On fixations located in remote positions relative to the target. the position of the critical object is processed (localisation or where-processing). The subsequent saccade then sends the eyes near to that location in order to identify the target (identification or what-processing) and to stop the search.

2.10.3 Modelling conspicuity

Search tasks have been developed as a means of evaluating the conspicuity of coloured targets and the extent to which stimulus attributes such as target shape, size or orientation differences affect performance. In order to understand the processes involved for the purpose of modelling visual search, eye movements were recorded and analysed by Barbur *et al.* (1991) during these experiments. A model of visual search was developed involving several stages which determine the direction and amplitude of successive eye movements, based largely on the number and position of background distractor elements, the remembered scanpath and the 'effective' lobe size. As the lobe size is an important parameter in the model, an experimental model which allows its estimation from eye-movement measurements was also developed. Measurements of lobe size for different stimulus attributes, and the corresponding results obtained from search experiments have been used to optimise the parameters which affect the predictions of the model. The results show that it is possible to accurately predict the mean number of fixations required to locate the target during a search task, the variance associated with the mean search time and also to simulate successfully the patterns of the fixations and hence typical scan paths for selected pseudo-random patterns of stimuli. Barbur *et al.* claim that the model provides a better understanding of the processes and the performance limits associated with visual search.

2.11 SOME SPECIFIC APPLICATIONS OF VISUAL SEARCH

2.11.1 Advertisements and distraction

In most public buildings such as railway stations and shopping centres, users need sign-posting to find their destination. To be effective, these signs should be conspicuous. In other words, they should be easily noticed at least by people intentionally searching for routing information. In many actual situations, this requirement is clearly not met. This lack of conspicuity can have various sources: the sign itself (e.g. too small); the location of the sign (e.g. far from normal line of sight); objects in the vicinity of the sign (e.g. distracting architectural elements or advertisements); the actual layout of the environment (e.g. a tortuous corridor); the cognitive state of the user (e.g. nervous or uncertain). Much work has been conducted on searching for a target in cluttered visual fields (e.g. Williams, 1966; Treisman & Gelade, 1980; Jenkins & Cole, 1982; Cole & Jenkins, 1984). In these experiments the stimuli were composed of disjunct simple geometric forms and it is therefore difficult to generalise from the results of this work to complex realistic scenes and practical conditions. However, some studies used realistic scenes or slides of such scenes as stimuli in examining the conspicuity of targets. Holahan *et al.* (1978) studied the effect of distracting stimuli in the area around a stop sign on the reaction time to the sign. Shoptaugh and Whitaker (1984) measured the reaction time to directional traffic signs embedded in photographed street scenes. Cole and Hughes (1984) determined the detection performance of car driving subjects for objects situated along streets. All three studies demonstrated that distractors present in a scene reduced subjects' search performance. It can be safely assumed that advertisements present in a scene can decrease the conspicuity of routine signs. Considering the notion of conspicuity in more detail, Cole and Hughes (1984) distinguished between *search conspicuity* and *attention conspicuity.* With search conspicuity, performance is mainly determined by the set of the observer who searches intentionally for a specified target. On the other hand, if an unexpected object attracts attention, the object is said to have a degree of attention conspicuity (Boersema & Zwaga, 1990).

Research presented by Boersema and Zwaga (1990) tried to determine and analyse the effect of advertisements on the conspicuity of routing signs. The conspicuity of a routing sign can be operationally defined as the probability that the sign will be noticed by an observer within a fixed time or, conversely, as the time required by an observer to notice the sign. Three experiments were described in which the distracting effect of advertisements were measured. The stimulus material consisted of colour slides of realistic scenes containing systematically varied amounts of advertising. Given the task of the subjects in these three experiments, it seems that in all cases search conspicuity was measured as the subjects intentionally searched for a target (a routing sign). Although all signs were blue with white lettering, they had different shapes and sizes. Therefore, the subjects did not know exactly what the target would look like and so attention conspicuity may also have been required. Consequently, the conspicuity measured is a combination of search and attention conspicuity. The task of the subjects was to find in a scene's routing information, as quickly as possible, both a previously specified target word (a destination) and its direction arrow. Reaction time and eye movements were recorded.

The results indicated that reaction time is not a sufficiently sensitive measure of conspicuity. Search time derived from eye movements, however, increased significantly with the number of advertisements in two of the three experimental scenes.

A few theoretical implications of the results can also be mentioned. Firstly, results indicate that the search conspicuity of routing signs, as induced by the set of the subjects cannot fully compensate for the attentional conspicuity of even small advertisements. Secondly, the results do not support the notion of parallel processing for the case of realistic scenes. Although the subjects knew a unique characteristic (the colour) of the target, this did not enable them to ignore all other elements in the stimulus field. Thirdly, the results indicate that visual search is, at least in part, driven by stimuli outside the centre of the visual field.

Using colour slides of realistic scenes raises practical and methodological problems. To assess the feasibility of computer generated images being used as stimuli, Boersema & Zwaga (1991) created a compatible but simplified computer image for each of the realistic slides used previously. Search performances with these two kinds of stimuli were experimentally compared. The results for both sets of stimuli corresponded perfectly. In an additional experiment with the computer generated stimuli to further substantiate their feasibility, the hypothesis was confirmed that moving a distracting element (representing an advertisement) from a peripheral to a more central position decreased the conspicuity of the target element (representing a routing sign). It was concluded that the simplified stimulus fields are good representations of the realistic slides and an be used as stimulus material in future experiments.

2.11.2 Chess

The analysis of chess players' thought-aloud protocols shows that chess players' problem-solving depends on the task relevant cues (i.e. important chess-specific patterns such as threats and standard piece formations). These kind of learned patterns play a very important role in chess players' recall of chess positions, and it has been shown that in random positions skilled chess players are not able to benefit from their vast storage of chess specific patterns. A series of visual search experiments was conducted to test if randomisation also affects information intake analogously to the recall experiments. In these studies, it was noticed that skilled chess players are far superior than less skilled, both in game and random positions. Further experiments suggest that location coding is a key issue in analysing the asymmetry in the results of recall and perceptual classification experiments. It is suggested by Saariluoma (1990) that in perceptual classification subjects are able to make the decision concerning the presence of the target long before they code its location, but in recall they must code the locations of the pieces.

2.11.3 Racket sports and vision

Rapid and efficient perceptual mechanisms are required in open skills like tennis.

Opponent's behaviour and ball trajectory have to be analysed in order to select and execute the proper motor response. The quality of flow of visual information has an important influence on the cognitive and motor skills on which the game performance is dependent. Research in ergonomics (e.g. Mourand & Rockwell, 1972; Schoonard *et al.*, 1973; Papin *et al.*, 1976) show the expertise can influence visual search patterns. Results from experiments by Goulet *et al.* (1990) show that for expert tennis players, valuable information is selected during the 'preparatory' phase of service (that between the ritual and execution phases). Novices need to see the ritual phase until ball/racket impact to be more accurate. The findings emphasise the importance of combining the sampling of eye movements to behaviour parameters to sharpen the understanding of the perceptual processes underlying motor performance.

2.11.4 A cognitive approach to visual search in sport

Helsen and Pauwels (1991) have recently presented their cognitive approach to visual search in sport. This study investigated the relative importance in the determination of expertise in a sports discipline of attributes of the visual signal determined largely by the efficacy of the CNS versus cognitive attributes. The 'hardware' or 'task independent' attributes of the visual system assessed were simple reaction time, static visual acuity, dynamic visual acuity, peripheral visual range, and visual correction time. The 'task dependent' or 'software' attributes were complex decision speed and accuracy, number of fixation durations and fixation duration. The results of this approach revealed the importance of the 'software' dimension of visual search in the determination of skill in football. None of the 'hardware' factors predicted performance. If the goal is to predict performance, the more task-specific the test is to the criterion skill, the better. This is presumably due to greater overlap of knowledge bases required to perform the test and actual skills. In this way, these conclusions of visual search in solving tactical game problems support the findings of Starkes and Deakin (1984) and Starkes (1987, 1990). The interaction between level of skill and processing game-specific information thereby provides a firm foundation for evaluating the cognitive dimension of skilled performance and perception.

2.11.5 Driving

It is commonly accepted that approximately 90 percent of information input to the driver is visual. However, the introduction of in-vehicle navigation systems means that the driver of the future may have to cope with a greater flow of information which may overload the visual resources to the detriment of safe driving. In a eye movement study of on-road driving behaviour, Fairclough and Maternaghan (1991) provided evidence that a high visual load originating from inside the vehicle also results in decreased usage of rear-view mirror and differential driver search patterns when the navigation aids are in use. These results are discussed with implications for driving behaviour as a whole and for the design of in-vehicle information systems.

Many eye movement studies have indicated that fixation duration is influence by *mental workload* (e.g. May *et al.*, 1990). However, the results are by no means unambiguous, with some studies reporting longer fixation times when task difficulty is increased, and other studies report shorter fixation times. Unema & Rötting (1990) argue that within a theoretical framework based upon a combined cognitive stage model of information processing, most of these differences can be reconciled. The difference is thought to lie in the type of task being performed, and more specifically in the mental state for performing the task. As such, fixation duration and scanning behaviour may be indicative of the mental state of the subject. The data gathered with bus and car drivers seem to support the notion that arousal is a major factor influencing fixation duration. Unema and Rötting found shorter fixations among less experienced drivers and with tasks requiring high arousal, as opposed to the longer fixations found in tasks requiring more conscious processing and lower arousal. In other words, the fixation duration measure appears to be sensitive to differences in types of tasks. Fixation duration might therefore be a measure that can distinguish between tasks with a high load on early processes versus tasks with a high load on later (conscious) processes, applicable in a variety of *real* work situations. It therefore qualifies as one instrument in a canon of instruments needed to measure mental workload.

2.11.6 Pilots' scanning behaviour

The modern battlefield has become a proving ground for new technology. The cockpit of aircraft must be highly computerised and must allow for the simultaneous presentation of large quantities of information. A recent experiment by Oatman *et al.* (1991) compared three information display formats for presenting instructions to helicopter crews regarding which of several potential targets they were to engage. Text (alphanumerics) graphics (symbols), or numeric (digits) instructions were presented visually for the subsequent target acquisition search task while subjects were either serving as copilot or simultaneously piloting a simulated aircraft. The goal was to determine which format was most compatible with the subsequent target acquisition task and which would be least detrimental to flight performance.

One finding was that the instructions using text required the longest time to acknowledge and were perceived as the most difficult to use. Performance was faster with the numeric format than with the symbolic format, although when co-piloting, participants reported that the symbolic format was the easiest to use.

2.11.7 Mammographic screening

The national breast screening programme, recently instituted to detect early breast cancer by systematic X-ray examination of women, is expected to generate over 1.5 million breast X-rays (mammograms) per year. At early stages of the disease, the relevant signs may be subtle, and high error rates have been reported. Computer based methods may provide a means of improving the accuracy and objectivity of mammographic interpretation. Computer-based cue generation methods can be designed to respond preferentially to a particular characteristic (e.g. edge, shape or local grey-level properties) of the targets they are intended to detect. However, such methods will also respond to non-targets which share the target characteristic, and their performance in complex or noisy images may therefore be poor. By combining information from a number of independent image cues, each responding to a different characteristic of the target, the specificity and robustness of detection can be improved. Astley *et al.* (1991) have presented results to demonstrate the validity of this approach.

2.11.8 Industrial inspection

Whether by humans or machines, industrial inspection may be regarded as a multistage process involving at its core visual search and decision-making. Implications for this model for both the practice of inspection improvement and theories of human behaviour are represented by Drury (1990). In particular, models of the search process are used to understand the variables involved in visual search. Recent work aimed at defining the optimum time to spend inspecting a single item or product is presented. The mathematical consequences of these models have serious implications for inspection tasks. Particularly, the scanning strategy (i.e. how an inspector's inter-fixation distance is chosen) becomes critical. Drury presents an experiment showing that humans can indeed behave optimally in a search task. Drury comments that for the future, understanding of human and automated inspection in the same terms will be necessary to ensure optimum allocation of function between human and machine components of inspection tasks.

Chapter Three

The Role of Array Shape in Search Tasks

3.1 OVERVIEW

Reducing search time for computer menu displays and other VDU work has considerable value in terms of potential financial savings. One area which applied psychologists can consider is the optimal design configurations of display screens for typical HCI tasks. There are implications for both windowed interfaces and for the issue of desk-top publishing. As a preliminary step, a laboratory type visual search task was used to investigate performance effects with various array shapes. Contrary to expectations based around the 'newspaper column' phenomenon, subjects performed significantly better with an extreme landscape array. In order to examine possible reasons for this, eye movement studies were conducted using a scleral coil technique. Results showed greater saccade amplitudes for horizontally elongated formats. In addition, a model based on visual lobe differences was proposed. A further eye movement study was conducted using reading material displayed within different array shapes, but no significant differences in eye movement parameters emerged. It was concluded that alphanumeric information (although not text passages) presented on a VDU is more easily searched for and located when arrayed in horizontally elongated configurations.

3.2 GENERAL INTRODUCTION

Computer users interact with an enormous number of displays. Tullis (1983) reports that employees of a single insurance company will view over 4.8 million displays per year. Employees in the Bell System using only one particular software package will extract information from over 344 million displays per year. The reduction of only a fraction of a second in the time it takes users to process each display could lead to large savings in both time and money. Tullis further estimates that 55 man-years of labour costs can be recouped by bringing about a mere one second reduction in the time users spend analysing the information on each screen of a computerised service system accessed in a given year. Visual search is one of the most common subtasks carried out with displays. It has been demonstrated by Fisher *et al.* (1989) how highlighting, for example, can form one effective way of reducing search times. Factors as diverse as array dimensions, position and conspicuity of information, choice of iconic versus verbal labelling, text font style, optimisation of peripheral vision, etc. perhaps could all enhance search performance. The first experiment concerned the potential for VDU screen shape to optimise search times required to locate target characters.

The screen format of VDUs has traditionally followed the 'TV model' in its horizontal × vertical rectangular dimensions. A typical VDU screen ratio is in the region 1.4: 1 horizontal to vertical area. However, this trend is changing. Firstly, we are witnessing greater diversification in TV screen dimensions, with 'square' to 'cinemascope' choices becoming available. Will VDU design now follow this 'tube-technology'? Secondly, the proportion of non-cathode ray tube (CRT) VDUs is increasing; and therefore the format considerations of new screens are less restricted. Thirdly, and most importantly, the use of multiple-window systems allows for a wide variety of virtual windows to be created simultaneously within the screen area. It is, for instance, conceivable that a horizontally elongated (landscape) format would enhance performance on the more graphical tasks, as this is more consistent (within an ecological model; e.g. Gibson, 1979) with our natural horizon-based perception. Equally, learning may have brought about a readiness to perceive written information in something akin to $A4$ layout.

3.3 SEARCH AND SHAPE

VDUs have evolved from static, linear entities of sequential data lists into rich and complex graphic environments which communicate on a more 'immediate' level. Windows can be regarded as the conceptual framework for the capture, development, organisation, and highlighting of information, whether textual or graphic. Gait (1986) defines 'pretty windows' as those windows which have the dimensions of golden rectangles—the ratio of their height and width form Euclid's golden ratio: approximately 1.618—but also on a more general aesthetic level he includes squares within this. Psychologists, artists, architects, printers, calligraphers, product designers, advertisers, and others responsible for creating visually pleasing objects have used golden rectangles for many years. The empirical foundation for this preference rests on a series of experiments performed by Fechner in 1871, where subjects demonstrated a preference for rectangles of dimensions with a ratio of 1.62—close to the golden ratio's approximate 1.618.

Gait notes that in view of Fechner's discovery, display interfaces which encourage pretty windows and pretty panes within windows can benefit by a cognitive ergonomics advantage, with the user environment leading to greater productivity by maximising the display's communications potential. He also notes that current product advertisements offer suitable sample windows for an *ad hoc* experimental verification of Fechner's findings. An analysis of one such advertisement showed that the distribution of the relative dimensions of the 72 windows depicted had golden rectangles as its primary peak, while a secondary peak corresponded to square windows.

Thus, the aesthetics of 'golden rectangles' within screens and windowing systems has been considered by Gait. Here, we address the *functionality* of different configurations. As a laboratory measure of task performance whilst using various aspect ratios, a simple classic visual search task was chosen. Visual search is required whenever a visual target is not immediately apparent, and as such is a task generalisable widely to 'real-world' VDU tasks. The degree to which search area is defined, the method by which the boundaries are marked, and its size and shape all may vary. In situations where search aids such as cues or sector markings are found to be of assistance they probably reduce the effective search area.

3.3.1 Previous studies

An extensive literature search revealed only one series of studies relating visual search and display shape (Bloomfield, 1970). Concerning the shape of the search display, Bloomfield (1970) described three experiments in which the shape was varied. In all three the displays varied in their horizontal and vertical extents.

In the first, three displays of constant area were used. They were a vertical rectangle (8 × 32 stimuli), a square (16 × 16), and a horizontal rectangle. Only seven observers were used and there were large individual differences. The mean times for search increased as the display changed from vertical rectangle (3.2 s) to horizontal rectangle (4.6 s) to square (5.3 s). There was a large variability in times on particular trials, and these differences were not significant.

In a second experiment, in which much more extreme examples of vertical and horizontal displays were used, observers were required to search through the display for every stimulus in turn. Use was made of an eye movement monitoring technique described by Oldfield (1960). Material consisted of one-inch circular discs showing two three-letter combinations. When presented with a display of discs, the observer commenced with a particular disc, read the second group of letters on it, then searched the first letter groups of the other discs for the same letters. When these were located, the new second group was read and search recommenced. This procedure continued until the observer arrived back with the first stimulus.

Each set was arranged in eight displays of varying horizontal and vertical dimensions (30 × 1, 15 × 2, 10 × 3, 5 × 6, 6 × 5, 3 × 10, 2 × 15, and 1 × 30). Two observers searched sixteen displays (eight shapes and two arrangements of letters) which were arranged in random order, in each of three sessions. Search time was measured with a stopwatch. The first session was considered as practice, and data from the second and third sessions were pooled. There was little difference in times for the second set of vertically printed letters. For the first horizontal set of letters, however, the displays could be roughly divided into two groups. The three with the greatest vertical extents (30×1, 15×2, and 10×3) produced times similar to the vertical set of letters. However, the remaining five displays, two almost square (6×5 and 5×6) and three with the greatest horizontal extents ($3 \times 10, 2 \times 15$, and 1×30), were searched quicker. In other words, the vertical displays required more time than the square or horizontal ones for horizontally oriented alphabetic material. With vertically oriented material, the times were all as long as for the vertical displays for the horizontal orientation.

In the third experiment, with targets differing in size from non-target discs, significant differences were found between the cumulative distributions although no clear pattern emerged. The horizontal display was half as long and twice as wide as the square, whilst the vertical display was twice as high and half as long. Projected, their sides subtended visual angles of approximately 8°–10° and 32°–40°. There were two runs with each target in each session, and five sessions with each display. The sessions were randomised between display shapes. There were 100 readings per target per display for *one* practiced observer. For the target most different in size from the background, both horizontal and square displays were superior to the vertical ($p < 0.01$). For the intermediate target, performance on the horizontal and square were better than the vertical ($p < 0.01$). Additionally, the square was superior than the horizontal ($p < 0.05$). For the target closest in size, performance on the vertical was superior to either of the other two displays ($p < 0.01$). From Bloomfield's depiction of these results, it was apparent that for considerable sections of the graphs, the horizontal and vertical displays were superior to the square although the differences were only significant when they were inferior. He remarked that: 'It is difficult to draw any conclusions about this experiment. For a difficult target the vertical display seems clearly to require less time than the horizontal and square displays, while for an easy target just as clearly, it requires more... One might conclude that display shape may have a consistent effect, if the display contains material to which the observer has developed stereotyped responses. The majority of search tasks probably do not enter into this category. In them there may be display shape effects though they are probably much less consistent' (pp. 120–121). This suggestion is returned to within §3.16.

3.3.2 'Exhaustive, but inefficient search'

Bloomfield produced several equations for efficient, exhaustive coverage of the display and for an independent 'glimpse' strategy. It is sometimes assumed that these two strategies will result in the best and worst possible search performance respectively, and that all performance data should lie within these limits. In fact, neither strategy is very likely to be adopted by observers for most search tasks. The observers used for the experiments reported by Bloomfield were asked to comment on how they carried out their searches. They all described regular strategies of one kind or another.

For example, the most common strategy for the more difficult targets in the regularly arranged display consisted of a horizontal scan across the top of the display, followed by a short downward movement, then another horizontal scan across the display and so on until the display had been covered. This, is here termed a modified reading strategy, or 'snaking'. Another strategy for the same targets and display involved a spiralling scan, beginning in one corner and moving in a clockwise direction into the centre of the display. As a further example, the most often described strategy for the irregular displays was similar to the first mentioned horizontal scanning method. First one block of stimuli, then another, was searched; the order of blocks being determined by the overall scan pattern.

Although subjective descriptions may not match an objective record of eye movements, it was clear from their reports that the observers were attempting to use regular scanning strategies of one sort or another. It also appeared that they attempted to use the same strategy on most of their searches. The filmed eye movement records were inadequate to use for precise eye position analysis.

After the above considerations of observers' comments, experimenter's impressions, Enoch's early (1959) description of eye movements, the difficulties of efficient coverage when the display area is not very large compared with the visual lobe area, and the expectation that observers will naturally adopt an exhaustive inefficient coverage strategy, Bloomfield then built up a somewhat different picture of search. In this, the observer is attempting to scan a display in a methodological fashion. His scan pattern will probably vary slightly from trial to trial. However, he basically tries to cover the whole area of the display by looking at small adjacent areas of the display one after another, making sure that he does not leave any areas uncovered. As a result, Bloomfield suggests, the steps between the fixations are probably small compared with his visual lobe area. Additional fixations would produce new information. They will also yield an increasing amount of information that has already been available. When the observer detects a target in the retina, he will then make a final confirming fixation. Also, if he is relatively inexperienced, he may make other eye movements in order to carry out comparisons between a suspected target and other display elements.

Bloomfield further remarked that if the average distance between successive fixations is very small compared with the visual lobe area axis, then the search strategy will be very inefficient and will result in very long search times. As the fixation distance increases, the coverage will become more efficient and search time will decrease. If the fixation distance continues to increase, eventually it will be such that there is the minimum overlap producing the most efficient possible search performance for that display and visual lobe size. After this, as the distance continues to increase, search performance will again decrease. It could reach a point at which it produces search times equivalent to those that would be achieved with an independent fixation strategy. These times would not be as poor as those which were produced when the difference between the average fixation distance and visual lobe area axis was very small. The fixation distance will not become as great as this for any observer, and it is even less likely to increase after this. If it did so, the times would also continue to increase, particularly when the fixation distance was large enough that two successive fixations would not both fall within the search area.

EXPERIMENT 1: A STUDY IN SHAPES

3.4 INTRODUCTION

The aim of this first study was to expand on these findings with particular relevance for VDUs. It was intended to apply a classical visual search method to this contemporary issue. In order to shed light on the search strategies generally employed, it was also intended to follow up this study with a fine analysis of users' eye movements.

3.4.1 Visual lobes

The notion of 'visual lobes' was introduced in §2.4. Although its dimensions vary with type and difficulty level of task, it is known that in a task such as that employed within the present study, the visual lobes are elliptical rather than circular (e.g. Rayner & Fisher, 1987).

Over the years, many investigators have attempted to relate concepts such as the 'visual lobe', the 'perceptual span', the 'useful field of view', or the 'area of conspicuity' to eye movement control in visual search and in reading (e.g. Paterson & Tinker, 1947; McConkie & Rayner, 1975; Mackworth, 1976; Engel, 1977; Bouma, 1978; Jacobs, 1986). Several studies have also tried to decompose into perceptual, attentional, and linguistic components the factors that influence this limited area from which information is available in visual search and in reading (O'Regan, 1979; see Rayner, 1983, or McConkie, 1983, for reviews). One of the few attempts to systematically predict eye movement parameters directly on the basis of quantitative measures of visual constraints imposed on eye movement behaviour is that of O'Regan *et al.* (1983). In this attempt, which was designed to distinguish between the effects of early visual mechanisms and higher level processes on perceptual span, this team proposed a definition of visual span as the size of the region around the eye's fixation point within which letters can be recognised with a given accuracy, without the use of contextual information. The limits of this span depend on purely visual factors, such as viewing distance, letter similarity, or inter-letter spacing, and can be estimated using the model proposed by O'Regan (1983). Within the hypothesis that there should be a relationship between the limits of this span and eye movements in reading and visual search, saccade sizes should be found to parallel changes in visual span, induced by varying visibility conditions (e.g. viewing distance, letter spacing). However, O'Regan *et al.* (1983), using a reading task, and Lévy-Schoen *et al.* (1984), using a complex visual search task, failed to show a direct dependence of eye movement parameters on visual span changes. It was suggested that, in such tasks, higher level factors related to linguistic processing and cognitive load, determined eye movement behaviour more than did the spatial sensory limits constraining visibility of letters at each fixation.

Stemming from that work, Jacobs (1986) suggested that if cognitive load could be reduced to the very minimum in a visual search task, then the control of saccades should depend directly on the sensory constraints limiting the number of letters visible around the eye's fixation point. The global hypothesis was that, under such conditions, each saccade should bring the eye to a zone where new visual information can be gathered; i.e. the limit of the visual span. As a consequence, if visual span is altered as a function of the experimental visibility condition, then saccade sizes should, on average, change in parallel. This extreme hypothesis, which considers eye movement behaviour as a direct function of spatial sensory limits, was termed the 'visual span control hypothesis'.

In a psychophysical experiment, three methods were used to manipulate the *spatial* visibility limits (visual span), as measured by a psychophysical procedure; changing viewing distance, inter-letter spacing, and target-background similarity. The results of this experiment were then used as a reference for predicting mean saccade sizes and fixation durations in a visual search task in which the same visibility changes were made. About 80 percent of the variance of mean saccade sizes could be accounted for by adjustment of saccades to changes in visual span, so the visual span control hypothesis was confirmed. As for the *temporal* characteristics of scanning behaviour, less than 50 percent of fixation duration variance seemed to be determined by visual span changes. It was suggested that other, higher level factors, possibly related to decisional processes intervening in the triggering of saccades and the computation of their spatial parameters, might play an important role in determining fixation durations in a simple search task.

To avoid the problematic higher level factors related to linguistic processing and cognitive load, noted by previous workers such as Lévy-Schoen *et al.* (1983) and Jacobs (1986), the stimulus materials in the present study (Expt. 1) were limited to random, alphabetic characters.

3.4.2 Effects of line length

Although the interest in the effects of line width upon legibility of print and its educational implications go back more than a century, relatively little research has been performed in this field within the past fifty years. Ophthalmologists, educators, psychologists, advertisers, editors and typographers from 1881 to 1923 made casual observations and carried out a few uncontrolled measurements. These led to definite recommendations based on opinion and inadequate data. This material was reviewed in detail by Pyke in 1926. Although there was a range of opinions, most listed the optimum line width at about 22 picas (about 3.7″; one pica = about $1/6^{th}$ inch) for 10 to 12-point type. 'On the whole, the suggestions favor use of relatively short lines. But the recommendations tend to be somewhat contradictory. Obviously, carefully controlled experimentation is needed to determine the range of line widths that may be used safely with each size of type' (Tinker, 1963, p. 74).

A wide diversity of practice is employed for line widths in all manner of printing. A total of 1500 journals and books were surveyed by Paterson and Tinker in 1940 (1940*a*). With double-column printing, the line widths used in non-scientific journals were spread quite evenly from 14 to 22 picas. However, in the double-column scientific journals, most were printed in 13 to 18-pica line widths. In single-column printing, there was greater uniformity, with journals using mostly 23 to 18-pica line widths. Most of the textbooks concentrated on 19 to 24-pica lines.

In an early study, Starch (1923) required 40 subjects to read the same material printed in line widths of $1\frac{1}{2}''$ (9 picas), $2\frac{3}{4}''$ (16.5 picas), and $5''$ (30 picas). The material in 16.5 picas was read 16 percent faster than that in the 9-pica width, and 7 percent more quickly than the 30-pica text. Starch considered that an optimum line width would be somewhere in the region of 18 picas.

Paterson and Tinker (1940*b*) analysed the eye movement pattern for reading 10-point set solid (no leading) in a 19-pica line width in comparison with 9 and 43-pica lines. For the 9-pica versus 19-pica line, the fixation frequency, pause duration, and perception time were significantly increased. Readers apparently were unable to make maximum use of peripheral vision in reading the 9-pica lines. For the 43-pica versus 19-pica line, the fixation frequency, pause duration, perception time, and regression frequency were all significantly increased. It is worth noting that regression frequency was increased by 56.7 percent. With this excessively long line of 43 picas (about $7.17''$), the major difficulty appeared to be in accuracy of relocating the beginning of each new line when the return sweep was made. It is probable that this difficulty so disrupts the reading process that re-establishment of the most efficient oculomotor patterns in reading each successive line becomes impossible.

However, it is likely that there must be some optimum line width, not in terms of absolute line width as may be construed from Tinker's statements, but in terms of subtended retinal angle. In the materials described here and within Experiment 2, although line width was longer than this in the 25×100 array, the subject was seated with a viewing distance of 80 cm, which is considerably more than a typical reading distance.

3.4.3 The edge effect

Also of note for the present study is the 'edge effect'. In 'free search' tasks, i.e. where there is no fixed order of search and the observer is free to construct their own scanpath through a display, biases are found in the number of fixations given to different parts of the display. In particular, the outer half of the display (i.e. that broad border constituting 50 percent of total area) appears to get, on average, fewer fixations than the inner half (Enoch, 1959; Ford *et al.*, 1959). The lower the average fixation density for a particular part of the display, the slower will be the average search times for targets occurring within that area. Therefore, targets in the outer half of the display have generally slower search times than those in the inner half (Baker *et al.*, 1960). This 'edge effect', as it has come to be known, 'appears to be one of the more robust effects in visual search' (Monk, 1981*a*, p. 615).

Despite its practical importance in military and industrial settings, little research seems to have been carried out on the edge effect since the pioneering studies. A few experiments have found an edge effect incidentally (e.g. Monk, 1974, 1976, 1977), but possible interactions of the effect with other search factors appear to have largely been neglected. The only exception seems to be the work of Monk (1981*b*) which explored the interaction between the edge effect and target conspicuity using manipulations of both factors. Thus, the applied questions to be answered were of the lines: 'What does increasing the target conspicuity do to the magnitude of the edge effect?' and 'What does increasing the magnitude of the edge effect do to the search times of conspicuous and inconspicuous targets?'

Monk concluded that the edge effect appeared to be unaffected by changes in target conspicuity. By contrast, when the magnitude of the edge effect was varied (by altering the edge ratio (ER); see Monk, 1971), there were different consequential effects on conspicuous and inconspicuous targets. As the ratio of search times of outer to inner targets decreased, that of inconspicuous to conspicuous targets increased. The results were explained in terms of a model based upon the size of the mesh of fixations used to cover the display. It was suggested that the mesh and the magnitude of the edge effect appeared to be inextricably linked.

3.4.4 Experimental rationale

This first experiment was primarily designed to explore the effect of array shape on search performance. As the edge effect has been regarded as a robust phenomenon, it might be expected that this would interact with a shape effect, if not manifest itself as a main effect. From what was outlined in §3.4.2, it was hypothesised that the vertically elongated arrays would demonstrate significantly faster search times than more horizontally elongated arrays.

3.5 METHOD

One hundred files of random alphabetic (upper case) arrays were generated by computer programme. Blank spaces were also included. The material may therefore be described as 'irregular'; regular search material consisting of a uniform spread of characters. These files consisted of five different configurations: 25 characters high by 100 characters wide, 30 × 83, 50 × 50, 83 × 30, and 100 × 25 characters high and wide. Examples are shown in Figures 3.1 to 3.5. Area covered remained constant. All files contained only one *X*, the line position being checked and edited to ensure no systematic location bias. Also, the position of the *X*s were edited so that in half of the files of each dimension, these characters were positioned in the inner 50 percent of the array and in the outer 50 percent in the other half of the files. From a hard copy print-out of these arrays, monochrome photographic slides were produced with positive polarity.

Slides were presented in randomised order to subjects via a carousel projector, with the sequence being changed for each subject. The projector was positioned 264 cm from a rigid white screen. Subjects were seated so as to have a viewing distance of 50 cm. An f = 150 mm lens was used. Screen images were projected at eye level and were of the following dimensions: 13.5 × 33.5, 16.5 × 27.25, 27.25 × 16.5, 10 × 45.5, and 55 × 8.5 cm height by width. Thus, images consistently occupied an area of some 450 cm^2. Each letter was approximately 4 mm high, and extended over about 0.46° of visual arc. Screen luminance was 33 cd/m^2. Ambient illumination was approximately 23.0 lux, measured vertical to the screen, achieved with two fluorescent strip-lights in an otherwise darkened vision laboratory. This was regarded as a compromise between improving display visibility whilst retaining some semblance of a 'real' environment.

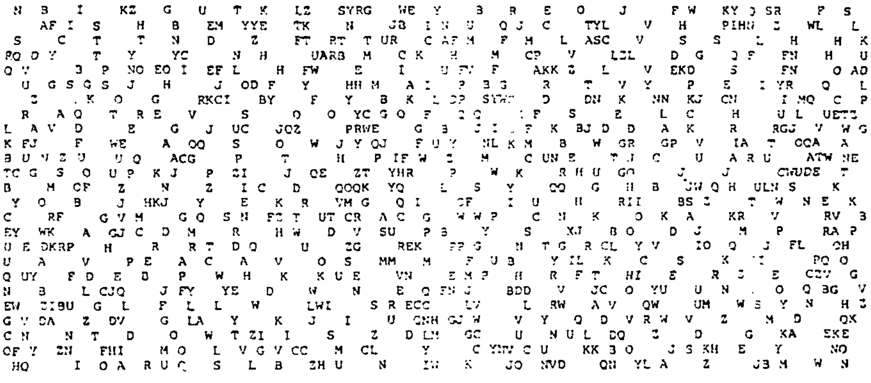


Fig. 3.1: An example of a presentation in 25 × 100 format.

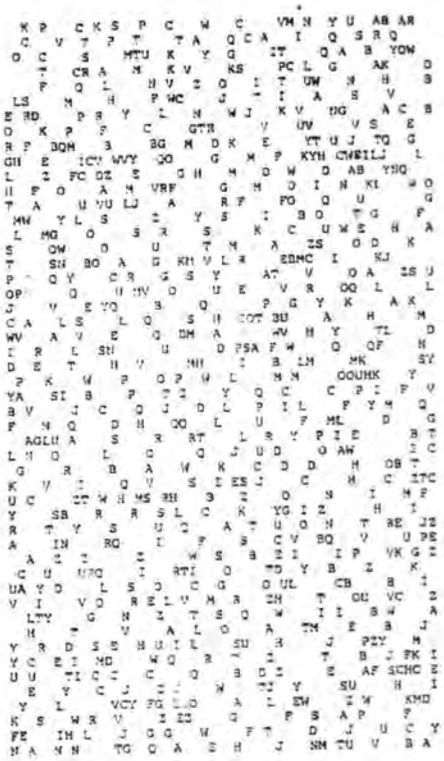


Figure 3.2: An example of a presentation in 50 × 50 format.

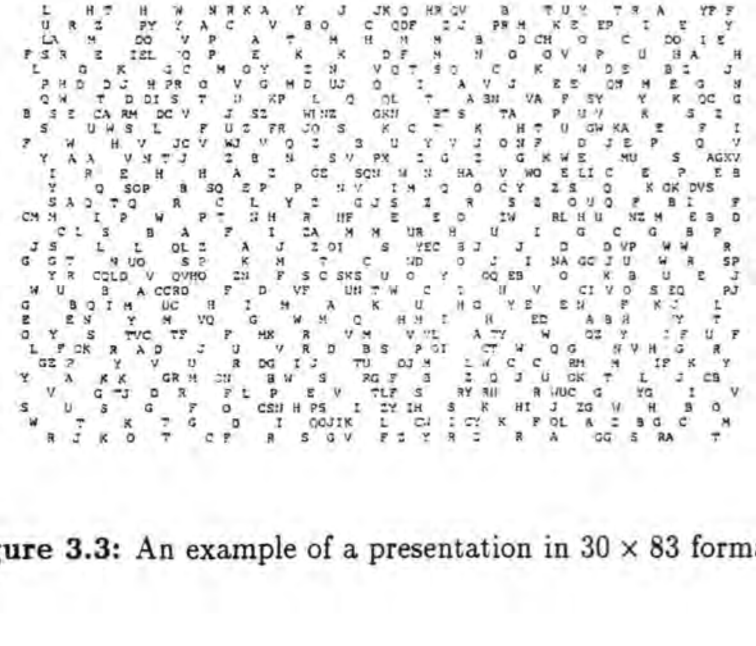


Figure 3.3: An example of a presentation in 30 × 83 format.

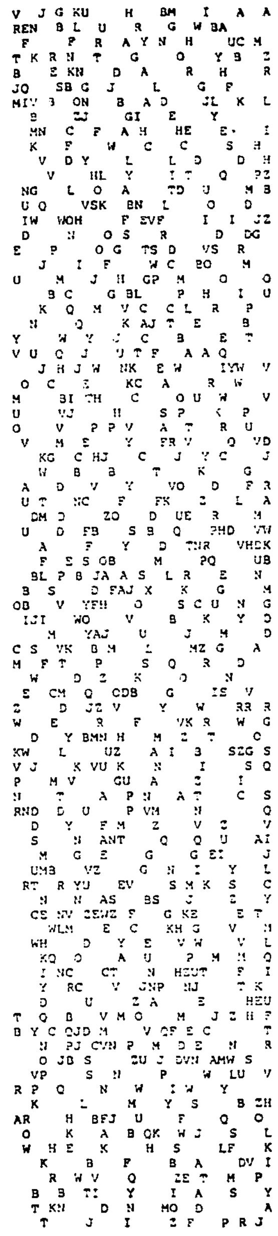


Figure 3.4: An example of a presentation in 83 × 30 format.

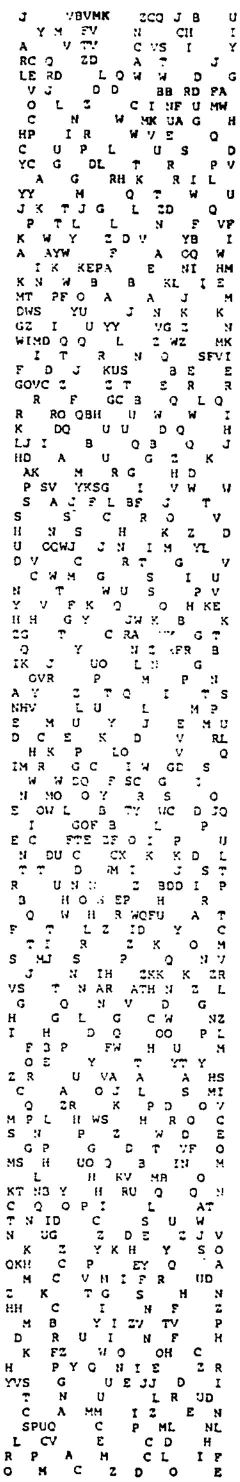


Figure 3.5: An example of a presentation in 100×25 format.

The 20 subjects (9M: 11F), who were paid, had an age range of 13 to 55 ($\bar{X}$ = 23.8, sd = 11.2). They were drawn from a wide range of the population. Eye sight (corrected where appropriate) was tested by an abridged version of the 'Curpax' (passage N8) Test Type (Curry & Paxton, 1980).

Subjects were instructed that they would be shown a total of 100 slides and that they were to search for the target letter X. They were informed that they could use whatever strategy of search they wished. Full instructions are included in Appendix 3.1. Thus, a 'free search' strategy as opposed to a 'serial' or 'systematic' strategy was allowed. They were required to press a morse key when they had located the target letter. This key, along with the projector remote control, was wired to an electronic timer. They were required to state to the experimenter what letters were immediately to the left and right of the X in order to verify their success.

A total of 2000 data recordings were thereby obtained.

Pilot runs (in which one subject took over eight minutes to locate the target) suggested that after some time, strategies became increasing more erratic and the probability of locating the target at any particular point in time progressively decreased. Therefore, a cut-off point of 180 seconds was taken and trials were terminated if the target had not been located within this time.

Subjective comments regarding the strategies of search used by each subject were also recorded.

3.6 RESULTS

Table 3.1 depicts overall search times (s) on the five arrays, inner versus outer performance, and mean overall search times.

Mean search time 51.9 (20.1)					
Inner 53.2 (18.9)			Outer 50.7 (21.2)		
	Array dimensions				
	25×100	30×83	50×50	83×30	100×25
	47.3 (20.7)	55.6 (19.6)	49.2 (18.6)	49.9 (21.0)	57.6 (19.5)
Array dimensions x inner vs outer position					
Inner	50.0 (20.9)	59.6 (20.6)	50.8 (15.0)	45.8 (20.6)	59.6 (17.1)
Outer	44.5 (20.6)	51.4 (18.0)	47.7 (22.0)	54.1 (21.0)	55.7 (22.0)

Table 3.1: Breakdown of search times.

A two-way repeated measures ANOVA was conducted on the data. There was a significant difference ($F(4, 76) = 3.75, p < 0.01$) between array dimensions, no significant differences between inner or outer effects, and no interactions between these two variables (see Table 3.2).

Source	df	SS	MS	F	p
Subjects	19	42722.711			
Position	1	305.606	305.606	1.410	NS
Subs. x Pos.	19	4117.640	216.718		
Dimensions	4	3142.349	785.587	3.746	< 0.01
Subs. x Dims.	76	15937.315	209.701		
Pos. x Dims.	4	1605.397	401.349	2.364	NS
S x D x P	76	12900.781	169.747		
Error	180	38009.087			

Table 3.2: Two-way repeated measures ANOVA conducted on array dimensions and position.

A Tukey's test conducted on the means of the array dimensions revealed significantly faster search times ($p < 0.05$) for the 25 high $\times$ 100 wide array compared with the 100 high $\times$ 25 wide array ($WSD = 8.563$). There was a non-linear pattern in search times across the array shapes between these two extremes but the differences between these intermediate measures were not statistically significant.

The results were therefore in the opposite direction to that expected.

Figure 3.6 depicts mean search times as a function of array shape.

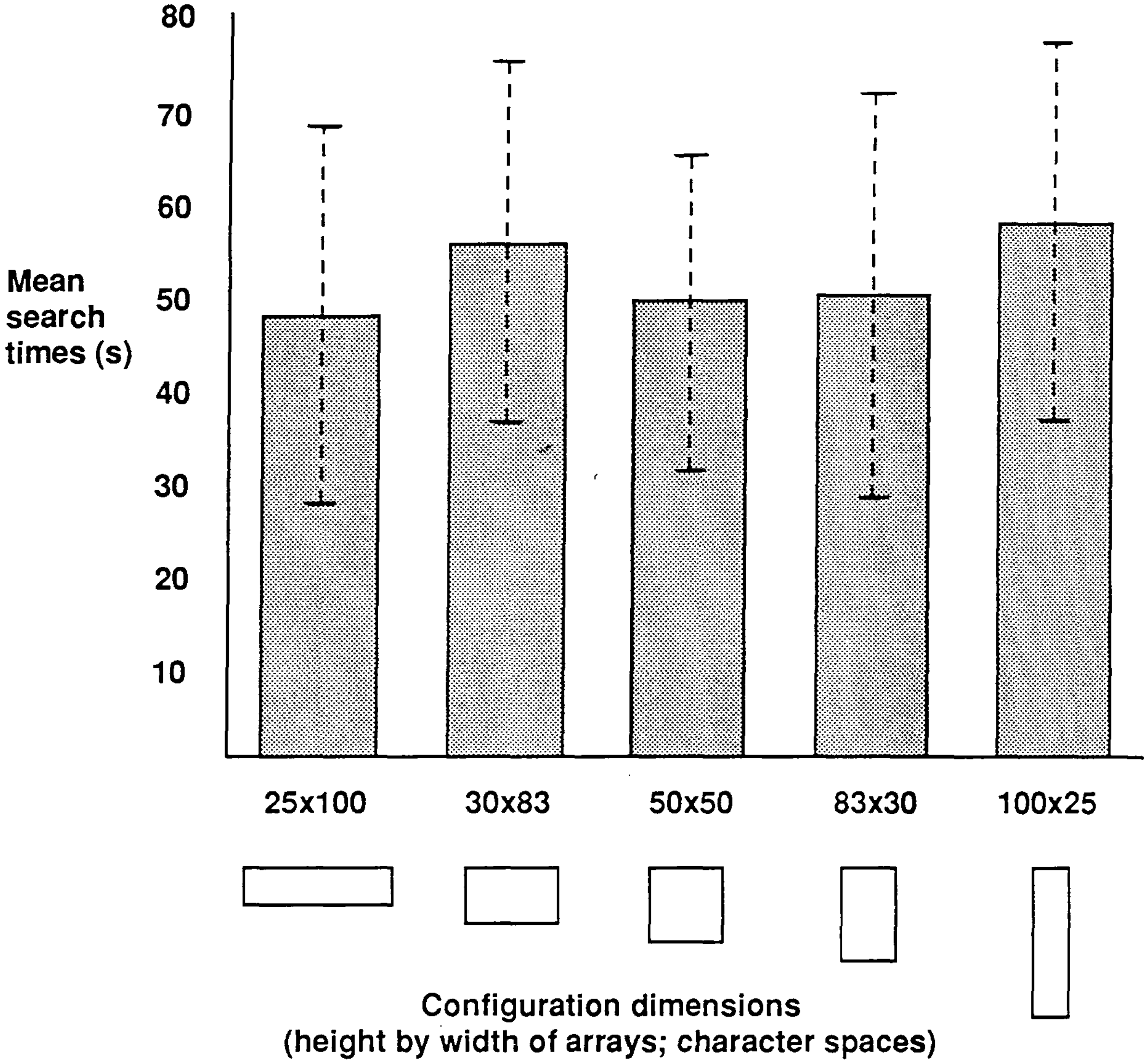


Figure 3.6: Pattern of search time means for the five arrays (random characters search; N=20).

The distribution of search times was positively skewed (see Fig. 3.7). Using a log transformation on the above analyses, similar patterns were obtained for overall scores on the five arrays and mean overall score. Table 3.3 depicts overall scores on the five arrays, inner versus outer performance, and mean overall score.

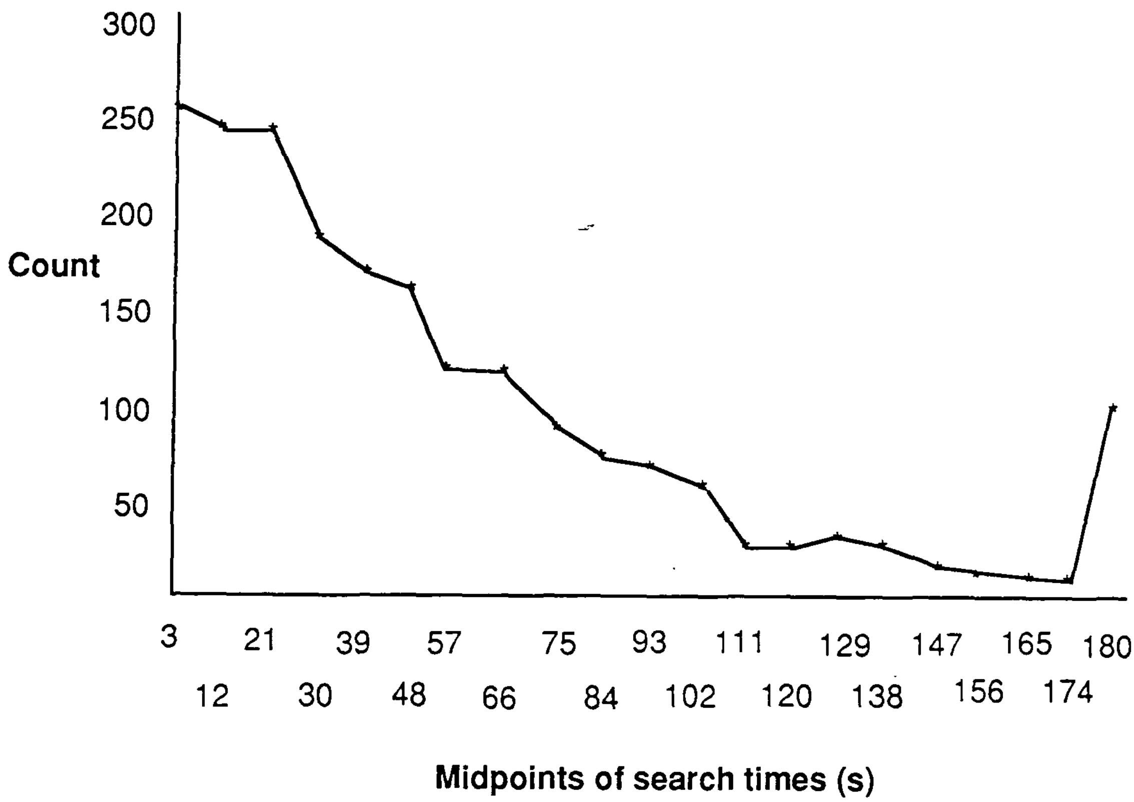


Figure 3.7: Distribution of search times.

A two-way repeated measures ANOVA (see Table 3.4) was conducted on this transformed data. All patterns found still remained and were in fact emphasised. There was a significant difference between array dimensions, no significant differences between inner or outer effects, although there was an interaction between these two variables.

	Mean search time (N = 2000)				
	1.679				
Inner (N = 1000)			Outer (N = 1000)		
1.696			1.663		
	Array dimensions				
	25 × 100	30 × 83	50 × 50	83 × 30	100 × 25
	1.632	1.711	1.656	1.659	1.734
Array dimensions x inner vs outer position					
Inner	1.664	1.752	1.658	1.616	1.757
Outer	1.601	1.670	1.627	1.703	1.711

Table 3.3: Showing breakdown of search times (log transformed) for all slides presented to all subjects.

There were a total of 96 instances in which subjects exceeded the search time limit of 180 seconds. To test for significantly greater frequencies whereby subjects exceeded this limit between the different array presentations and also between the inner versus outer variable, a multiple χ^2 analysis was performed (Table 3.5).

Source	df	SS	MS	F	p
Subjects	19	3.718			
Position	1	0.052	0.052	2.371	NS
Subs. x Pos.	19	0.418	0.022		
Dimensions	4	0.285	0.071	4.176	< 0.01
Subs. x Dims.	76	1.297	0.017		
Pos. x Dims.	4	0.185	0.046	3.214	< 0.05
S x D x P	76	1.097	0.014		
Error	180	3.334			

Table 3.4: Two-way repeated measures ANOVA conducted (with log transformations) on array dimensions and position.

Source	χ^2	df	p
Position	0.666	1	NS
Shape	6.573	4	NS
Interaction	7.132	4	NS
Total	14.371	9	NS

Table 3.5: χ^2 analysis of frequencies of > 180 s search times on variables of position, shape, interaction and total.

3.6.1 Correlating search time and target position

The correlation (r) between sequential position (number of character spaces which would be traversed in a true reading strategy) of the target position as measured in terms of spaces including blanks and the target's position in terms of actual characters was 0.981 ($df = 99; p < 0.001$). The overall correlation between sequential position (number of characters plus spaces) and mean search times was 0.510 ($df = 99; p < 0.001$). The corresponding correlation between search times and target position measured solely in terms of characters was 0.520 ($df = 99; p < 0.001$). The diversity of systematic/asystematic strategies was represented by the fact that individual subjects' correlations between search times and character position (excluding spaces) ranged from −0.049 to 0.675 with a mean correlation of 0.263 ($sd = 0.217$) (df = 19; NS). A scatterplot of a subject with an apparently highly systematic search strategy is shown in Figure 3.8. Figure 3.9 illustrates a typical pattern found from a subject who generally applied random strategies.

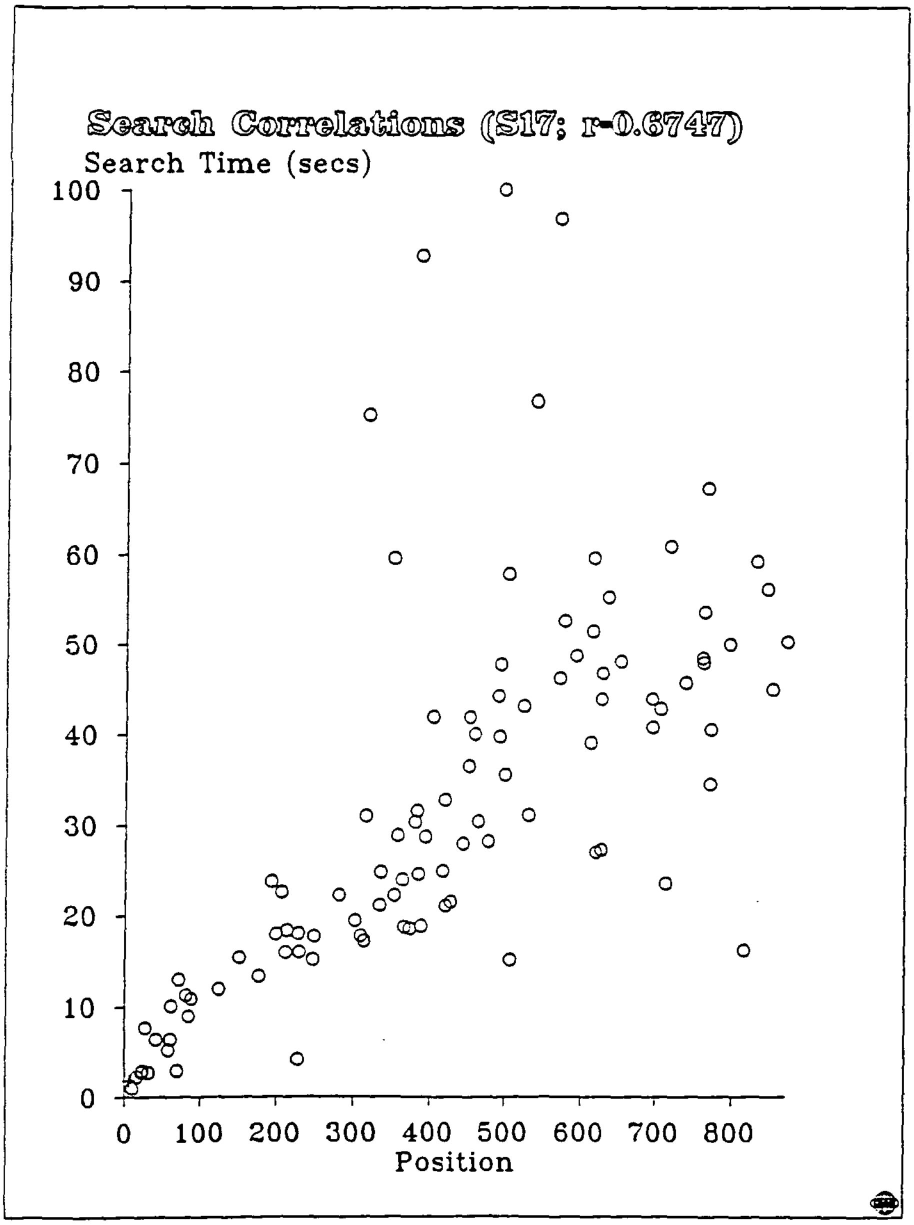


Figure 3.8: Scattergraph of one subject (S17) demonstrating a high correlation between search times and serial target position (systematic search strategy).

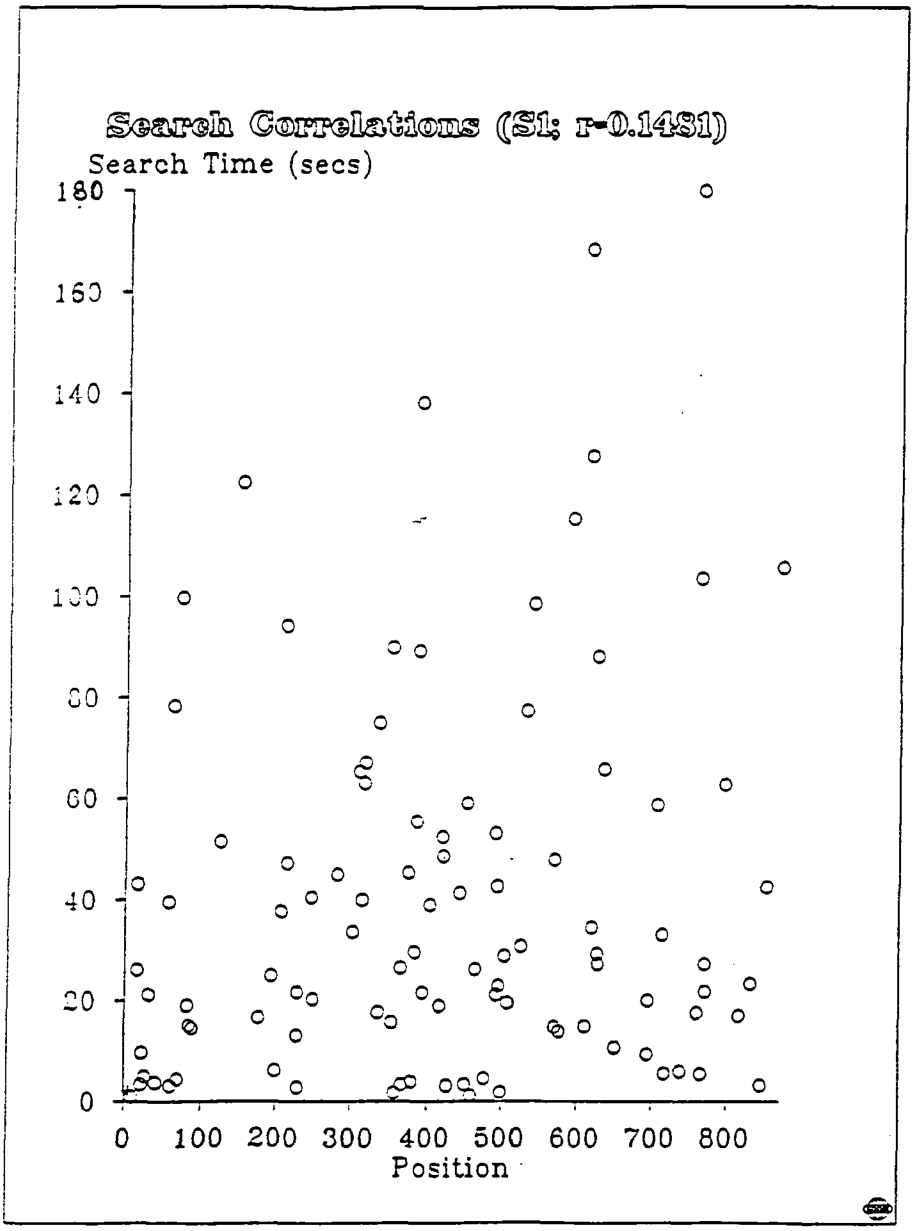


Figure 3.9: Scattergraph of one subject (S1) demonstrating a low correlation between search times and serial target position (random search strategy).

3.7 DISCUSSION

With vertically elongated material (e.g. newspaper columns) shorter return sweeps are naturally made to the beginning of the next line to be read. These shorter saccades are less likely to under- or over-shoot and thus locating the start of the line is less of a problem (see §3.4.2). As such, columnar material is generally felt to be easier to read. Due to the apparent preferences for newspaper column style vertically elongated presentations of text, it might have been expected that should a difference emerge within the present study, performance would have been superior on the extreme vertical array. In fact, the opposite pattern emerged.

So, actual performance suggested that people's visual search on this sort of task was superior on horizontally drawn out material despite subjective comments suggesting a general preference for the vertically elongated displays. Two or three subjects made comments of the nature that they found themselves reading the vertically elongated arrays so fast that they felt they were not fully attending to the content: 'We have been taught to read... should read vertical strips faster... easier for horizontally elongated... Read too fast with newspaper type. I enjoy taking things in.' Another subject reported that with the horizontally elongated arrays he gave a quick look first and then would employ a reading strategy. He added: '... but with wide screens I can't read very quickly... Because when I reach the edge I still think that I have missed something and should go through again... Still concentrating on what I have read before. For me, normal page type is the best... Slower with wide screen'. This may have been responsible for the overall shorter search times to target location, whereas a faster visual search may have produced numerous misses of the target and re-searches.

3.7.1 Relating the differences to visual lobe shape

The above comments suggest that with the extreme portrait arrays, accuracy may have been sacrificed over an unintentional fast speed of reading. This possibility could be tested in two ways. One way simply would be to compare frequencies of $>$ 180-second times across the presentations. There were, however, no significant differences between the dimensions, in terms of frequencies of exceeding the time limit. The second method, and one much more precise, would be to record eye movements over the presentations and to compare measures such as number of fixations, saccades and efficiency of scan paths.

The $>$ 180 s rule frequency distribution over the various configurations was investigated by applying χ^2 procedures. This failed to reveal any significant differences in terms of frequencies of cut-offs between any of the variables. The $>$ 180 s cut-off was, of course, an arbitrary one. Although it can not be categorically stated that choosing an alternative cut-off would not have produced different results, the experimenter's experiences whilst collecting 2000 search records led to progressively greater confidence in the chosen criterion as that best reflecting a point at which search strategies (on this task) break down.

A further possible explanation for this horizontal superiority may be in terms of visual lobe area. As mentioned earlier (§3.4.1), the visual lobe is defined as the peripheral area around the central fixation point from which specific information can be extracted and processed. Rayner and Fisher (1987) found that the size of the span in their random characters search task was very similar to the size of the visual lobe when subjects are reading rather than searching for a letter. In both situations, the span is asymmetric in the direction the eyes are moving extending 3–4 letters to the left of fixation to about 15–16 letter spaces to the right of fixation. As the visual lobes are obviously not square or rectangular, there will always be some minimum amount of overlap. In addition, unless the array dimensions are in some multiple of visual lobe dimensions, there inevitably will be some 'wastage' from lobes extending beyond the array boundaries.

A model is now presented in abstract form and then applied to data from this study. Figure 3.10 shows a classic reading passage from Buswell (1937) complete with fixation points with superimposed visual lobes according to Rayner and Fisher's (1987) findings. Data are that of a good reader on prose of modest difficulty. It should be stressed that it is the horizontal dimensions which are important. The vertical extent of the lobe has been reduced from its likely size in order to avoid clutter.

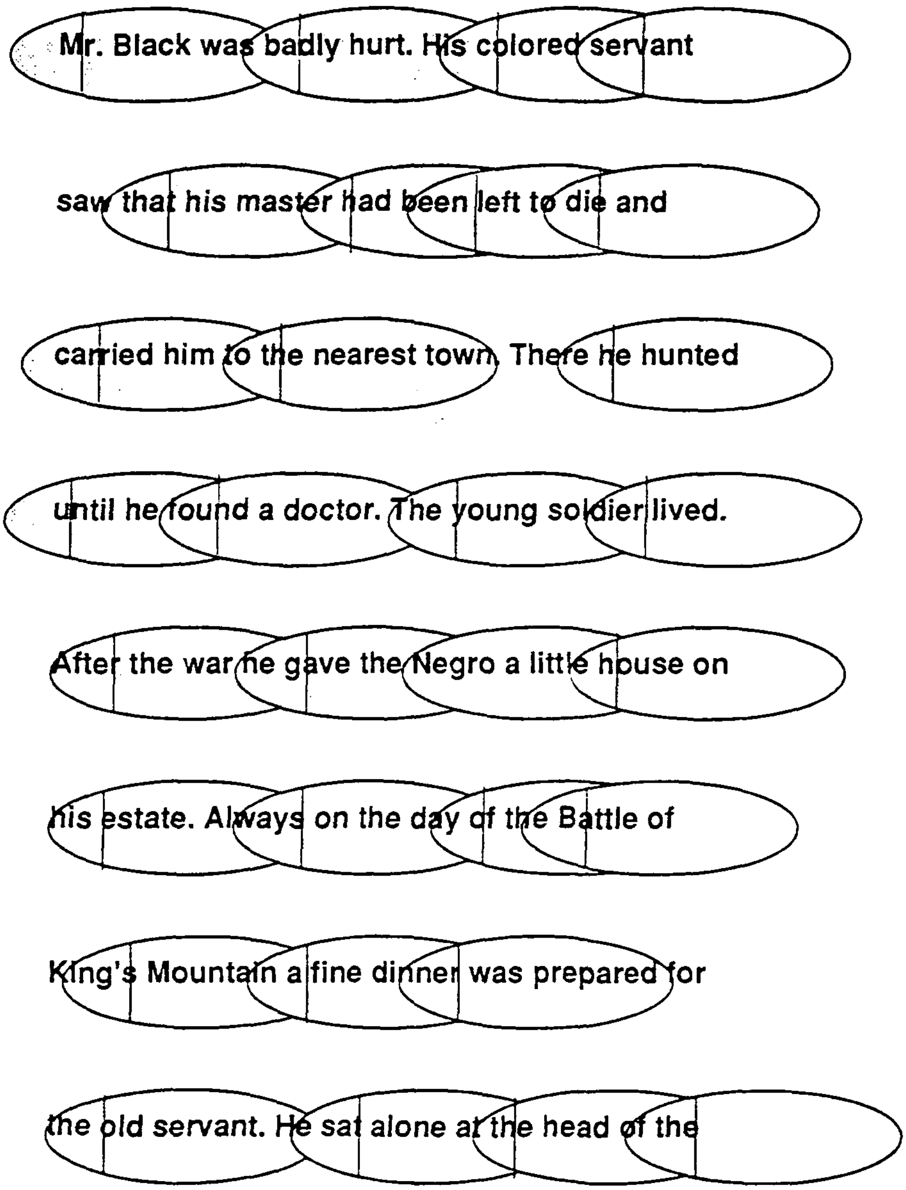


Figure 3.10: Superimposed visual lobe areas upon classic reading data. Vertical lines denote approximate fixation points. Adapted from Buswell (1937).

Figures 3.11 and 3.12 depict this model as applied to actual eye movement data of a subject (AH) showing differences in visual lobe areas covered on the extreme horizontally and vertically elongated arrays. Other subjects demonstrated similarly characterised patterns.

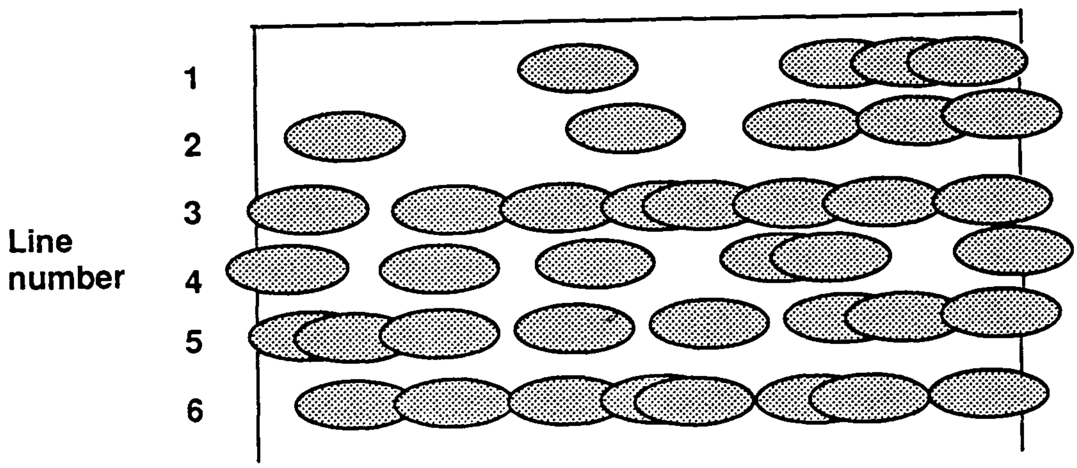


Figure 3.11: Model of visual lobe areas shown during search of 100 characters wide by 25 characters high array: part of actual eye movement data obtained from one subject (AH).

There is a more general reason why one might expect search strategies to be inefficient and to involve considerable amounts of overlap. Much of human behaviour is associated with redundancy. Humans prefer to operate at very high levels of certainty. Typically, more information than the minimum that would be required by an ideal, efficient machine is acquired by man before he makes a decision or takes a particular course of action. In everyday life we typically acquire far more information than we need for our immediate purposes and it is likely that this approach to information acquisition will also be used in experimental situations such as this search task. Thus, one might expect observers to use an exhaustive but inefficient coverage strategy with much redundant information being collected (Bloomfield, 1970). Due to the visual lobe geometry, it is also likely that more 'wastage' will occur with material consisting of shorter lines; i.e. with the vertically elongated arrays.

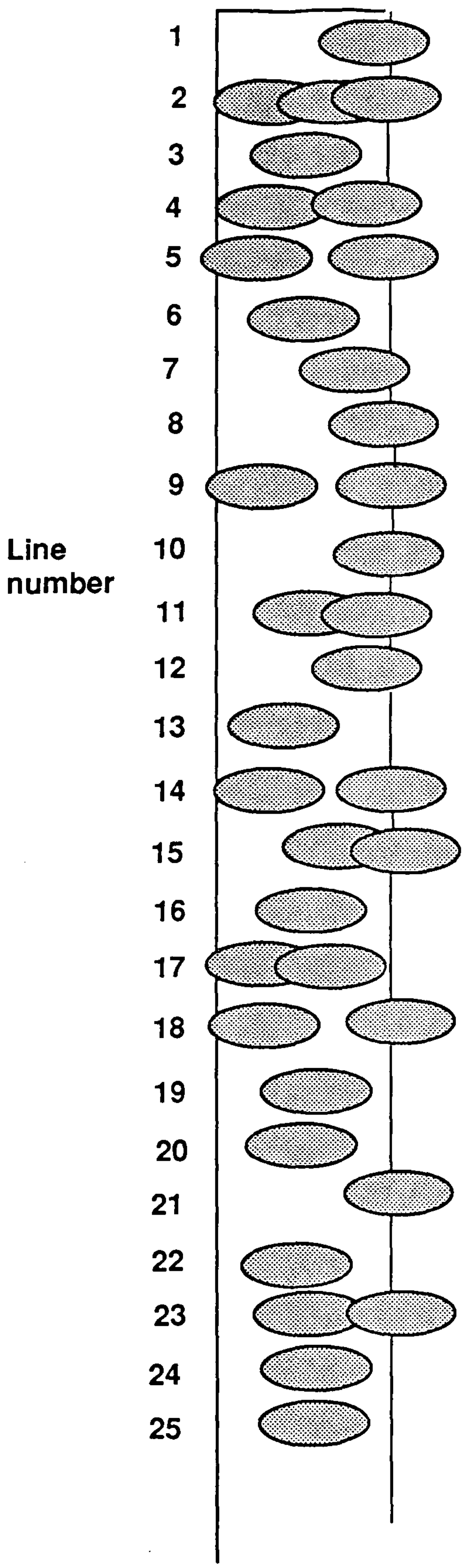


Figure 3.12: Model of visual lobe areas shown during search of 25 characters wide by 100 characters high array: part of actual eye movement data obtained from one subject (AH).

One assumption concerning the visual lobe model rests on the lobe areas being elliptically biased during reading along the horizontal plane. Stark (personal communication, 1990) reports that with Japanese readers, visual lobes are found to be biased down the vertical axis. On this point, it is worth noting that the Chinese language can be written in either longitudinal or latitudinal format. Although traditional Chinese writings were arranged longitudinally, today many technical or scientific writings have been changed into latitudinal format. Hwang *et al.* (1987) have explored the effects of these alternatives on the work efficiency of VDU operators and have offered relevant recommendations for the design of information displays.

It was expected that the edge effect would clearly emerge. However, this phenomena only manifested itself as an interaction with the shape effect within the log transformed analysis, and then in a direction opposite to that expected. The most likely reason for this simply seems to be that some subjects reported that they would often commence their search with a quick scan of the periphery of the arrays. This (subjectively reported strategy) was generally used as a tedium-relieving strategy. Successes in quickly locating the targets within the trials where targets actually were in the outer 50 percent of area are likely to have countered any effect based on fewer fixations to the outer half of arrays. This finding calls into question the previously claimed 'robust' nature of this effect, at least as far as free search strategies are concerned.

As noted in the §3.2, it is possible that the pattern found within this study is linked to mankind's 'horizon-based' perceptual system. However, it is known that acuity falls off peripherally from the fovea almost equally in both the horizontal and vertical planes (Rovamo & Virsu, 1979). It is probable that what is more significant here is an interaction between eye movement tendencies and text line arrangement.

As stated, the findings were incongruous with the 'newspaper column' phenomenon, in which better scanning performance might be expected with shorter line lengths. This, in part, led to the development of Experiment 2.

EXPERIMENT 2: EYE MOVEMENTS AND ARRAY SHAPE

3.8 INTRODUCTION

Although (simply on the grounds of providing a detailed analysis of search) it was intended to supplement the main study with a small-scale eye movement analysis, a further reason emerged which emphasised this need. This involved our reasoning about visual lobes. Although, given the variability in saccade sizes, the idea of visual lobes being arranged neatly to fit into the array framework is likely too simplistic, the different configurations may well induce scanning strategies with different degrees of visual lobe overlap. It was therefore decided to investigate this possibility, using the scleral coil method.

3.8.1 Transition matrices

A specific form of analysis used, and one which was employed within three experiments contained in this thesis, was that of analysis by transition matrices. This particular methodology was introduced in §1.11.1 and it is now time to describe what is involved in more detail here. Other sequential analyses, it must be stated, apart from this sequential transition matrix analysis, are possible (see, e.g., Kleigl & Olson, 1981).

Transition matrix analysis stems from the mathematical principle of Markovian 'chains'. 'A Markov chain is sometimes a suitable probability model for certain time series in which the observation at a given time is the category into which an individual falls. The simplest Markov chain is that in which there are a finite number of states or categories and a finite number of equidistant time points at which observations are made, the chain is of first-order, and the transition probabilities are the same for each time interval. Such a chain is described by the initial state and the set of transition probabilities; namely, the conditional probability of going into each state, given the immediately preceeding state' (Anderson & Goodman, 1963, p. 241). It has a long history in psychology with its applications covering areas such as learning (e.g. Atkinson & Crothers, 1964) and decision making (e.g. Rapoport & Chammah, 1965).

A saccade can be described as a movement from fixation (x_1, y_1) to fixation (x_2, y_2). The Euclidian distance between these two fixations is the saccade amplitude. When the difference ($x_2 - x_1$) is not zero, the angle subtended by the saccade would be z degrees, where

$$z = \arc \tan \frac{y_2 - y_1}{x_2 - x_1} \qquad (3.1)$$

When the difference $x_2 - x_1 = 0$, then the angle z may be one of three possibilities: 0 degrees (when $y_1 = y_2$—although not actually a saccade as no movement is involved); 90 degrees (when $y_2 > y_1$); or 270 degrees (when $y_2 < y_1$).

In this method, each saccade has been classified according to its angle, $z(0 \leq z < 360)$, as shown in Table 3.6 and Figure 3.12. One can then use the eight categories *(N, NE, E, SE, S, SW, W* and *NW)* to assign each saccade according to its direction.

Those falling within *SE, SW, NW,* and *NE* categories are termed *diagonal* saccades. It should be noted that the term 'diagonal' does not strictly relate to every saccade which is oblique or not moving at a 90° or multiples thereof. Thus, from Figure 3.13, saccades from point X (centre) to any point traversing a shaded sector would be considered diagonal. The saccade direction change A-X-B, representing a typical return sweep as seen in reading patterns, although obviously not horizontal nor orthogonal, would *not* be termed diagonal.

z (angle in degrees)	classification
$z \leq 22.5$ or $z > 337.5$	*E*
$292.5 < z \leq 337.5$	*SE*
$247.5 < z \leq 292.5$	*S*
$202.5 < z \leq 247.5$	*SW*
$157.5 < z \leq 202.5$	*W*
$112.5 < z \leq 157.5$	*NW*
$67.5 < z \leq 112.5$	*N*
$22.5 < z \leq 67.5$	*NE*

Table 3.6: Criteria for classifying saccade directions.

The *probability vector* has eight elements; its first element being the proportion of saccades $E, p(E)$. The proportion of saccades SE will be the second element, *p(SE);* and so on. The probability vector thus informs us about the frequency of each saccade type.

The *transition matrix* has 64 (8×8) elements. If the number of saccades registered is C, then the number of consecutive pairs of saccades is $C - 1$. There are $8 \times 8 = 64$ different consecutive pairs of saccades *(E, E; E, SE; E, S;... E, NE;...* and *NE, NE).* We can let $C(E,\ E)$ be the number of consecutive pairs of saccades *(E, E); C(E, SE)* be the number of pairs *(E, SE)*; etc. We can also let $C(E) = C(E, E) + C(E, SE) + C(E, S) + \ldots C(E, NE)$; $C(SE) = C(SE, E) + \ldots C(SE, NE)$; etc. The first element in the transition matrix, $p(E|E)$, is the number of pairs (E, E) divided by the number of pairs with E as the first saccade of the pair

$$p(E|E) = \frac{C(E,\ E)}{C(E)} \tag{3.2}$$

The second element,

$$p(SE|E) = \frac{C(E,\ SE)}{C(E)} \tag{3.3}$$

is the number of pairs $(E,\ SE)$ divided by the number of pairs with E as the first saccade of the pair, etc. The 64th element,

$$p(NE|NE) = \frac{C(NE,\ NE)}{C(NE)} \tag{3.4}$$

is the number of pairs *(NE, NE)* divided by the number of pairs with *NE* as the first saccade of the pair. Therefore, each row of the transition matrix informs us about the frequency of each *subsequent* saccade type, when the *previous* saccade is of a particular type.

From eye movement recordings it is possible to calculate the proportion of each saccade type (*probability vector*) and the proportion of a particular saccade type being followed by a saccade type *N*, saccade type *NE*, ... type *NW*; i.e. *the transition matrix.*

Transition matrices, as defined, possess a particular problem. When, for instance, C(E) is small, all the eight elements of the first row of the transition matrix have a high standard error and this first row can give misleading information. To solve this problem, a reduced transition matrix can be used in place of the full 8×8 matrix. Two considerations must be borne in mind. Firstly, each reduced matrix should contain a high proportion of the total pairs of successive saccades. Thus, there is no substantial wastage of information and the reduced matrix contains almost the same information as the full one. Secondly, each element of the reduced matrix should have a reasonable standard error because the reduced matrices are constructed by selecting only the more common saccade directions.

From the probability vector and the transition matrix, it is then possible to show the characteristics of the search strategy. Some information contained in the probability vector can be expressed as the proportion of diagonal saccades (those categorised as *(NE, SE, SW* and *NE)*. If the visual task conveys information in rows and/or columns, a reading-type search would be expected to contain only a few diagonal saccades. Conversely, a random search is expected to provide a larger number of diagonal saccades.

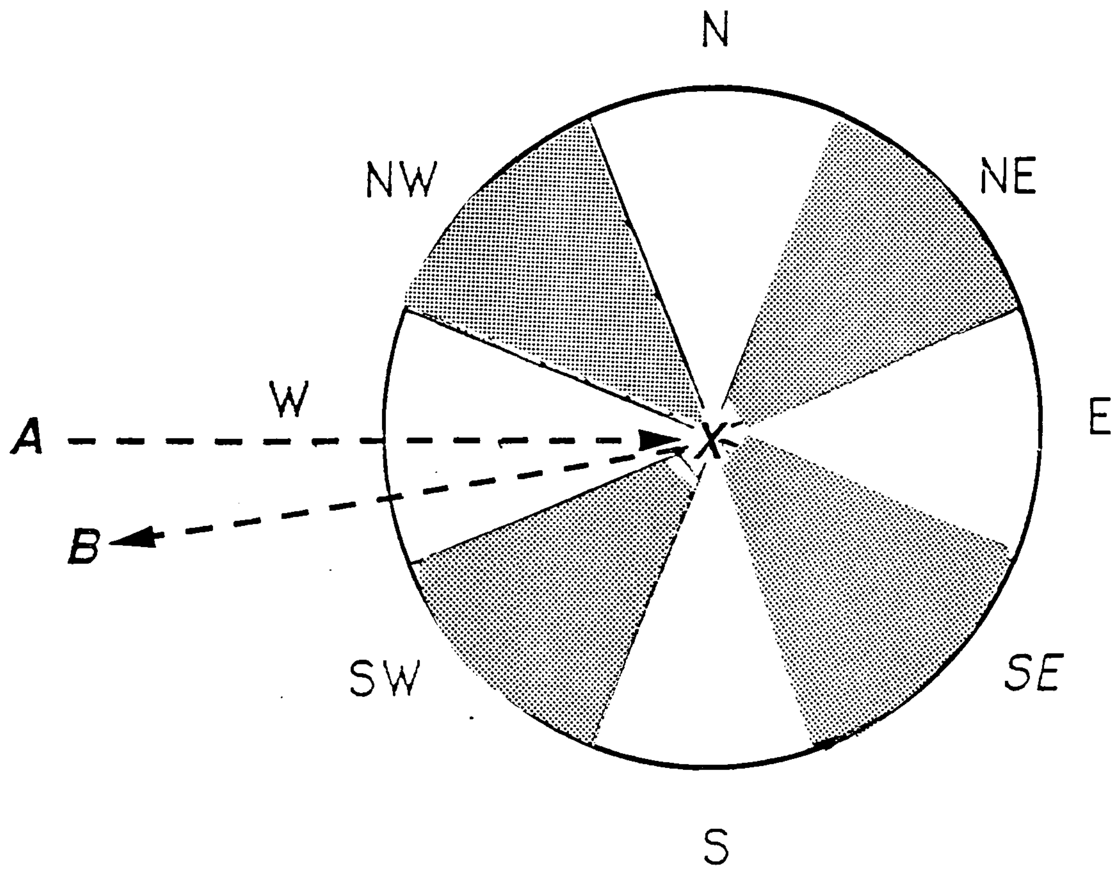


Figure 3.13: A depiction of saccades classified as diagonal.

It is possible to obtain useful information from the transition matrix. By inspecting rows, one can discover some properties of the search. For instance, a reading search would provide large figures for some values (e.g., P(E—E). A modified reading search would be expected to provide similar figures for some symmetrical elements (e.g. P(E—W) and P(W—E)).

In all three transition matrix analyses described within this thesis, the visual material was arranged in such a way as to require only certain types of saccades. In particular, no diagonal saccades are actually required for the tasks involved. Therefore, it was decided to determine if the proportion of diagonal saccades could convey information regarding the possible asystematicities associated with the experimental conditions. As might be expected, the incidence of avoidable (i.e. diagonal) saccades is low but differ in significant ways between the conditions. For the transition matrix analysis, because of the low number of diagonal saccades, only horizontal and vertical saccades were used in the analysis. In this study, the main purpose of the probability vector and transition matrices analyses was to test if, as expected, modifying search instructions would change the proportion of diagonal saccades.

3.9 METHOD

Three subjects (none of whom had participated in Expt. 1) were required to exercise a strictly systematic (reading) strategy on five slides only and then to employ a free search strategy on a further five slides only. Viewing distance was 80 mm, and projection distance was 150 mm. With a variable focus (70 mm–120 mm) lens, it was possible to obtain projected images corresponding to those shown in Experiment 1. Luminance and illumination levels were as for Experiment 1.

Horizontal and vertical eye movement data were recorded using the programme 'SAMSAMP' with a CED 'Alpha' computer. Sampling rate was 20 ms. The SAMSAMP programme transformed the raw eye movement data to give, for each fixation, values of fixation duration (in ms) and fixation position (matched to the rows and columns of the display array).

Files were thus obtained containing lengthy columns of fixation durations (in ms) and a similar column of movement amplitudes (calibrated at per character unit). Data were then statistically analysed using the $SPSS_X$ package. For saccades, all negative (and zero) amplitude lines were edited out, whilst regression data were obtained by editing out all lines containing positive data and large negative values which clearly were return sweeps.

3.10 RESULTS

3.10.1 Traditional eye movement parameters

Table 3.7 shows mean data derived from subjects on the measures of fixation duration, saccade and regression amplitudes (individual subjects' data in Appendix 3.2).

One-way, repeated measures ANOVAs were conducted on this data. Concerning the systematic strategy results, there were no significant differences between arrays for fixations ($F(4,8) = 1.277$), saccades ($F(4,8) = 1.847$), or regressions ($F(4,8) = 1.719$). With the free search results, no significant differences were again found for fixations ($F(4,8) < 1$), saccades ($F(4,8) = 3.635$), or for regressions ($F(4,8) < 1$). Figure 3.14 illustrates the clear trend of decreasing saccade amplitudes from horizontally elongated to vertical elongated arrays for both systematic (reading) and free search strategies.

Array HxW	Fix. dur. (ms)	Sac. dis. (spaces)	Regressions (spaces)
SYSTEMATIC SEARCH			
25x100	303.0 (33.7)	10.2 (4.3)	4.4 (0.6)
30x83	281.3 (20.0)	9.8 (3.0)	4.1 (0.8)
50x50	281.0 (10.0)	8.6 (0.1)	4.4 (1.1)
83x30	276.7 (3.8)	8.3 (1.2)	5.0 (1.3)
100x25	296.3 (15.0)	6.4 (0.6)	5.2 (1.5)
FREE SEARCH			
25x100	301.0 (71.3)	12.7 (5.3)	4.4 (1.9)
30x83	280.0 (20.3)	9.0 (3.5)	4.2 (1.1)
50x50	284.6 (33.6)	9.1 (2.1)	5.4 (0.8)
83x30	277.3 (34.2)	6.6 (2.2)	5.8 (0.5)
100x25	318.0 (12.5)	5.8 (1.9)	5.8 (2.4)

Table 3.7: Mean data from subjects showing mean fixation duration, saccade distance (character spaces travelled; horizontal component only), and regression distance (standard deviations in parentheses).

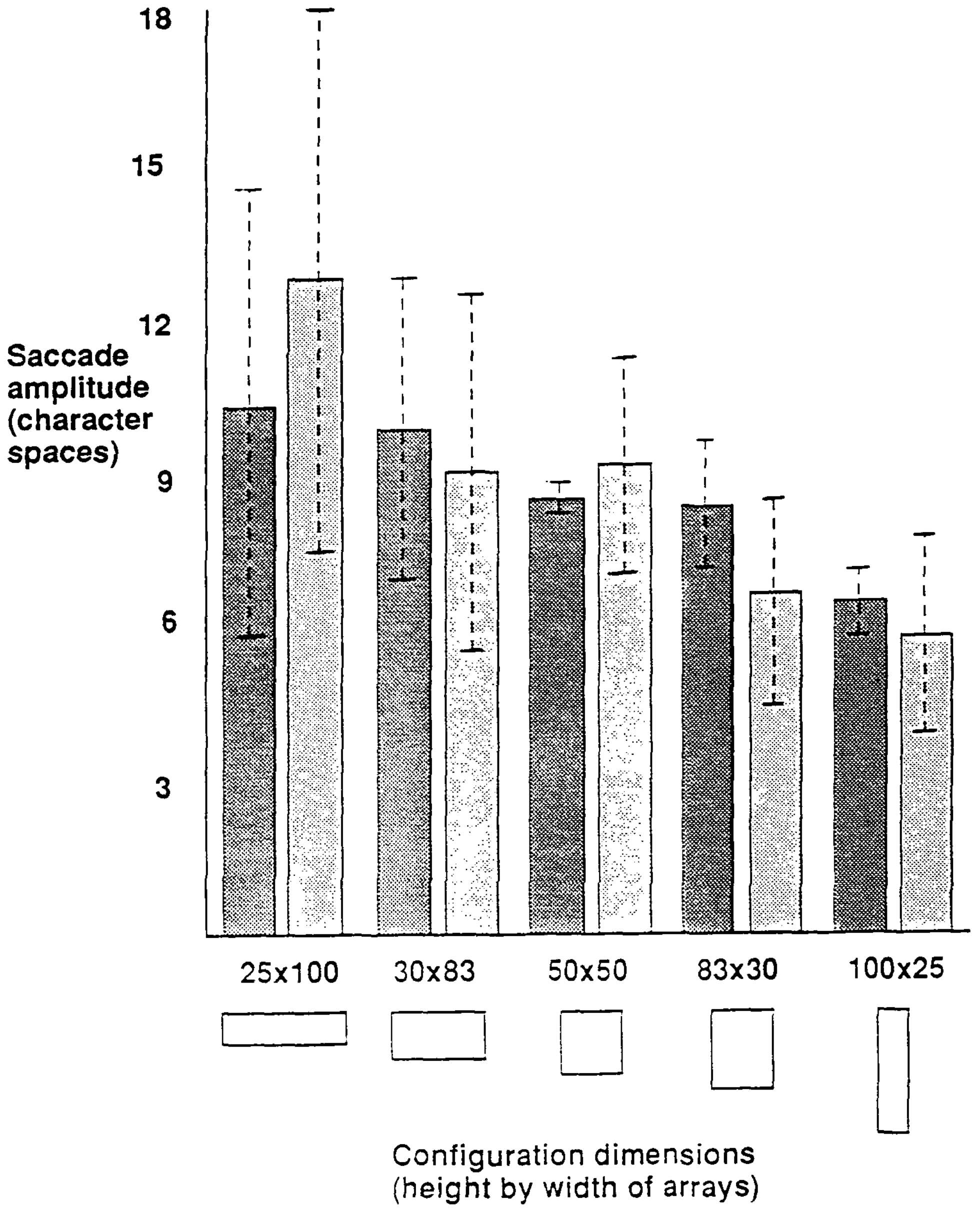


Figure 3.14: Histograms of mean saccade amplitudes over the five configurations on a systematic search (dark bars) and a free search (light bars). Vertical lines denote standard deviations.

Figure 3.15 shows the equivalent pattern for regression amplitudes for the two types of search strategy over the five configurations.

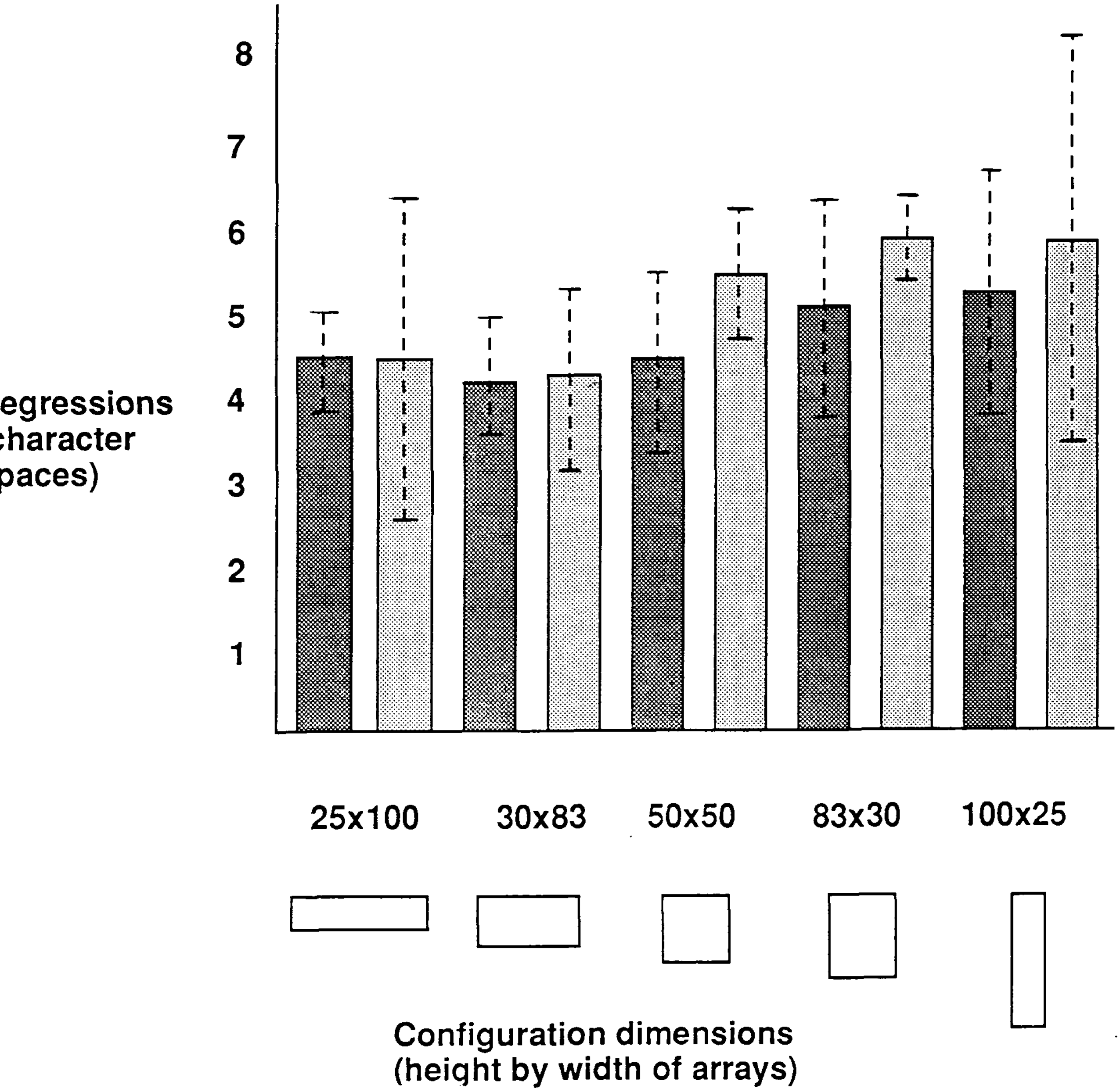


Figure 3.15: Histograms of regression amplitudes over the five configurations on a systematic search (dark bars) and a free search (light bars). Vertical lines denote standard deviations.

3.10.2 Diagonal saccade proportion

Concerning diagonal saccade proportion, saccade proportions may be obtained from the probability vector of each observer by summating the proportion of SE, SW, NW and NE saccades (see Table 3.8). Here, the interest was in differences between extreme horizontally biased arrays and extreme vertically biased arrays. Therefore, in order to simplify the analysis, only the 25×100 and 100×25 arrays were included.

	Free search		Systematic search	
	25x100	100x25	25x100	100x25
AH	0.171	0.263	0.000	0.038
JF	0.102	0.200	0.042	0.163
MS	0.077	0.151	0.012	0.015
$\bar{X}$	0.117	0.203	0.016	0.072

Table 3.8: Proportion of diagonal saccades in two search types and two extreme arrays.

3.10.3 Transition matrices

Concerning transition matrices, due to the fact that saccades E and W were by far the most common, only these are considered within Table 3.9. In the free search condition, the 2×2 matrices contain 68 percent of the total number of pairs of successive saccades (considering the mean across subjects and shapes). This figure increases to 89 percent in the reading situation. One transition matrix was omitted as the particular subject's data contained only a very small number of saccades in this condition. In each of the eleven cells of this 2×2, the first element is

$$p(E|E) = \frac{C(E,E)}{C(E,E) + A(E,W)} \tag{3.5}$$

the second is

$$p(W|E) = \frac{C(E,W)}{C(E,E) + C(E,W)} \tag{3.6}$$

etc.

Each 2×2 transition matrix contains four elements. Two relate to the entries $p(E|E)$ and $p(W|W)$, and do *not* imply direction change, whilst the other two ($p(W|E)$ and $p(E|W)$) *do* imply change in direction.

			Free search		Systematic search	
Sub.	Shape		E	W	E	W
AH	H.S.	E	0.70	0.30	0.77	0.23
		W	0.44	0.56	0.45	0.55
	V.S.	E	*	*	0.48	0.52
		W	*	*	0.58	0.42
JF	H.S.	E	0.72	0.28	0.73	0.27
		W	0.62	0.38	0.66	0.34
	V.S.	E	0.54	0.46	0.44	0.56
		W	0.62	0.38	0.68	0.32
MS	H.S.	E	0.77	0.23	0.89	0.11
		W	0.28	0.72	0.64	0.36
	V.S.	E	0.62	0.38	0.66	0.34
		W	0.43	0.57	0.73	0.27

Table 3.9: 2 × 2 transition matrices of horizontal and vertical shapes (E = saccades to East; W = saccades to West; * = insufficient data for analysis). See §3.8.1 for a description.

3.11 DISCUSSION

3.11.1 Traditional eye movement parameters

The lack of significant results within the ANOVAs must be taken in the light of the small number of subjects involved. It is, of course, particularly difficult for results to attain statistical significance with a sample size of three. However, recruiting volunteers for scleral search coil studies is notoriously difficult. In this case, greater weight may be attached to trends and percentage differences. There is, for instance, a linear trend for saccade amplitudes within both free and systematic searches (see Fig. 3.14). Mean saccade amplitudes diminished from 12.2 (sd = 4.3) character spaces travelled in the horizontally elongated condition to 6.4 (sd = 0.6) character spaces in the vertically extended array; a percentage decrease of 37 percent. There was an even greater decrease in the free search condition with saccade amplitudes diminishing from 12.7 ($sd = 5.3$) character spaces within the 25×100 array to 5.8 character spaces in the 100×25 array; a difference of 54 percent. As greater saccade amplitudes may be taken as representing greater efficiency of search (all other factors, including accuracy, being equal), then these results clearly show substantially greater efficiency of search within the horizontally elongated arrays.

Where efficiency is represented by *greater saccade* amplitude, there might also be a converse *shortness of regressions* expected as the realisation that some characters were not fully perceived initially would be apparent more immediately, and hence the shorter amplitude of regressions to rectify this perceptual doubt. This pattern was also found. Figure 3.15 shows results from the analysis of regressions (in terms of character spaces travelled) for the two types of search. There was an 18 percent increase in regression amplitudes from the horizontally elongated array ($\bar{X} = 4.4; sd = 0.6$) to the vertically elongated array ($\bar{X}; sd = 1.5$). Similarly, in the free search, there was a 32 percent increase in regression amplitudes from 4.4 ($sd = 1.9$) in the 25×100 array to 5.8 ($sd = 2.4$) in the 100×25 array.

In other words, within this small sample, saccade distance varied systematically with array configurations; subjects tending to show larger saccades with more horizontally elongated arrays. This suggests that eye movement parameters when viewing material may be influenced by the overall visual layout of the material. As far as can be ascertained, this has not been previously reported. There was also a trend, regardless of whether a systematic or free search strategy was used, of shorter regressions with the horizontally elongated materials. This combined pattern of eye movement efficiency may underlie the shorter search times required to locate target characters on the screen.

Whilst it might be argued that the validity of regressions outside of a reading task may be questionable, saccade and regression amplitudes have been included within the data from this free search task (Table 3.7). Perhaps distances travelled by saccades have little value when subjects may be searching (perhaps even diagonally) in an entirely random manner. However, subjects reported that, in general, they gave the area a quick initial random scan and then settled into a snaking strategy. Therefore, although there was certainly a greater proportion of diagonal saccades in the free search condition (see §3.10.2), it is likely that these figures do have some validity as far as the traditional eye movement measures are concerned for the free search.

The only other known findings in any way similar to these results appear to be those of Rayner and Fisher (1987). Although examining string length of random characters rather than array shape, they found that arrays consisting of two letters per string were searched more slowly than other sized arrays. Most of the effect was due to the fact that subjects made shorter saccades in this condition. With respect to the size of the perceptual span, Rayner and Fisher also note that the size of the span in their search task was very similar to the size of the perceptual span when subjects are reading rather than searching for a letter.

Thus, the available eye movement data do seem to support this visual lobe argument. A further finding from this work is the evidence for greater efficiency of visual search due to longer saccade amplitude on more horizontally biased arrays.

3.11.2 Diagonal saccade proportion

Firstly, as expected, the proportion of diagonal saccades is greater in free search than in systematic search. This holds also in each of six possible comparisons of three subjects and two shapes. In this experiment the array where the target needed to be searched for was made from different numbers of rows. Thus, the observers, when asked to use a systematic strategy, produced less diagonal saccades than when instructed to use any strategy. This result is not surprising. It is suggested that the proportion of diagonal saccades may be regarded as an index of asystematic search when the array to be searched comprises of columns and/or rows.

Secondly, the diagonal saccade proportion is larger in the vertical shape than in the horizontal shape. This is also true in each of six possible comparisons. Therefore, in the vertical shape, the search is less systematic than in the horizontal one.

Thirdly, if the angles of saccades possessed a rectangular distribution (with parameters 0° and 360°), then the proportion of diagonal saccades would be around 0.5. In fact, that figure is 0.46; because the rows by columns matrix on which the calculations are based was not completely square. The results are far removed from this figure and do not correspond with any purely random search strategy incorporating the rectangular distribution supposition.

3.11.3 Transition matrices

In the *reading-type search,* the three observers show the same pattern. The proportions of saccades corresponding to changes in direction are larger in the vertical shape than in the horizontal shape condition. This result is not surprising as the subjects had been employing a reading search. 'Reading' the vertical shape requires more changes in direction than 'reading' the horizontal one.

In the *free search,* there is no common pattern. The matrices seen within subject JF are very similar to this individual's systematic search matrices, although it is of note that horizontal shape matrices are more similar to one another than are the vertical shape matrices. Subject AH's horizontal shape matrix is also similar to the corresponding horizontal shape systematic matrix. The matrices from MA deserve two comments. On the one hand, $p(W|E)$ and $p(E|W)$ are greater in the vertical than in the horizontal shape. On the other hand, in each matrix, $p(E|E)$ is similar to $p(W|W)$ and $p(E|W)$ is also similar to $p(W|E)$, and this did not occur in the systematic search matrices. A possible explanation is that MS (more so than others), employed something akin to a reading strategy during the free search. This would explain the presence of more changes in direction within the vertical shape, although the 'reading' actually consists of 'snaking'. This would also explain the above similarities.

It may be stressed that the important issue here concerns the general pattern (e.g. efficiency) of eye movements, rather than merely the specific location of fixations.

3.12 CONCLUSIONS

A tentative explanation for the performance differences found within the different shapes can be offered based on eye movement analyses. At least three findings have emerged which relate to the difference between horizontal and vertical arrays.

Firstly, longer saccade amplitudes are found within horizontal arrays. It has been shown that different shapes are related to different saccade amplitudes; longer saccades are more associated with horizontally elongated arrays than with vertically biased arrays.

Secondly, there is a more systematic search with horizontal arrays. It has been shown that the proportion of diagonal saccades is related to the array shape; the proportion of these redundant saccades is greater in vertical configurations. This conclusion can be supported by data from the transition matrix analysis of the reading task (Table 3.9) because there is greater similarity between the free and systematic 2×2 transition matrices when both are horizontal than when they are vertical.

Thirdly, there are more changes in direction within the vertical arrays (larger values of $P(E|W)$ and $P(W|E)$) than within the horizontal arrays, as would be expected due to the greater number of rows to be searched within the vertical condition.

These three results can be related to search time differences. The first result can be related to the lobe/perceptual span hypothesis. The second result can be related to a more redundant search within the vertical arrays, making it more possible to search the same position more than once. The third result could also predict more search time in the vertical array as these arrays require more return sweeps which waste time. However, it is unlikely that return sweeps, by themselves, have more than a marginal effect on reading time.

Therefore, both longer saccade amplitudes and a less systematic search can be the explanation of the performance differences attributed to shapes. Explanations have been offered as to how differences in saccade amplitude can account for difference in search times. Another possibility may now be added; namely, the possibility of a more redundant search within vertical shapes due to greater asystematicity. Naturally, the more redundant is the search, the more likely observers are to repeat the visual scan over the same locations, and the greater is the time likely required to successfully locate the target (assuming no significant use of peripheral vision).

From both (a) an analysis of the frequency of saccades redundant for the task involved (diagonal saccades) and from (b) the analysis of dependency between saccades necessary for the task (E and W), it would appear that vertically biased configurations are characterised by a more asystematic search.

EXPERIMENT 3: ARRAY SHAPES AND READING

3.13 INTRODUCTION

Due to the uniqueness of the findings and important implications from Experiment 2, it was decided to obtain eye movement recordings during a simple reading task where the reading materials were arranged in the five array shapes described above.

As was so of Experiment 1, the rationale of this study was based on the classic work on reading and line width which was detailed earlier (see §3.4.2).

3.14 METHOD

Five 'pages' of text were prepared using stories from the well standardised Neale Analysis of Reading Ability (Neale, 1958). These were arranged within the five array shapes described above. Examples of the two extreme shapes are shown in Figures 3.15 and 3.16. It can be seen that these materials were not right justified, as to do so would have resulted in inconsistent inter-letter spacing with confounding effects for eye movement recordings.

(1) It was midnight. A mournful wailing sound echoed through the deserted castle. The girls ceased
exploring abruptly. "Ghosts!" whispered one girl. "Nonsense!" replied the other, but nevertheless
she proceeded cautiously in the direction of the mysterious noise. Gathering courage, and with
mounting curiosity, the girls approached the old kitchen. Then scarcely daring to breathe, they
swung open the door. Their torches searched the darkness and immediately their excitement turned to
pity. Before them, almost exhausted, lay the farmer's dog. He had been imprisoned while hunting for
rats by a gust of wind.
(2) After a brief encounter with the Turks, Lawrence and his Arab force made a mock retreat.
Although out-numbered, Lawrence guessed that surprise tactics might retrieve the campaign.
Accordingly, as his followers withdrew, they concealed themselves in the rocky crevices of a narrow
gorge leading to the city. Meanwhile the women, acquainted with the circumstances, prepared to
defend the city gates. The success of Lawrence's plan depended on whether the Turks would assume
that the Arab retreat was genuine. There was an interval of terrible tension. Then the unsuspecting
Turks stormed in hot pursuit into the pass. At once, concentrated rifle fire swept their column.
The troops fell into a panic, for the confined space permitted no counter attack.
(3) Fascinated by the prospect of recording the spectacle of a long-dormant volcano smouldering
again, the two scientists approached the crater's edge. Intent on their photography, they ignored
an ominous rumbling. In reproof, the subterranean cauldron suddenly exploded violently, ejecting a
great quantity of rocks. Fortunately these fell on the opposite slopes. Greatly alarmed by this
premature eruption, the men hastily began the descent. Instantly a gigantic avalanche of fiery
boulders hurtled around them. Aware that their apparatus hindered progress, they abandoned all
equipment except their precious cameras. Then came an anxious moment. As one man was evading a
flying fragment, he was struck off-balance by a rebounding boulder. A lengthy halt would have been
disasterous. It was, therefore, with immense relief that they discovered his injuries to be
superficial and resumed the fantastic scramble to the safety zone.

Figure 3.15: Example of 25 × 100 arrayed reading material.

(1) Susan hurried to the
starting position for the
relay race. Last year her
team had been disqualified
for not transferring the
baton properly. Now they
were determined to avenge
their defeat. But what was
this? Susan inspected one
shoe. The sole had broken
lose in the obstacle
event. Her heart sank. The
track was unsuitable for
running barefoot. Her
plight, however, had been
observed. "Try mine,"
insisted Philip, a reserve
runner, unfastening his
shoes. Luckily they
fitted, and later, Philip
shared the honours when
his school was awarded the
athletic shield.
(2) Among animals the fox
has no rival for cunning.
Suspicious of man, who is
its only natural enemy, it
will, when pursued,
perform extraordinary
feats, even alighting on
the backs of sheep to
divert its scent trail.
Parent foxes share the
responsibilities of cub-
rearing. Through their
hunting expeditions they
acquire an uncanny
knowledge of their
surroundings, which they
use in an emergency. This
is well illustrated by the
story of a hunted fox
which led its pursuers to
a neglected mine-shaft
enclosed by a circular
hedge. Swiftly it mounted
the barrier. The hounds
followed, only to be
drowned in the accumulated
water fifteen metres
below. The fox, however,
apparently on familiar
territory, skirted the
hedge and subsequently
escaped.
(3) Each April, at the re-
appearance of the cuckoo
in its familiar haunts,
bird-watchers must marvel
at the accurate flights
with which birds span
their seasonal abodes.
What causes these regular
journeys? The theory that
rigorous winters compel
birds to migrate is
insufficient, for many
migrate in summer.
Likewise, it cannot be
argued that the fledglings
imitate the older
generation, for the
offspring generally
migrate alone. The best
explanation suggests that
migration is an inborn
custom, probably
originating in some
ancient era when the
flights were necessary for
survival. Most species
favour particular routes.
Thus, on one occassion
when some storks from east
Germany were captured and
released among storks from
west Germany, they did not
accompany their relatives
along the western
migration route. Instead,
with unerring instinct,
they re-discovered the
traditional south-
easterly path of their
eastern ancestors.

Figure 3.16: Example of 100×25 arrayed reading material.

Five subjects participated.

Stimulus dimensions and experimental conditions were as described in Experiment 2.

The scleral coil technique was again used for this study.

3.15 RESULTS

3.15.1 Search paths

Figures 3.17 and 3.18 depict typical eye movement patterns obtained during reading the five passages of differing shapes. In both cases the straighter line shows the vertical movement and the step-like channel shows horizontal movement.

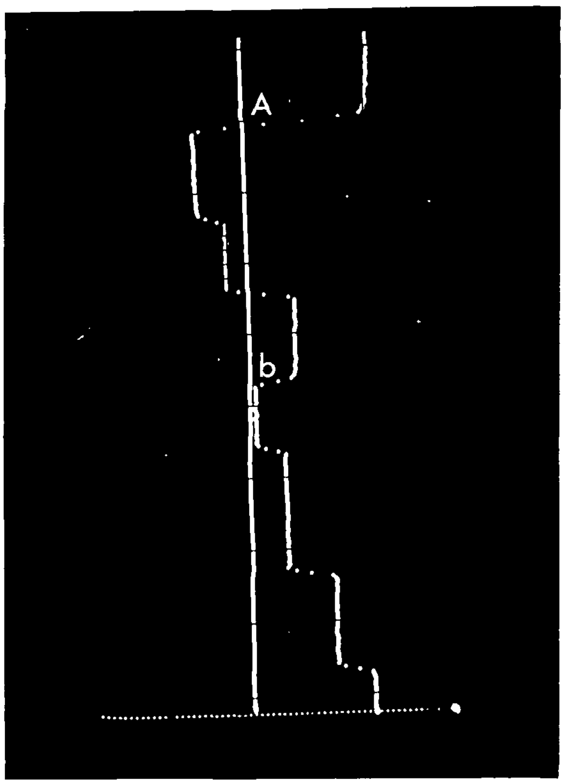


Figure 3.17: A section of eye movement recording from the 25 × 100 arrayed reading material.

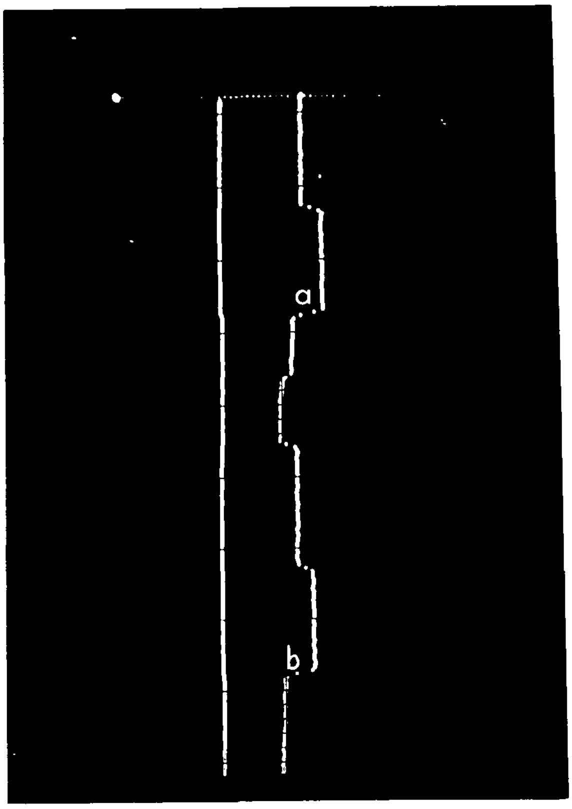


Figure 3.18: A section of eye movement recording from the 100 × 25 arrayed reading material.

Table 3.10 shows mean data derived from five subjects on the measures of fixation duration, saccade and regression amplitudes.

Array HxW	Fix. dur. (ms)	Saccades (spaces)	Regressions (spaces)
25x100	282.7 (58.6)	7.5 (1.3)	6.7 (2.7)
30x83	288.5 (61.1)	9.0 (2.8)	6.9 (3.9)
50x50	291.8 (71.9)	8.4 (1.3)	8.3 (3.0)
83x30	304.2 (55.1)	8.9 (0.8)	7.4 (2.4)
100x25	297.1 (48.0)	8.7 (2.7)	7.1 (1.3)

Table 3.10: Mean data from the five subjects showing mean fixation duration, saccade distance (character spaces travelled), and regression distance (standard deviations in parentheses).

One-way, repeated measures ANOVAs were conducted on these data. There were no significant differences between arrays for fixations ($F(4,16) < 1$), saccades ($F(4,16) < 1$), or regressions ($F(4,16) < 1$). These findings neither support nor refute the lobe distribution hypothesis.

A close inspection of the data revealed nothing of a more positive nature. There was only a five percent increase in fixation duration from the horizontally biased to the vertically biased array, a 14 percent *increase* in saccade amplitude, and a 6 percent increase in regression amplitude. Concerning saccade amplitude, it should be noted that in line with the outlined rationale, the opposite pattern to that found would have depicted greater eye movement efficiency with the horizontally biased arrays.

3.16 DISCUSSION AND CONCLUSIONS

Firstly, Figures 3.17 and 3.18 both show vertical eye movements (the straighter line), horizontal eye movements (the step-like channel), return sweeps (marked A, and regressions (marked B). The two patterns, however, are clearly different; largely due to the shorter return sweep to return sweep (line length; B—C) of the 100×25 array (Fig. 3.18).

On a general level, the studies described in this chapter of search with arrays of different shapes demonstrate a superiority for horizontally elongated arrays over vertically elongated ones. Eye movement patterns suggest two possible reasons for this. Firstly, horizontal arrays may be better matched to the natural visual lobe for text-like material. Secondly, the array shape may affect the size of scanning saccades, leading to more efficient scanning behaviour.

As noted earlier, Bloomfield (1970) remarked that: 'Because reading habits are so stereotyped, observers probably found the vertical letter set more difficult to scan whatever the display shape. Also, the vertical displays probably cannot be scanned very efficiently with the horizontal letters. This argument leads to the idea that display shape may only have an effect with material, such as the alphabetic symbols used here, that has associated with it stereotyped fixation patterns' (p. 118). 'One might conclude that display shape may have a consistent effect, if the display contains material to which the observer has developed stereotyped responses. The majority of search tasks probably do not come into this category. In them there may be display shape effects though it is likely that they are probably much less consistent' (p. 121). It is now felt that these statements are in need of some correction and qualification. It is suggested, on the basis of the experimental evidence provided here and on what is known of the general nature of stereotyped tasks, that verbal (alphanumeric) characters are susceptible to the shape effect. However, text (reading material) appears to be relatively immune, most likely due to the stereotyped nature of the visual and cognitive processes involved. In other words, the absence of susceptibility to the shape effect is likely due to higher level factors related to linguistic processing and cognitive load as mentioned earlier (§3.4.1) and as previously noted by Lévy-Schoen *et al.* (1984) and Jacobs (1986). Bloomfield seemed to be equating the three-letter combinations as used in his relevant work (and which were sometimes vertically arranged) with reading material. It is not surprising that he found it 'difficult to draw any conclusions' (p. 111) about some of his own results.

One implication from these studies relates to menu information (at least iconic menus and possibly also to brief verbal desriptors). In one of the numerous studies (see Parkinson *et al.*, 1988) of breadth versus depth of menu structure, it has been found that increasing depth (up to a single screen of 64 items rather than, say, four options with three frames) is an efficient mechanism for reducing menu search time (Snowberry *et al.*, 1983). In fact, number of responses seems to be the important factor. This suggests that the array configuration of large menu screens would be an important application area for this work.

It would be interesting to test whether iconic materials more generally are prone to the shape effect. Our evidence would suggest that a search performance superiority would also be found within horizontally biased configurations for this form of non-reading material. Once again, there are obvious implications for the spatial layout of iconic information within VDUs, and in particular, multiple-windowed VDUs.

Chapter Four

Optimal Presentation of Status Information

4.1 OVERVIEW

Visual search is required whenever a target or material is not immediately apparent. In the previous chapter, the optimisation of visual search was examined in relation to screen configurations. Attention is now turned to the effects of reducing visual search within a typical VDU-based task. Specifically, the possible benefits were investigated of displaying necessary information at the 'point of regard'; i.e. reducing the need for a visual search of relevant material. This chapter describes a study utilising an automated human-computer interface monitoring system to record the efficacy of two methods of displaying state information (insert versus type-over mode) within a text-editing system. Information presented only at the cursor position resulted in faster overall performance time than that when information was presented only on a status line. The results are discussed in terms of visual processing and cognitive factors. A second study integrated eye movement monitoring records with those obtained with a protocol analyser.

EXPERIMENT 4: CURSOR DISPLAYED INFORMATION

4.2 INTRODUCTION

Interactive computing with VDUs is rapidly progressing from a situation in which the VDU was employed simply to present sequential text to a situation where the full potential of the technology may be exploited so as to distribute information at a variety of different screen locations. Findlay *et al.* (1988) recently suggested ways in which psychological knowledge of the visual, attentional and memory systems can be directed to this application, and presented a pilot study (see §4.3) showing how the concepts may be amenable to experimental investigation.

Most single tasks can be subdivided into components. It is recognised in the study of *task analysis* that a variety of different hierarchical levels may be considered. Findlay *et al.*'s analysis considered a low level in the hierarchy; the fine structure of user actions. It was also noted that some of the concepts involved would seem to be useful for consideration of higher levels too.

The seminal (yet rather simplistic) keystroke analysis of Card *et al.* (1983) showed that certain activities, most notably text editing, could be rigorously analysed into a sequence of elementary operations. A limitation of their approach is the concentration on a single goal-oriented processing stream. A feature of human cognitive activity which becomes of increasing importance when more complex displays are considered is that much processing occurs outside the main conscious processing stream. The psychology of attention offers a model of human operation in which activity (thought or action) in one principal processing stream may be accompanied by activity in several subsidiary processing streams. Attention is regarded as a flexible resource with tasks which are overlearned and/or of minimal information content proceeding in an automatic manner as the individual simultaneously engages in a principal task (e.g. Norman & Bobrow, 1975).

With vision, for instance, our eyes take in information from a wide area of the visual field. At any instant, the gaze is directed to one particular location. Normally, it is the visual material to which the gaze is directed which receives focal attention. However, visual processing outside the active region is still occurring. As an example, any substantial change will be capable of eliciting an orienting response wheresoever it may occur in the peripheral visual field.

Investigations of reading processes (e.g. Rayner, 1983) demonstrate how parallel processing is employed dynamically. Detailed textual information is absorbed from a fairly small region where gaze is directed, but less detailed information (e.g. word shape, word boundaries) is being simultaneously assimilated from more distant regions to assist the eye guidance and provide some preliminaries to the detailed analysis. It has been suggested that certain discriminations in parafoveal and peripheral vision can be made at *no cost* to the attentional resources required for central processing, whereas other discriminations can be made only if some conscious effort directs attention to the peripheral location (Treisman, 1985). Normally, direction of attention to a peripheral location would involve direction of an eye movement to that location although the two operations are partly dissociable (Shepherd *et al.*, 1986).

For the user, one particular stream will be occupying the current focus of consciousness and an important consideration is how this focus shifts from one stream of activity to another. Such multiple streams may well be occurring both on the human and on the computer side of the interaction. As Findlay *et al.* (1988) note, surprisingly little appears to be understood of this aspect of cognitive functioning. Miyata and Norman (1986) consider this problem and use the computer analogy of 'interrupt handling'.

Within this analogy, conscious human activity can be viewed as consisting of bouts of 'processing' which are terminated by an 'interrupt'. Interrupts may arise *externally*, through the senses; e.g. a telephone call or the arrival of a visitor. However, the interrupt idea seems to have face validity as a description of *internal* causes of processing terminations (see §4.2.1). Suddenly remembering a task which needs urgent attention would be one example.

Situations arise in which interrupts are deliberately preset as *reminders* in order to ensure that a particular task is accomplished. The issue of reminders can link the abstract study of attention and processing streams with practical design considerations since the provision of reminders is rendered feasible by a multi-element display.

4.2.1 Internal and external memory

The topic of reminders illustrates the close connection between the topics of attention and memory. A reminder is essentially a memory aid and such phenomena are used widely; e.g. a shopping list or an examination-aid mnemonic. External aids to memory are often employed when other, intervening, cognitive events might interfere with the processes of learning and recall, when accuracy is at a premium, and when memory load is to be minimised to facilitate the allocation of attention to other activities. In general, it would appear that individuals prefer to employ external aids to memory rather than rely on their own internal memory (Intons-Peterson & Fournier, 1986). This suggests that the effort involved in the use of external memory props is less demanding than the cognitive effort necessitated in encoding and retrieving information from internal memory sources.

4.2.2 State monitoring

A common situation in human-computer interaction where the problem of internal memory arises is 'state monitoring'. A characteristic of a complex system is that it can manifest a variety of different 'states', and as systems grow in complexity the number of possible states also increases. When interacting with a complex system, the problem of how to 'keep track' of the current state of the system may be termed the 'state monitoring' problem.

One situation in which state monitoring is necessary occurs with many word processors and text editors. These have available two or more modes that may be used for inserting text. Typically, there is an *insert* mode in which new text may be entered without deleting any existing text, and alternatively there is a *type-over* mode in which new text is typed over existing text. Experience shows that both facilities can be useful at different times and commonly it is possible to switch from one mode to another with a simple command or keypress.

This situation illustrates the choices occurring at the fine structure level relevant to the design of an interactive system. From the user's point of view, we are considering an instance where a processing stream must be interrupted. New information must be obtained about the system state before the next task—the next insertion—is performed. As mentioned above, both psychological theory and design practice suggest that in this particular case external rather than internal memory should be used. The salient question is then: How can this information be best presented on the screen?

Many word processing systems present a reminder about the current state of the system in a 'status-bar' menu, generally at the top or at the bottom of the screen. To use the information, it is necessary to look at it. This involves a routine which takes the following form: make a saccadic eye movement to the appropriate screen location; retrieve the information; find the text position (flashing cursor); and make a saccadic eye movement back to the text position. Findlay *et al.* (1988) used the term 'minimum switch' or 'mini switch' to describe this sequence, and this was contrasted with both more or less extensive interrupt handling operations as follows:

(1) **Normal Switch:** Used to obtain information which is not immediately available on the screen. For example, a sequence of mouse operations could be used to open a window to obtain the information.

(2) **Mini Switch** (as described above): Used to obtain information that is available on the screen but not at the central viewing location. To obtain the information it is necessary to move the eyes to the screen location.

(3) **Micro Switch:** Information could be presented in the visual periphery which has sufficient salience for an overt eye movement not to be required although a covert attentional movement would still be involved. It is probable that this covert movement would entail some attentional cost.

(4) **Information at the Fixation Point:** If a state indicator could be made available close to the fixation point, one would predict that even less attentional capacity might be required for its use.

Analysis of the component activities shows that fewer actions are required as the list is descended. We may expect that the elimination of actions will lead to faster and, perhaps, more easily used systems.

4.3 FINDLAY *ET AL.*'S (1988) PILOT STUDY

A normal text editor was modified to present the state information at the fixation point by modification of the cursor from a solid block in one condition to a pair of lines above and below the character. The aim of the study was to compare performance when state monitoring required a 'mini switch' of a saccadic glance to a peripheral location (the unmodified text editor) with that when information was available at the fixation point (the modified version). The text editor was standard on the University mainframe, and subjects all had some familiarity with the unmodified form of the text editor. They were given some training followed by an editing session in which both indicators were available. They then carried out a second session in which only one of the indicators remained. Results were analysed in terms of changes between Session One and Session Two between the two groups. In the pilot study, it was found that those subjects having information displayed at cursor-plus-status-line followed by cursor-only experienced no difficulty with the change from the normal form of the editor and actually made significantly fewer pauses immediately before a mode change on the second trial/condition than did those having the same former condition but followed by status-line-information-only. This same subject group also demonstrated a greater reduction in time between first and second trials than did those with status line information only on the second task. This was interpreted in terms of attentional interrupts and eye movements necessary to view information only available at the status line. It was suggested that there might be an advantage within a system designed with fewer operations requiring a the switch of attention.

There was, however, a subjective element to the particular measures unavoidable at the point of development at which the data were collected in that a 'pause' was not measured objectively but was taken as a noticeable delay with no cursor movement during fast replay ('equivalent to a pause of about 3 seconds in real time').

4.4 MONITORING EQUIPMENT

In 1984, IBM published a description of the *Playback* system, which was designed to assist in an ergonomic evaluation of the usability of an application system, particularly in the evaluation of word-processing systems (Neal & Simons, 1984).

More recently, Morris *et al.*, (1988) have described the work involved in producing a similar acquisition and analysis tool, involving physical portability, multi-channel recording, synchronised video and sound recording and a synchronised system for acquisition of analysis notes. This HIMS monitor (Human Interface Monitoring System) has been developed at the University of Manchester, Institute of Science and Technology, U.K., by the Department of Computation, in collaboration with ICL and the HUSAT Research Centre, under an Alvey Project Grant from the UK Government. With the HIMS a special purpose workstation is used to capture a detailed record of the interactions of one or more people with a computer system during a complete work session. This information can then be used to evaluate particular input/output devices, software interfaces and total systems. Additionally it may be used to investigate some aspects of human behaviour. The device has, for example, been used for assessing programming plans within the context of skill development in programming (Davies, 1990).

The hardware of the system is based on a VMEbus with a single-board computer supplying the main processing power. This contains a Motorola MC 68010 microprocessor. The processor board is interfaced to an 80-Megabyte hard disc and an IBM PC/AT-compatible floppy drive. A second VMEbus board provides eight RS 232 serial input/output channels. The whole system is normally mounted in a five-position card cage measuring 130 × 320 × 430 cm. The remaining three card positions are available for interfaces specific to a particular use of the system. Two common examples of these are the keyboard and mouse line taps. Fig. 4.1 shows an example of monitoring.

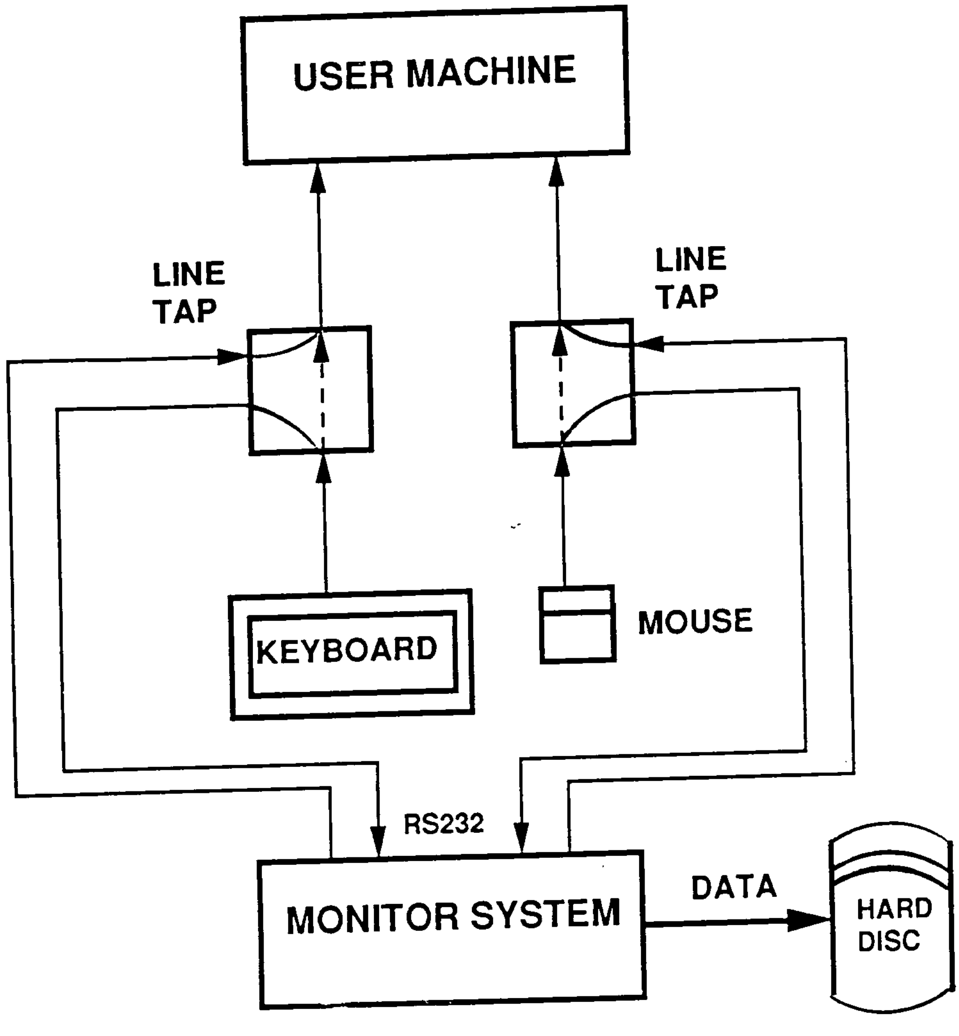


Figure 4.1: A schematic diagram of an example of monitoring with the HIMS. From Morris *et al.* (1988), 'Human-computer interface recording', *The Computer Journal,* **31,** p. 439, Cambridge University Press. Reproduced with permission.

The data record can subsequently either be analysed or used to form a session replay in which the user machine is driven by the recorded data to replicate exactly the original sequence of operation; 'player-piano' style. Due to development in the HIMS system, the data collected were available for automated analysis in greater detail than that obtained in the pilot study. Whereas only overall timing could be automatically recorded for subjects in the pilot study, the following programmable facilities were used in the present study.

(1) **Pauses:** The analysis software could be set to record periods of keyboard inactivity of any length. For comparison with the pilot study, a three second period was chosen.

(2) **Pauses Immediately Before a Mode Change:** As mode changes could automatically be time-stamped to a hundredth of a second (see below), it was again possible to use the automated analysis to determine where pauses were immediately followed by mode changes.

(3) **Changes to Type-over Mode:** The HIMS could be set to record a number of specified keystroke sequences. The sequence 't-h-e' could, for instance, be so 'mapped'. In addition, non-displayed symbols (e.g. the control key) could be recorded by programming their hexadecimal equivalent. Thus, control-A could be mapped equivalent to keystroke hex 01. This was the change to type-over mode keystroke sequence.

(4) **Changes to Insert Mode:** Likewise, the hexadecimal 13 (control-S) was mapped to record incidences of changing to character insert mode.

The present report extends the pilot study using an improved monitoring system, removing all elements of subjectivity from the data collection.

4.5 METHOD

Results from a further 12 subjects were obtained and, where appropriate, combined with the original data. In all, 24 paid subjects (M 11: F 13) with a mean age of 24.0 (sd=4.8) participated.

A screen-editor programme ('SE'), previously used as the Northumbrian Universities Multiple Access Computer (NUMAC) mainframe editor, formed the basis of the software used in this study. Permission was granted to modify this programme for research purposes. In traditional form this provides a negative contrast (grey text on dark blue background with the monitor used) 80 column by 20 rows (excluding status line) format. Three modified forms of this screen-editor were constructed which could be run on an IBM PC. All modified programmes allowed the user access to a restricted command set. Forward and backwards scroll commands, cursor movement commands, character set and line deletion commands were common to each version. These all used the same pair of control keys for switching between 'character insertion' and 'type-over' modes. In the former case, characters could be inserted in the middle of text; in the latter case, characters entered would over-type existing material. It should be noted that the term used for the type-over mode within this particular screen-editor was 'line insertion' (see, e.g., Fig. 4.3). This related to the fact that when the user had reached the end of one line whilst in type-over mode, a new blank line (below) would be created, rather than have the user over-type the next line of text also. In each case the cursor flashed at a standard rate.

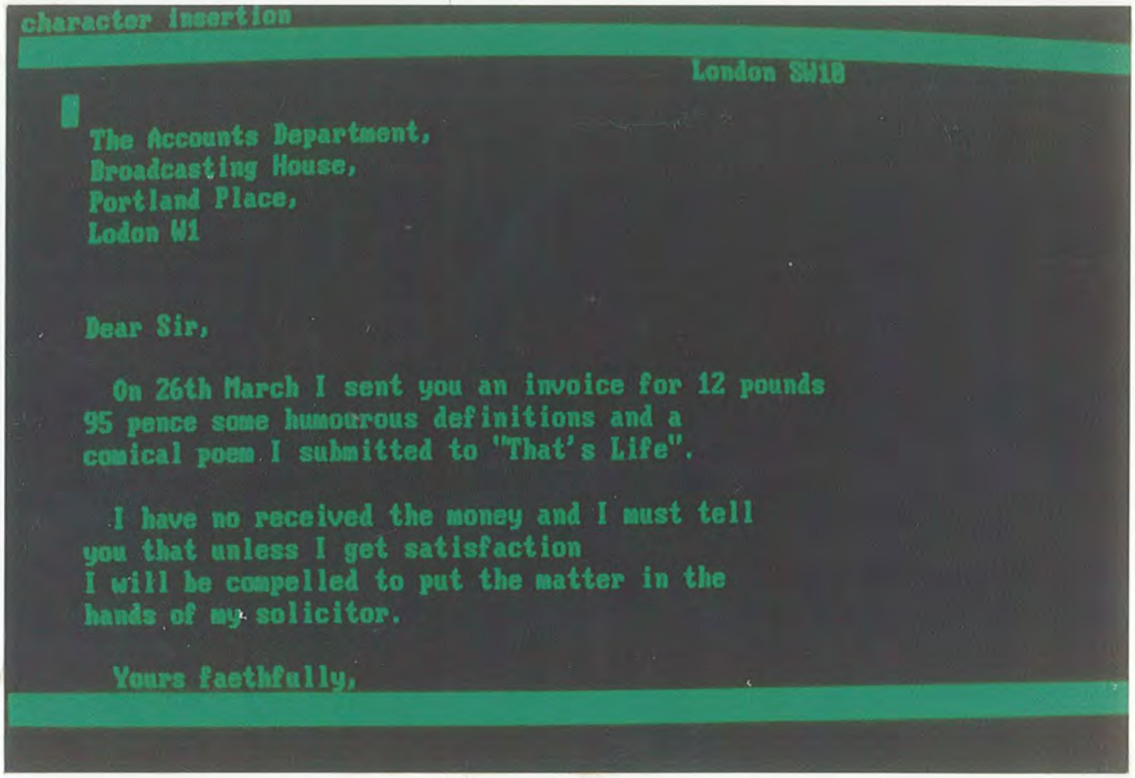


Figure 4.2: An example of a screen display from the Window Only (insert mode) condition.

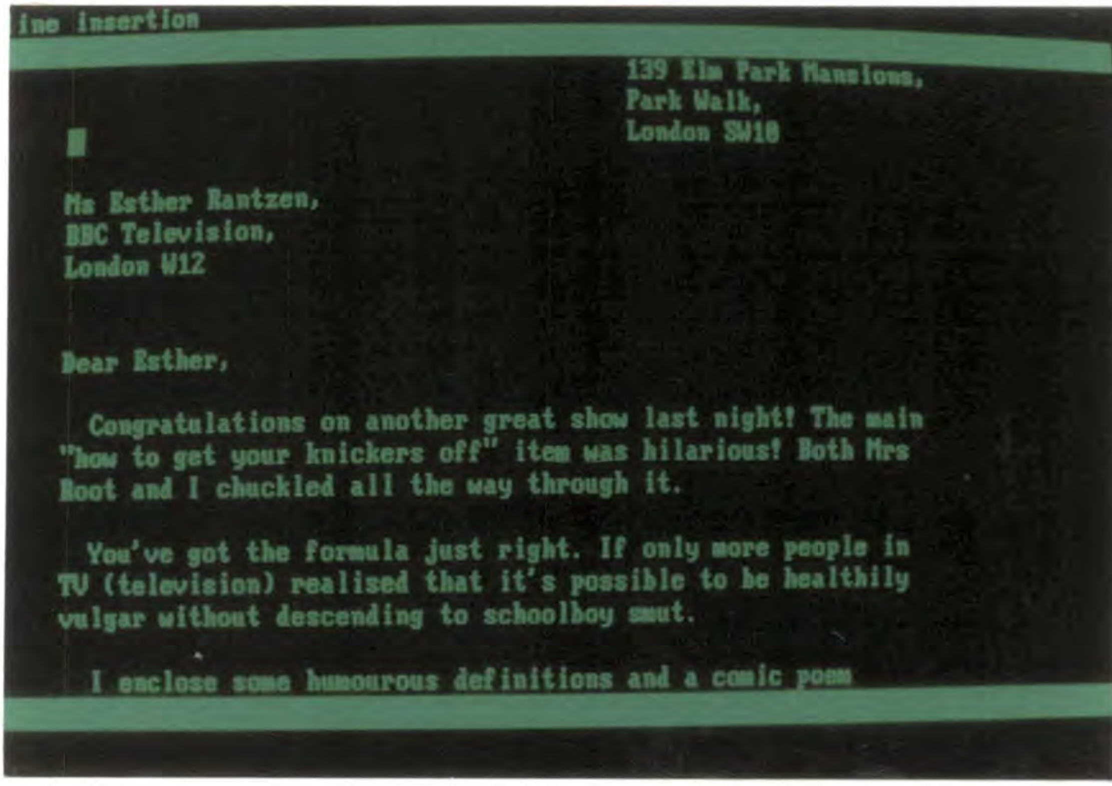


Figure 4.3: An example of a screen display from the Window Only (type-over mode) condition.

State information was conveyed in three different ways as follows:

(1) **Window Only:** A verbal message was presented in a one-line window on the top left of the screen to indicate which of the two insert modes were current (see Figs 4.2 and 4.3). The window measured 0.4 mm by 3.5 mm, equivalent at a viewing distance of 50 cm to an image subtending 0.5° by 4° at the retina.

(2) **Cursor Only:** The mode indication was achieved by a change in cursor character. In type-over mode the cursor was a solid block (■) (see Fig. 4.4). In insert mode, the cursor was changed to a pair of bars (=) (see Fig. 4.5). In each case the cursor flashed at a standard rate .

(3) **Window and Cursor:** Both these indicators were available (see Figs 4.6 and 4.7).

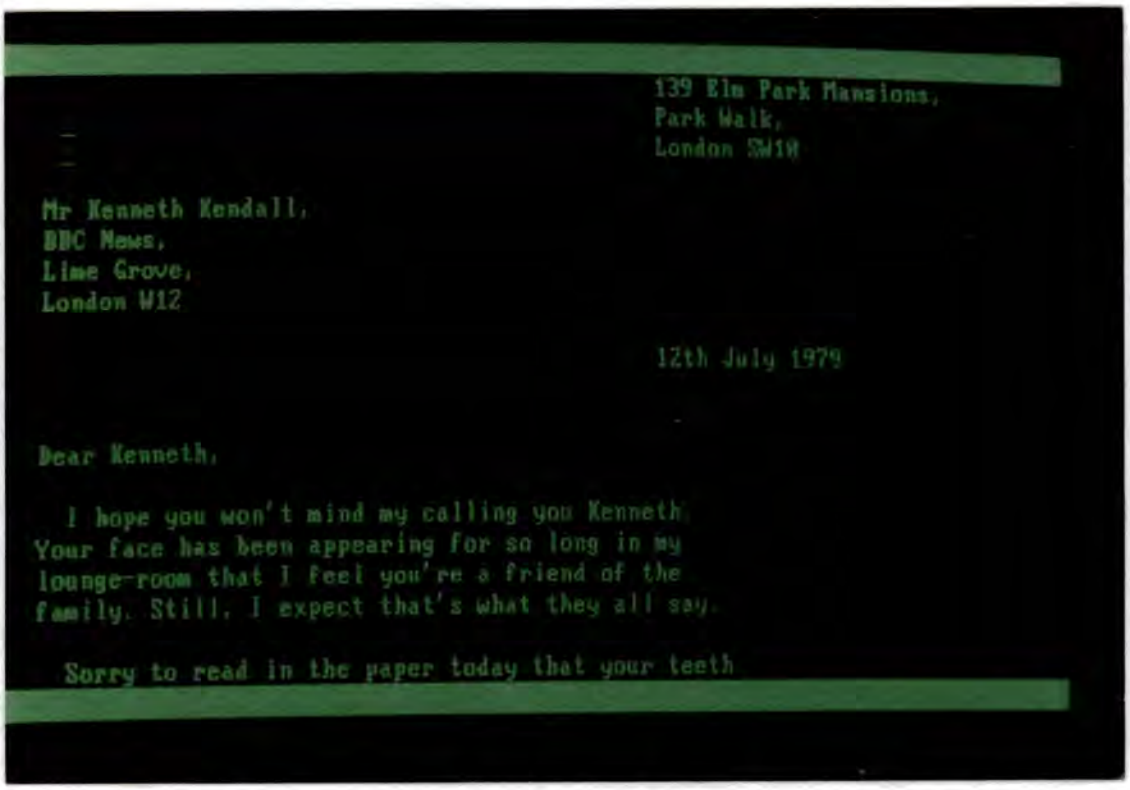


Figure 4.4: An example of a screen display from the Cursor Only (insert mode) Condition.

Two sets (approx. 550 words each) of fictitious letters were compiled. Files containing degraded versions of these texts were prepared for the PC in which a large number of errors were deliberately inserted. These were edited so that their correction would be most simply achieved by frequent alternations between insert and type-over modes. The number of corrections to be made were the same in both texts. Typed A4 copies of the files were made with the errors requiring correction highlighted in yellow (character insertions) and orange (erroneous letters where type-over gave the simplest means of correction). Examples of materials used are included as Appendix 4.1.

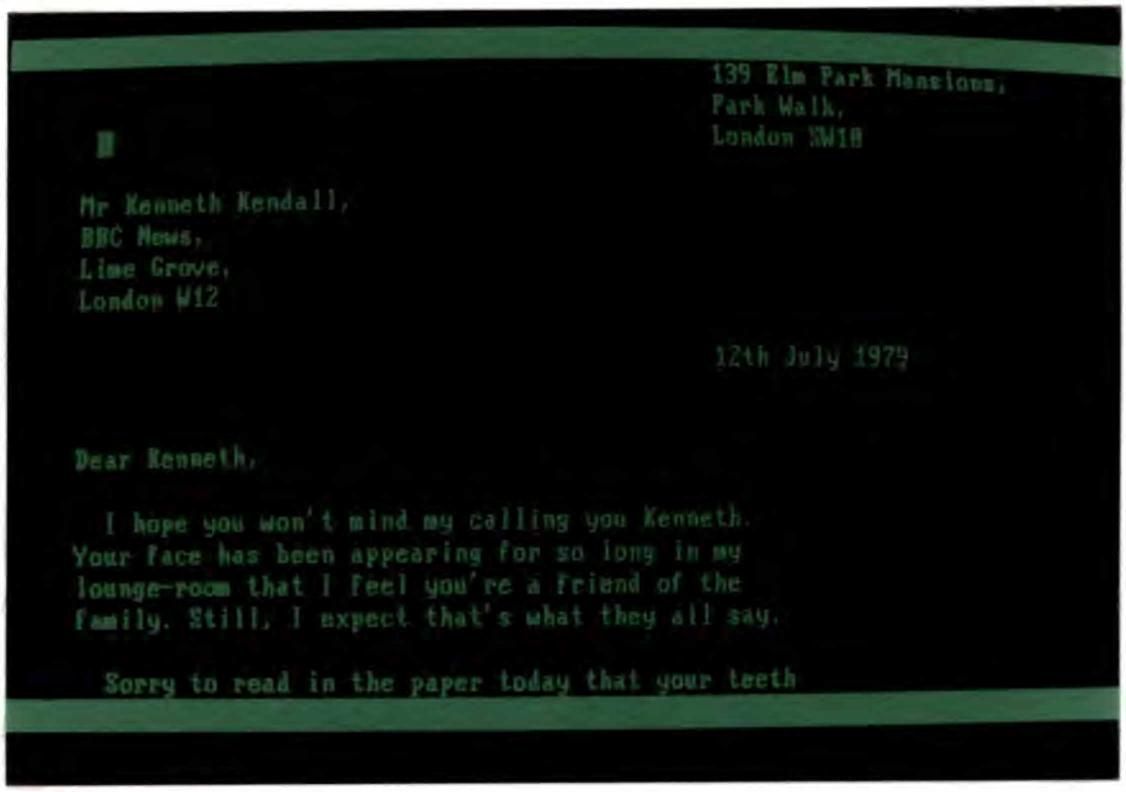


Figure 4.5: An example of a screen display from the Cursor Only (type-over mode) condition.

The following criteria were chosen for assessing performance: (a) number of typing mistakes made; (b) number of mistakes made which remained when the subject stated to the examiner that they had finished (either overlooked errors in the original text or typing errors incurred during the trial and left uncorrected; (c) incidence of 'verifying' pauses of three seconds or more immediately before a mode change (where the subject may well be making a 'mini switch' of attention in order to check from the status bar what mode they are in); (d) other pauses of three seconds or more made during the session; (e) number of attempts to type-over text whilst actually in insert mode, and (f) vice versa; and (g) number of mode changes made (in part a measure of errors made and which required correcting).

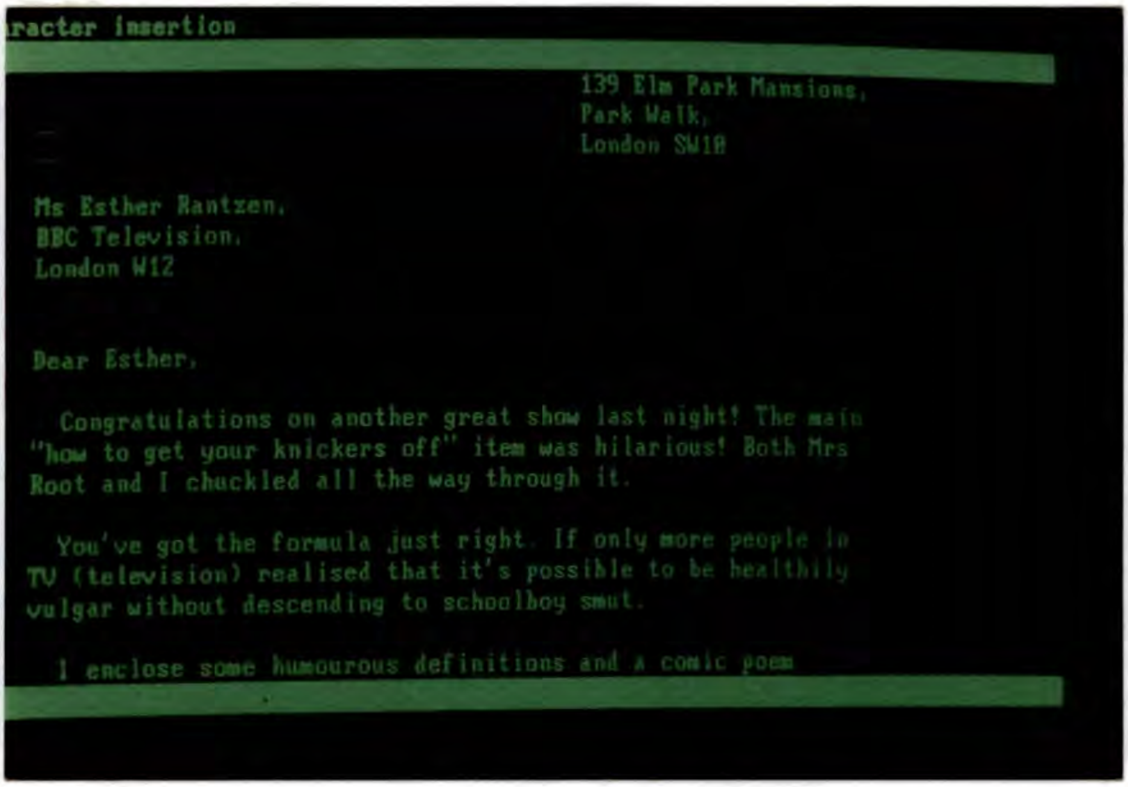


Figure 4.6: An example of a screen display from the Window and Cursor (insert mode) condition.

Subjects were randomly assigned to two groups. Both groups carried out an initial text editing session in the Window and Cursor condition and were then assigned randomly to the Window Only ('Window group') or Cursor Only ('Cursor group') condition for a second session. Instructions to subjects are included as Appendix 4.2.

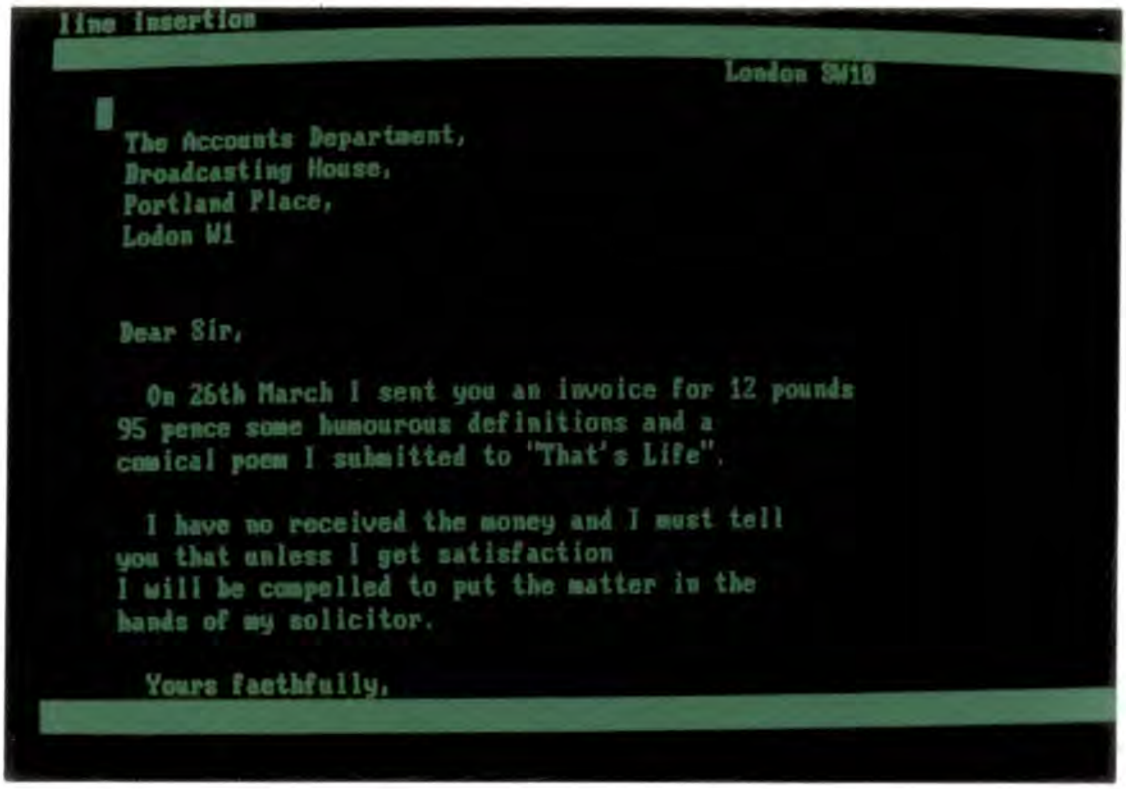


Figure 4.7: An example of a screen display from the Window and Cursor (type-over mode) condition.

4.6 RESULTS

Table 4.1 illustrates results obtained from all 24 subjects on various measures analysed. Figures refer to differences between the first and second sessions. Usually, this is in absolute terms but where more appropriate and possible, percentage differences are specified. In line with results from the pilot study, pauses were compared in terms of those over three seconds both immediately before a mode change and also elsewhere. Pauses immediately before a mode change were also compared in absolute terms.

A particular problem emerges with the analysis of the total time on task. The subjects used possessed different typing skills and also different amounts of experience with the word processing system. In consequence, improvement between session one and session two was expected on the basis of a simple practice effect. Moreover, the extent of improvement, assessed as a measure of time speed-up, was likely to be greater for subjects at the earlier stages of practice since skill learning curves are known to be negatively accelerated (Welford, 1968). This is apparent in the high correlations (0.90 for the Cursor group and 0.58 for the Window group) between the improvement score and the total time for the first task. To allow for this, a percentage measure was used in which the speed-up time was expressed as a percentage of the time taken on the first task.

	Cursor Group				Window Group			
	Task 1	Task 2	Change	% change	Task 1	Task 2	Change	% change
Typing mistakes made before correction	22 (10)	19 (11)	-3 (10)	87 (36)	24 (24)	22 (23)	-2 (8)	84 (49)
Mistakes remaining when 'finished'	6 (6)	6 (4)	0 (4)	158 (131)	8 (11)	10 (17)	2 (7)	96 (59)
Pauses of 3 s or more immediately before mode change	11 (9)	6 (6)	-5 (5)	59 (30)	14 (9)	9 (5)	-5 (9)	85 (52)
Other pauses (of 3 s or more) minus those above	65 (83)	37 (52)	-28 (35)	64 (47)	74 (89)	57 (64)	-17 (28)	57 (39)
Length of pause (s) immediately before mode change (∗)	3.8 (3)	2.3 (1)	-1.5 (3)	85 (77)	5.8 (0)	2.8 (0)	-3 (1)	56 (22)
Attempting to type-over whilst in insert mode (∗)	5 (3)	5 (10)	0 (8)	54 (110)	16 (22)	6 (3)	-10 (23)	35 (34)
Attempting to insert whilst in type-over mode (∗)	5 (2)	2 (2)	-3 (2)	33 (10)	4 (2)	2 (2)	-2 (8)	50 (10)
Mode changes	58 (9)	54 (8)	-4 (13)	10 (12)	57 (12)	47 (13)	-10 (7)	5 (20)

Table 4.1: Measures used and means obtained (standard deviations in parentheses). ∗ = Data from only 12 subjects analysed; criteria not comparable with any used in the pilot study. The figure for percentage change is the mean of each individual subject's percentage change.

Table 4.2 shows results obtained for overall times taken on both sessions for the two groups. Change in time taken is expressed as a difference between the two sessions and as a percentage of that difference.

Cursor group				Window group			
Task 1	Task 2	Diff.	% change	Task 1	Task 2	Diff.	% change
25.8	21.2	4.6	82.2	46.0	41.1	4.9	89.3
53.1	30.9	22.2	58.2	36.5	32.1	4.4	87.9
44.1	27.0	17.1	61.2	34.3	48.0	-13.7	40.0
30.2	25.9	4.3	85.8	48.8	34.1	14.7	69.9
23.3	24.1	-0.8	103.4	31.9	29.0	2.9	90.9
66.0	48.8	17.2	74.0	23.1	22.2	0.9	96.1
23.5	18.2	5.3	77.5	37.0	31.5	5.5	85.1
29.4	18.6	10.8	63.3	24.1	26.1	-2.0	108.3
22.0	14.5	7.5	65.9	30.5	32.1	-1.6	105.2
21.2	21.4	-0.2	100.9	24.5	20.5	4.0	83.7
61.4	35.2	26.0	57.3	30.1	21.2	8.9	70.4
24.5	22.0	2.5	90.0	52.5	36.4	16.1	69.3
		Mean = 9.7	76.6			Mean = 3.7	91.3
		sd = 8.9	16.1			sd = 7.9	19.9

Table 4.2: Overall times (mins) taken by each subject on the two tasks.

An independent t-test conducted on the *differences* between the groups in times to completion yielded a value of $t = 1.737, df = 22, p < 0.05$. The corresponding t value for differences in *percentage change* between the groups was 1.988, $df = 22, p < 0.05$.

4.7 DISCUSSION AND CONCLUSIONS

This study shows a significant difference ($p < 0.05$) in overall times taken between the two groups over the two sessions, with the time reduction (in minutes) by the Cursor group ($\bar{X} = 9.7, sd = 8.9$) much greater than that of the Window group ($\bar{X} = 3.7, sd = 7.9$). Expressed as a percentage of time taken from the first task, those in the Cursor group ($\bar{X} = 76.6, sd = 16.1$) improved significantly ($p < 0.05$) more than those in the Window group ($\bar{X} = 91.3, sd = 19.9$). This difference substantiates that trend shown in the pilot study (Cursor group $\bar{X} = 77.4, sd = 15.3$; Window group $\bar{X} = 95.6, sd = 21.5$). The speed differential may reside specifically around instances of mode change or be spread more generally through performance. The three-second time period was quite arbitrary and so post hoc analyses were performed on differences in pauses immediately before mode changes using time periods of 0.5 seconds through to seven seconds. No significant differences in the two groups emerged. It is probable that the superiority seen in the Cursor group is a reflection of a general ease of performance on the task.

The significant difference in number of 'verifying' pauses made immediately before a mode change found earlier was not replicated when the study was extended from 12 to 24 subjects. Nevertheless, the trend did remain in that, on the second task, the Cursor group subjects (where information was displayed in both ways followed by information at cursor only) made only 59.2 ($sd = 29.8$) percent of pauses made on the first trial, whereas the Window group subjects (where information was displayed in both formats followed by information at the status line only) made 84.7 ($sd = 52.4$) percent of pauses made in the first trial. There is, in both cases, a learning effect over the two tasks. It is important to bear in mind that the means of establishing pauses of three seconds or more in the pilot study was very approximate. In absolute terms, pauses made immediately before a mode change were reduced in both conditions from first to second trial ($\bar{X} = 4.1, sd = 1.9; \bar{X} = 2.6, sd = 0.6$ respectively), again demonstrating a learning effect. As a percentage change from the mean times of the first trial to the second trial, the Cursor group's pauses reduced much more than those of the Window group ($\bar{X}\% = 84.8, sd = 77.5; \bar{X}\% = 55.9, sd = 22.3$ respectively).

In absolute terms and discounting those pauses made immediately before a mode change, there was a greater diminution of pauses (of three seconds or more) made between the tasks for the Cursor group ($-26.8, sd = 35.2$) over the Window group ($-17.4, sd = 28.2$). Comparing the second task with the first for both groups there were, for instance, less mistakes left in the typing material (Cursor group $= 0.4, sd = 4.3$; Window group $= 2.1, 6.8$), fewer instances of attempting to type-over whilst in insert mode (Cursor group $= 0.0, sd = 8.2$; Window group $= 10.2, sd = 23.1$), and fewer instances of attempting to insert whilst in type-over mode (Cursor group = $0.2, sd = 2.5$; Window group = $2.0, sd = 8.4$).

Although both groups reduced the number of mode changes made between the two tasks, this reduction was more marked in the Window group (Cursor group = $-3.0, sd = 12.8$; Window group = $-10.0, sd = 7.2$). In view of performance differences, this marked reduction for the Window group would not seem to be beneficial. It is possible that where mode was easier to determine (when depicted at the cursor) subjects were more likely to use this, thereby making more frequent mode changes and, in turn, benefitting from the option.

Subjective remarks were also elicited. It was generally found that of those in the Cursor group (peripheral information removed), many reported that they had learned to use the cursor information on the first task and felt no effect of the change. Those in the Window group reported generally that the loss of cursor information in the second session was felt and that they had to give more consideration to which mode they were in. The notion of memory was spontaneously mentioned by more than half the subjects (including a number of those who were not students of Psychology). One subject remarked of the first task that she '...needed to look at the top at first to check the mode but after a while I could stop as I knew'. Of the second task this subject (Window group) noted that having to look up at the cursor line was a 'distraction... eyes follow text... better to have information at the cursor. It's much easier to register the cursor in the text where you are looking'. This subject (who was not a student of Psychology) felt that she was probably slower on the second trial (which she was) because of the lack of cursor information. She also suggested (as others had done) that individuals accustomed to using different editors at different times might not pay much attention to the cursor information because the status line seemed more definite, permanent, and/or consistent. She furthermore drew an analogy with a driving mirror and the attentional shift involved, particularly when it is not set to catch activity via peripheral vision, using the term 'having to unplug your head from one area'. Another subject said that repeatedly referring to the status line information was 'too much of a fuss to keep wandering up and down'. However, such a view was directly contradicted by two others, who reported that looking up was not disruptive.

The most plausible interpretation of these results would seem to be along the lines of the attentional differences outlines in §4.2.2. Russo (1978) pointed out that the use of eye movements in scanning is likely to be accompanied by a cognitive cost in terms of both the execution time and the preparation time for the activity. We believe that designers should give more careful consideration to these aspects. Although there are limits, both practical and in terms of cognitive processing, to the amount of information which can be provided with cursors, there are many possibilities (colour, shape, blink rate etc.) which may be explored.

However, a reminder is also, by its nature, a *distractor* and its value may vary accordingly. There are many times when reminders may be more welcome than others, and when human activity is hindered by an over-abundance of interrupts. These may be external (e.g. a deluge of incoming telephone calls), but lack of concentration seems also characterised by the appearance of too many distracting thoughts from memory. As Findlay *et al.* (1988) have commented, such situations may be contrasted with the low level of interrupts characteristic of an individual 'lost in thought', perhaps engaging in a search for inspiration.

It was said in the Introduction that the effort occasioned by the use of external memory props is less demanding than the cognitive effort required to encode and retrieve information from internal memory sources. However, in some situations memory aids can become cumbersome and unproductive. A study by Davies *et al.* (1990) investigated the use of internal and external memory props within the context of learning to use a word processor. Results showed that those subjects provided with an external memory prop in terms of a window containing a list of available commands performed no better than a group who had spent a short time committing the commands to internal memory. The performance of the former group was considerably disrupted by the removal of the memory prop information. In this case, apparently, use of a simple learning strategy could obviate the need for a memory prop.

In this situation, the command set, once learned, remained unchanged. Memory may also be used in a more dynamic way to follow a changing situation. Under appropriate circumstances, human memory is very good at such a task. For example, it is generally possible to 'keep track' of the topic in a dialogue or of the individual's location within the environment. However, it is a familiar experience within interacting with computers that users are liable to become 'lost' during an interactive session.

EXPERIMENT 5: COMBINING KEYSTROKE AND EYE MOVEMENT ANALYSES

4.8 INTRODUCTION

VDUs and computerised systems may be used to investigate eye movements, and some work (e.g. this thesis) involves eye movement research to investigate interactional VDUs (also see §1.15). There remains great potential for precisely monitoring eye movements not only in response to state information but also to information displayed on the screen generally. For instance, if information can be displayed peripherally without requiring a major attentional break, then this would have bearing on interface performance. More importantly perhaps, such a scenario would 'feel right' to the user and thus have positive commercial implications. It was therefore decided to combine keystroke monitored data and eye movement measurements.

In Experiment 4, it remained uncertain to what extent subjects were actually looking at information displayed on the status line, and more specifically, where visual shifts were made, how attentionally disruptive such movements were likely to be. Use of a peripheral indicator does require appreciable processing effort and it was felt that it would be of benefit for further investigations to explore whether such interrupts are *mini* or *micro* shifts. For instance, what might be the effects on performance if the status line were varied between the standard font messages of low prominence (which must be fixated to be perceived) and a highly salient, thick bar present only when in one particular mode and absent when in another (thus conveying state without requiring reading)?

4.9 METHOD

Two male subjects who were accustomed to participating in search coil studies but who had not been involved in the main study (Expt. 4) took part in this small-scale follow-up.

Equipment used was as described in Experiment 4. In addition, subjects were required to wear the scleral search coils, as described in Experiment 3. Subjects were simply required to perform a shortened version (123 words in length) of the editing task (with information presented at status line only). The text to be edited was presented at the right-hand side of a split-screen display arrangement with the 'model' correct version at the left-hand side of the screen (see Fig. 4.8).

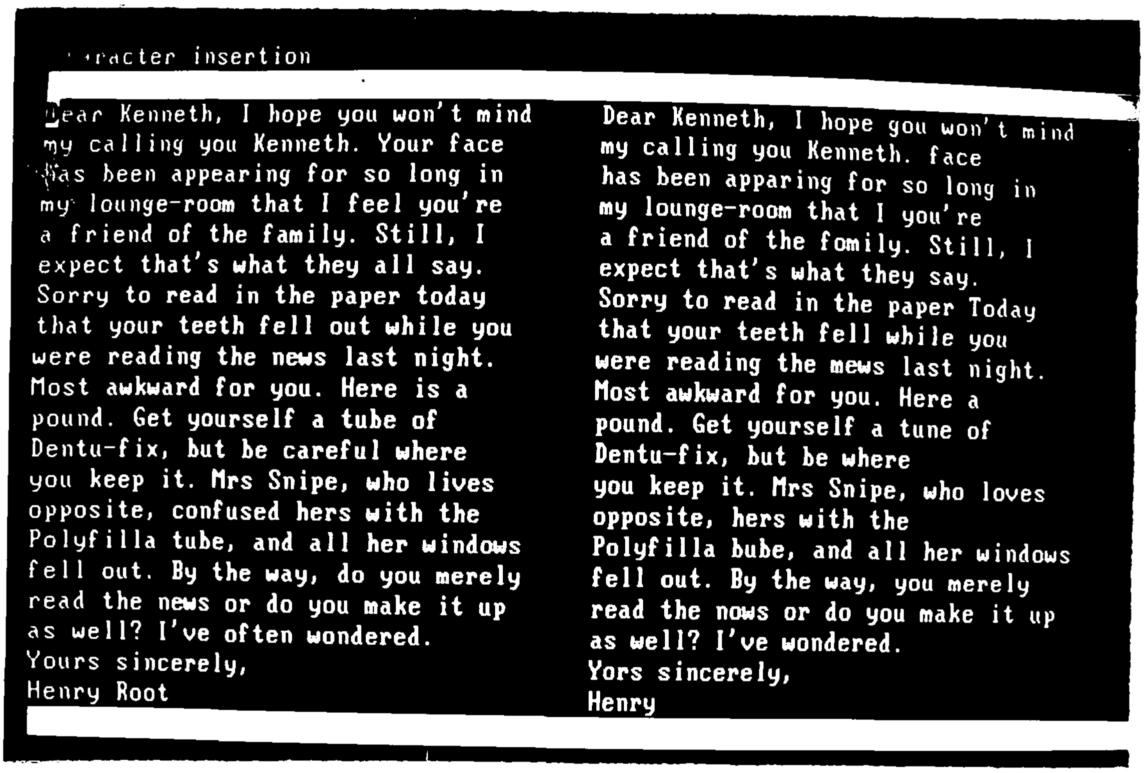


Figure 4.8: The text editing task involved in Experiment 5.

Programme 'SAMDIS' was employed for recording and analysing eye movements. After calibration with the 11″ IBM PC Display monitor at a viewing distance of 55 cm, the programme was able to provide x and y axis coordinates of eye position at sampling intervals of 100 ms. The programme was also set to give a verbal description of the approximate area of gaze (including off-screen fixations); e.g. 'screen left', 'screen right', 'status bar', and 'keyboard'. In addition, the programme gave a screen line number position of area of view (accurate to $\pm$ 1 vertical line).

From a print-out of such a record, it was then possible to correlate items from the time-marked HIMS record and to edit-in corresponding items onto the eye movement record. Table 4.3 shows sections from such a record for one subject.

4.10 RESULTS

Table 4.3 shows sections from a combination of the SAMDIS eye movement programme (columns 1–5) and the HIMS record (columns 6–7). This was achieved through manual editing. Column 1 gives the time-marked record in 100 ms sampling intervals. Columns 2 and 3 give the horizontal and vertical coordinates of eye position (arbitrary units). Column 4 denotes a 'short-hand' area of the screen where gaze is directed (derived from the x and y coordinates). The symbol '?' in this column denotes that the subject's fixation was momentarily positioned outside of the screen/keyboard area. Column 5 shows the line number where the eyes where directed (where appropriate). Column 6 displays keystrokes made within the particular 100 ms time interval. 'Ctrl-S' (control/S) depicts that the subject changed from type-over to insert mode; '$\Leftarrow$' depicts a leftward cursor movement; '$\Rightarrow$' depicts a rightward cursor movement; and '$\Longleftrightarrow$' depicts a spacebar press. Column 7 gives a verbal description (edited in) of what is clearly shown from a combination of the pre-set parameters available with the HIMS (e.g. 'Pauses for more than 180 s') and also what was apparent from a complete 'player-piano' replay via the HIMS of a session as shown on the monitor. Cursor position is given in row and column numbers (e.g. R4C23 denotes the cursor being at the fourth row down and the twenty-third column from the left).

From the data it emerged that the subjects consulted the status line an average of 11 times each within the average 55 seconds which they spent on the task. Glances at the status bar took an average of 300 ms each (sd = 100 ms). Although this task was contrived to cause much mode changing, the fact that these users were consulting status information at an average rate of once every ten seconds strengthened the belief that where necessary information is presented outside of the area of peripheral vision, then attentional interrupts must be considerable. It would have been interesting to see whether glances at the status line were significantly less often when cursor information was available, but the cursor information present/absent variable was not part of this particular experiment. This question of information availability/conspicuity in connection with eye movements is returned to in Experiment 7.

Time (s)	Pos x. Pos. y	Area	Line	Keystroke	Behaviour
.00	(-455, 735)	STATUS BAR		⇒	In type-over mode.
.10	(-455, 742)	STATUS BAR		⇒	Starts with cursor
.20	(-460, 745)	STATUS BAR		⇒	at R1C1
.30	(-460, 735)	STATUS BAR		⇒	
.40	(-460, 737)	STATUS BAR		⇒	
.50	(-455, 740)	STATUS BAR		⇒	
		....			
13.50	(-343, 625)	SCREEN LEFT	(1)		
13.60	(-343, 615)	SCREEN LEFT	(1)		
13.70	(-337, 622)	SCREEN LEFT	(1)		
13.80	(-343, 615)	SCREEN LEFT	(1)		
13.90	(-414, 776)	STATUS BAR			Consults status
14.00	(-414, 784)	STATUS BAR			bar (1.89s
14.10	(-419, 737)	STATUS BAR			before change)
14.20	(-645,-1099)	KEYBOARD		⇒	Moves cursor one
14.30	(-670,-1131)	KEYBOARD		⇓	place to the right
14.40	(-665,-1128)	KEYBOARD		⇓	and down three lines
14.50	(-619,-1154)	KEYBOARD		⇓	to R4C23, over-
14.60	(-619,-1152)	KEYBOARD			shooting two lines.
14.70	(-614,-1149)	KEYBOARD			Moves cursor up
14.80	(-343, 623)	SCREEN LEFT	(9)	⇑	two lines
14.90	(-327, 562)	SCREEN LEFT	(3)	⇑	to R2C23
15.00	(-332, 549)	SCREEN LEFT	(3)		but
15.10	(-332, 552)	SCREEN LEFT	(3)		overlooks
15.20	(-358, 609)	SCREEN LEFT	(2)		the need to
15.30	(-337, 633)	SCREEN LEFT	(1)		insert word
15.40	(-327, 656)	SCREEN LEFT	(0)		'Your' at
15.50	(-327, 630)	SCREEN LEFT	(1)		R2C25.
15.60	(-322, 615)	SCREEN LEFT	(1)	⇓	Moves cursor down
15.70	(-322, 615)	SCREEN LEFT	(1)	⇐⇐⇐	to R3C23 and left
15.80	(-322, 622)	SCREEN LEFT	(1)	⇐⇐⇐	to R3C13.
15.90	(-317, 612)	SCREEN LEFT	(1)	⇐⇐⇐	Changes to
16.00	(-317, 599)	SCREEN LEFT	(2)	Ctrl-A	insert mode.
16.10	(-322, 609)	SCREEN LEFT	(2)		
16.20	(-255,-1157)	KEYBOARD			
16.30	(-261,-1170)	KEYBOARD			
16.40	(-307,-1058)	KEYBOARD			

16.50	(-373,-1081)	KEYBOARD			
16.60	(-276, -245)	SCREEN LEFT	(20)	e	Types 'e' to
16.70	(-353, 458)	SCREEN LEFT	(5)		change 'apparing'
16.80	(-337, 458)	SCREEN LEFT	(5)		to 'appearing'.
16.90	(-322, 586)	SCREEN LEFT	(2)		
17.00	(-332, 555)	SCREEN LEFT	(3)		
17.10	(-332, 560)	SCREEN LEFT	(3)		
17.20	(-332, 542)	SCREEN LEFT	(3)		
17.30	(-184, 581)	SCREEN LEFT	(2)		
17.40	(-184, 570)	SCREEN LEFT	(2)		
17.50	(-189, 578)	SCREEN LEFT	(2)		
			....		
111.50	(-399, 132)	SCREEN LEFT	(12)	Ctrl-S	Changes to insert
111.60	(-399, 153)	SCREEN LEFT	(12)		mode
111.70	(-409, -625)	?			
111.80	(-435,-1357)	KEYBOARD			
111.90	(-424,-1399)	KEYBOARD			
112.00	(-430,-1446)	KEYBOARD			
112.10	(87,-1355)	KEYBOARD			
112.20	(199,-1337)	KEYBOARD			
112.30	(209,-1344)	KEYBOARD			
112.40	(107,-1558)	KEYBOARD			
112.50	(97,-1571)	KEYBOARD			
112.60	(92,-1574)	KEYBOARD			
112.70	(-66,-1436)	KEYBOARD			
112.80	(-76,-1433)	KEYBOARD		c	At L14C11 types
112.90	(-312,-1480)	KEYBOARD			'confused'
113.00	(-337,-1483)	KEYBOARD			and hits
113.10	(-343,-1470)	KEYBOARD			space-bar.
113.20	(-230,-1449)	KEYBOARD		o	
113.30	(-5,-1399)	KEYBOARD			
113.40	(0,-1391)	KEYBOARD			
113.50	(-230,-1394)	KEYBOARD		n	
113.60	(-435,-1410)	KEYBOARD			
113.70	(-435,-1404)	KEYBOARD			
113.80	(-445,-1375)	KEYBOARD			
113.90	(-363, -982)	KEYBOARD		f	
114.00	(-343, -104)	SCREEN LEFT	(17)		

114.10	(-337,	-99)	SCREEN LEFT	(17)		
114.20	(-332,	-75)	SCREEN LEFT	(17)	u	
114.30	(-317,	-72)	SCREEN LEFT	(16)		
114.40	(-322,	-78)	SCREEN LEFT	(17)	s	
114.50	(-317,	-78)	SCREEN LEFT	(17)		
114.60	(-317,	-91)	SCREEN LEFT	(17)		
114.70	(-322,	-75)	SCREEN LEFT	(17)	e	
114.80	(-225,	-44)	SCREEN LEFT	(16)		
114.90	(3450,	-1970)	KEYBOARD			
115.00	(-220,	-13)	SCREEN LEFT	(15)	d	
115.10	(-215,	26)	SCREEN LEFT	(14)		
115.20	(-225,	41)	SCREEN LEFT	(14)		
115.30	(-220,	65)	SCREEN LEFT	(13)	⟺	
115.40	(-220,	93)	SCREEN LEFT	(13)		
115.50	(-220,	114)	SCREEN RIGHT	(12)		
115.60	(266,	127)	SCREEN RIGHT	(12)		
115.70	(271,	122)	SCREEN RIGHT	(12)		
115.80	(424,	127)	SCREEN RIGHT	(12)		
115.90	(435,	75)	SCREEN RIGHT	(13)		
116.00	(409,	41)	SCREEN RIGHT	(14)		
116.10	(189,	33)	SCREEN RIGHT	(14)	⇓	Moves down and
116.20	(153,	172)	SCREEN RIGHT	(11)	⇐	left nine
116.30	(92,	114)	SCREEN RIGHT	(12)	⇐	spaces
116.40	(15,	99)	SCREEN RIGHT	(13)	⇐	
116.50	(15,	93)	SCREEN RIGHT	(13)	⇐	
116.60	(-20,	91)	SCREEN LEFT	(13)	⇐	
116.70	(-312	112)	SCREEN LEFT	(12)	⇐	
116.80	(-312	112)	SCREEN LEFT	(12)	⇐	
116.90	(-476	109)	SCREEN LEFT	(13)	⇐	
117.00	(-476	122)	SCREEN LEFT	(12)	⇐	
117.10	(-404	132)	SCREEN LEFT	(12)		

Table 4.3: Sections from a combination of the SAMDIS eye movement programme (columns 1–5) and the HIMS record (columns 6–7). See text for details.

4.11 DISCUSSION AND CONCLUSIONS

A specific benefit from this record is that it supports the notion of pauses of three seconds or more before a mode change as being characterised by verifying the mode status through an attentional 'micro switch'.

The major advantages of computer capture of behaviour, as opposed to video or verbal protocol techniques, are that it is automatic, error free and very large volumes of data can be efficiently collected. For example, Draper (1984) was able to use computer capture to investigate the use of UNIX commands by 94 users over a period of eight months by taking advantage of the UNIX system 'accounting' facility which records every process run and who ran it. Additionally, the data collected by computer capture are often in a format which is readily analysable statistically. 'The main disadvantage of computer capture, if used as the only observational method, is that, while often voluminous, the data are impoverished because all the other user behaviours that do not involve actually using the computer's input device(s) are not recorded. Thus, for the purposes of task analysis, computer capture is often best used in addition to other recording methods. In particular, computer capture can remove much of the burden of using pen-and-paper recording methods as it allows the observer to concentrate on only the user behaviours not captured by the computer' (Diaper, 1989, p. 229). In our case, this problem was obviated in that the other behaviour of interest (eye movements) were captured seperately and easily synchronised with keyboard performance. Indeed, it would have been a fairly simple matter to record these aspects of behaviour simultaneously, although this was not deemed necessary for present purposes.

It may be possible that the presence of the correct text consistently on the left-hand side of the screen, with status information displayed consistently in the upper left-hand corner may have had particular effects on eye movements. It is suggested that, for any future follow-up studies, that a counter-balancing procedure be employed with the correct text versions being displayed equally on right- or left-hand screeen sides.

These conclusions may seem somewhat weak, but it must be stressed that this synthesis of two unique behavioural analyses was never intended to be particularly quantitative. This dual record is innovative and this type of fine detail has considerable potential value to those working within a task analysis approach to HCI.

Chapter Five

Use of Peripheral Vision

5.1 OVERVIEW

The next study also involved the HIMS and a text editor. Four levels of peripheral salience (conspicuity) of status information were employed. This varied the possibility of using peripheral vision rather than saccades and attentional interrupts. Superior performance was found in the condition with high conspicuity. It was suggested that this was accounted for by the relative lack of disruptive saccades, reading and relocating one's place. Again, the main study was supported by a fine-grained eye movement analysis of the visual tasks involved. It was concluded that the peripheral presentation of text-editing status information is at present grossly underused.

EXPERIMENT 6: PERIPHERAL VISION, STATUS INFORMATION, AND CONSPICUITY

5.2 INTRODUCTION

The movements of one's eyes and head provide a mechanism for gross *selection* among visual stimuli. The position of the eyes at any point in time will obviously determine the pattern of stimulation of the retinal projections and, therefore, the stimuli which are of potential interest. In the natural environment, people want to look *at* what they are looking *for*. Variables which influence visual search are therefore major determinants of selection in normal perception.

Text editing may be regarded as a paradigmatic example of HCI. There are several reasons for this:

(1) The interaction is rapid. A user completes many transactions per minute for sustained periods.

(2) The interaction is intimate. As with all well designed tools, a text editor becomes an unconscious extention of the user, a device to operate *with* rather than operate *on*.

(3) Text editors are probably the single most heavily used programmes, and their use continues to increase within the office environment.

(4) Computer text editors are similar to, and can therefore be representative of, other systems for HCI. 'Like most other systems, they have a discrete command language and provide ways to input, modify and search for data. The physical details of their interfaces are not particularly unique. Because of these similarities, progress in understanding user interaction with text editors should aid the understanding of other systems' (Card *et al.*, 1983, p. 101).

The study of text editing is a task within the range of the analytic tools at the disposal of the cognitive ergonomist. Because of the intrinsic importance of the task itself, the similarities with other tasks, and the task's tractable complexity, studies of computer text editing are a natural starting point in the study of HCI.

One of the most important characteristics of VDU-based tasks such as searching for a menu item or a particular word is the distinction between what is in foveal as compared with peripheral vision at any particular moment (as is so of other everyday searches). Objects in central vision can be identified in detail, permitting the perceiver to determine whether it is the target being sought. As that determination usually will be negative—otherwise the search ends—the observer will have to shift his gaze to a different area of the field of view. The decision as to where to look next is presumably based on information available in peripheral vision. This information will not be in fine detail as pattern acuity drops 50 percent for an object located only one degree from the centre of the fovea and, for an object located eight degrees from the centre, it is only 15 percent of maximum (Riggs, 1965). Other kinds of information can be used apart from fine detail. A moving object, for example, is important information and might elicit a saccade, even when the moving object is perceived far out in the periphery and the object cannot be identified (Haber & Hershenson, 1980).

Therefore, the search process appears to consist of two components which occur almost simultaneously: (a) an identification process of parts of the retinal projection falling on the fovea, and (b) a decision process concerning the direction of the next eye movement based on information from the periphery.

5.2.1 Attentional fields

Information supplied by the stimuli from the periphery of the retina is obviously of considerable importance in normal perception. Sanders (1963) conducted a series of experiments concerning this issue which are followed here as described by Haber and Hershenson (1980). Two signals generally appear within these studies, one directly ahead and the other out to the right. The subject is typically required to press a morse key as quickly as possible to indicate the presence of one, both, or some combination of these signals. It is then possible to compare reaction times for the different responses.

On the basis of performance and of the mechanisms of visual orientation, three levels of the functional visual field can be discerned:

(1) the *stationary field,* where mere peripheral viewing already leads to efficient performance;

(2) the *eyefield,* where the supplementary use of eye movements is required; and

(3) the *headfield,* where head movements are also required.

In one series of experiments, Sanders mapped the relative sizes of these fields. An observer was presented with two patterns, each of which was a column of either four or five lights. He had to press one of four response keys to indicate whether the two stimuli both contained four lights, or both five, or the left one four and the right one five, or vice versa. It was necessary for the observer, who was instructed to fixate the position where the left-hand stimulus would appear, to view both stimuli. The right-hand stimulus was from 19° to 94° to the right. Both stimuli simultaneously appeared.

When the head and eyes were free to move, two discontinuities in performance occurred, one when the patterns were approximately 30° apart, and one when they were about 80° apart. When the observer's head was fixed so that head movements were not possible, performance rapidly diminished for larger display angles (81° to 94°). It was suggested by Sanders that the two slopes corresponded to the break points between the display and eye fields, and the eye and head fields, respectively. Thus, with relatively large and simple stimuli, peripheral vision is adequate out to about 30° and eye movements extend the visual field out to about 80°. What this means is that beyond 30° the eyes must move and beyond 80° the head must move.

A further study strengthened this hypothesis. In this situation, the observer was instructed to remain fixated on the left stimulus or to move their eyes between them. When no eye movement was permitted, peripheral vision terminated at about 30°. Even before this, reaction time increased exponentially as the visual angle increased from 0°. When observers were allowed to move their eyes, reaction times were constant out to about 90°, after which it rapidly increased. Even for small angles, this condition was slightly slower, which suggests that the observers' strategies must have been different in some way when they knew they could move their eyes.

Although peripheral information is used to determine the subsequent fixation, the target presently being fixated is also being analysed; i.e. two tasks are performed simultaneously. This raises several questions. Is this a division of attention or are the two tasks performed independently? If this does represent a division of attention, does one task interfere with the other? When the fixated signal and the peripheral signal were both within the eye field—and particularly when they were both within the display field—the RTs to the first signal were increased slightly over that required when only a single signal was presented and no peripheral processing was necessary. This demonstrated that the central and peripheral tasks interacted; at least in Sanders' experimental situation.

Haber and Hershenson (1980) suggest that Sanders' theory of selective acts can be reinterpreted in terms of preattentive processes that operate simultaneously over large areas of the retina (see §2.5.2). Within the display field, preattentive processes are capable of locating and segregating all figures and of testing global characteristics. Sanders' subjects only required information concerning the height or brightness of the stimulus to spot the difference between four or five lights. When the visual angle between the objects in the field exceeds the limits of the preattentive processes, the spatially parallel nature of the processing is lost and an eye movement is necessary to complete the testing. One of Sanders' most important findings is that even out to 80° in the periphery, some information is available. This information reduces the amount of time required to process the stimulus when it is directly fixated after an eye movement. Preattentive processes thus appear to operate over wide areas of the peripheral retina. When a slower head movement becomes necessary, the preattentive processes have obviously lost their spatially parallel character and two seperate glances are needed.

Finally, Sanders' notion of selective acts illustrates a further interaction between control and identification processes in visual search. When single targets are not too far into the periphery, it seems that some identification can occur, even whilst the eye is being directed to move towards them.

The issues of conspicuity and directed attention were introduced in §2.8.

5.3 METHOD

Subjects Sixteen paid subjects (7M:9F) participated. Their ages ranged from 21 to 55 with a mean of 28.9 ($sd = 10.2$). All had normal corrected or uncorrected visual acuity and no known impairments of colour vision. Computer/keyboard experience varied from almost nil to several years experience of word processing and/or text editing.

Four short passages of humourous material (April Fool anecdotes) of equal length formed the basis of the stimulus material. A 'degraded' version of each was prepared so as to include an equal number of misspellings and an equal number or errors of omission in each one. These two types of errors were randomly scattered throughout the passages so as to produce an extreme clerical-type task and one which optimised the need to irregularly change between insert and type-over mode, ideally with as many occasions where users would forget the current mode and seek external cues as to current mode.

The NUMAC screen-editor programme ('SE'), as described in §4.5, again formed the basis of the software used in this study. The first stage of modification involved producing a two-column display within which the correct text could be presented on the left hand side and text to-be-corrected (degraded text) presented in the right hand side (see Fig. 5.1). This was achieved with the feature that the cursor, at the end of a line, did not 'wrap around' to the very left of the screen, but merely located itself at column 41 of the next line.

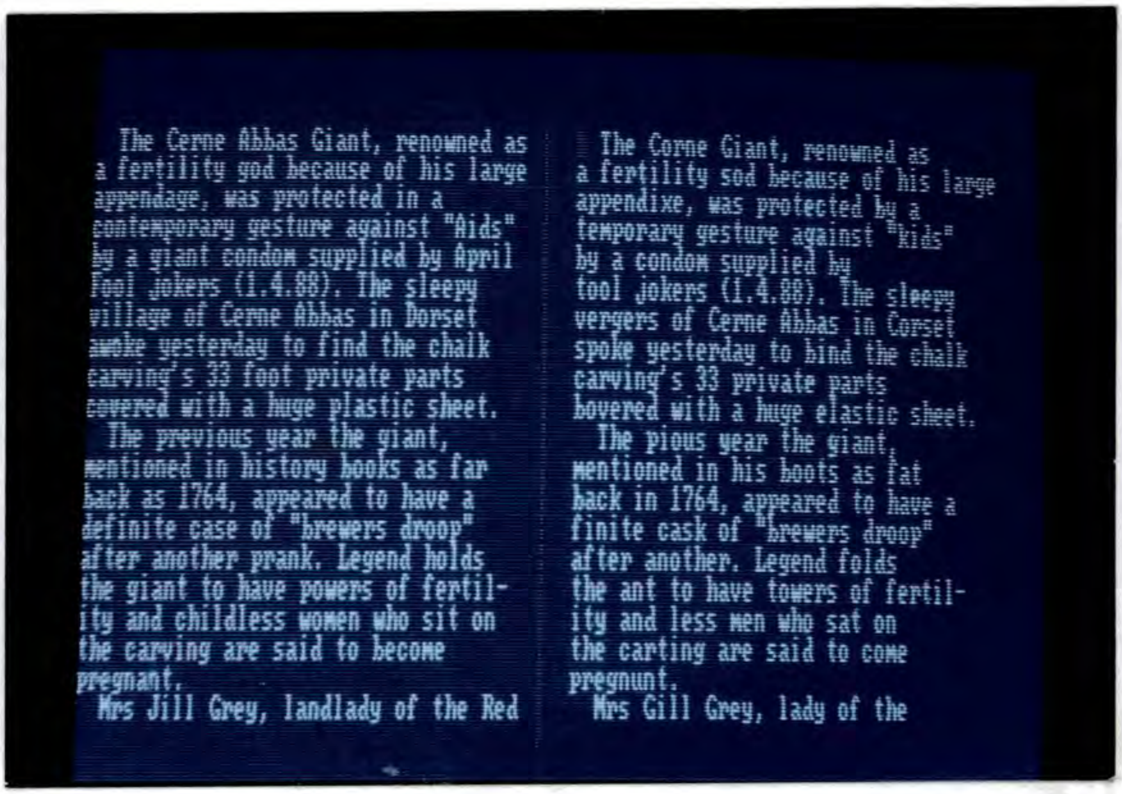


Figure 5.1: Example of split screen arrangement. This example shows how the screen would appear in the insert mode of the Low/intermediate, High/intermediate and High Conspicuity conditions.

A second stage of modification consisted of preparing four versions each with differences designed to produced a range of levels of conspicuity of status information relating to insert or type-over mode. This took advantage of: (a) colour (degree of contrast with background), (b) size (area of screen occupied by the information), and (c) spatial positioning of information (where on the screen it appeared).

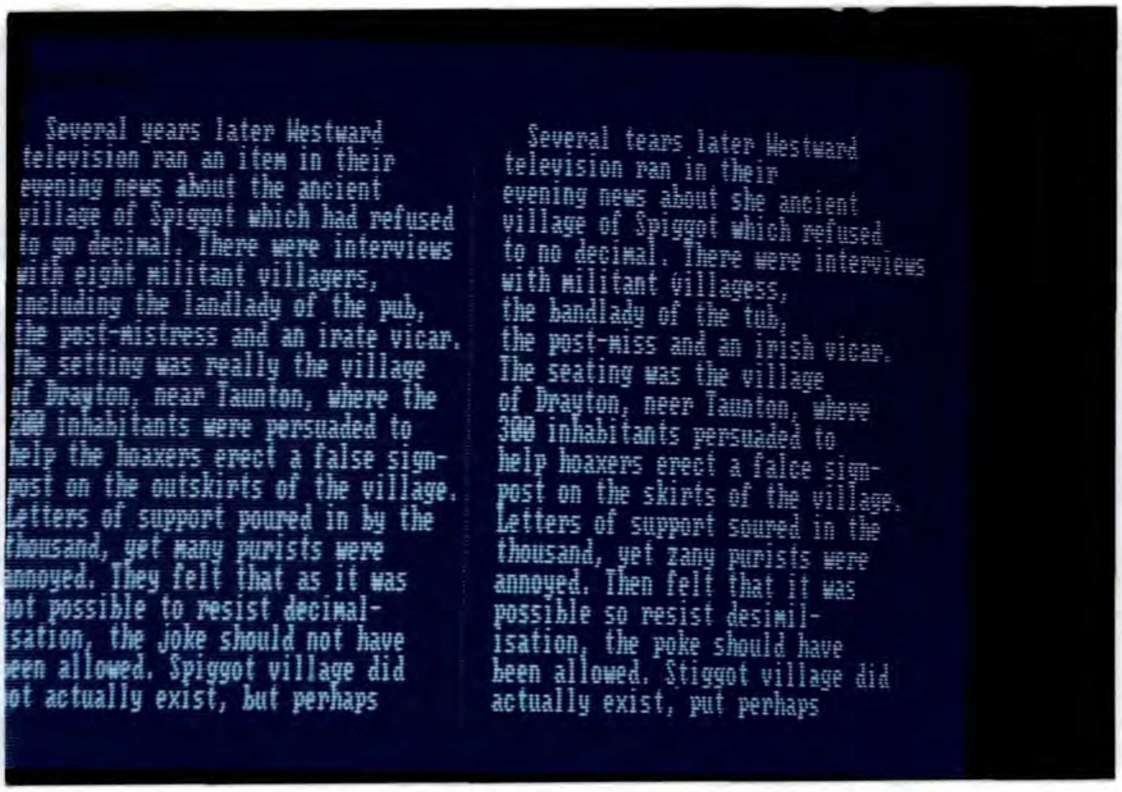


Figure 5.2: Example of Low Conspicuity condition in type-over mode.

The four levels of conspicuity or salience of information are described below.

(1) **Low Level:** Either 'type-over' or 'insertion' was present at the left hand side of a status bar at the top of the screen. These words were presented in maroon lettering against the dark blue background. Ctrl-A changed the mode display mode information to 'type-over' whilst Ctrl-S changed it to 'insertion'. Therefore, in both cases, no extra cues were available from word length, overall luminance, etc. (see Figs 5.2 and 5.3).

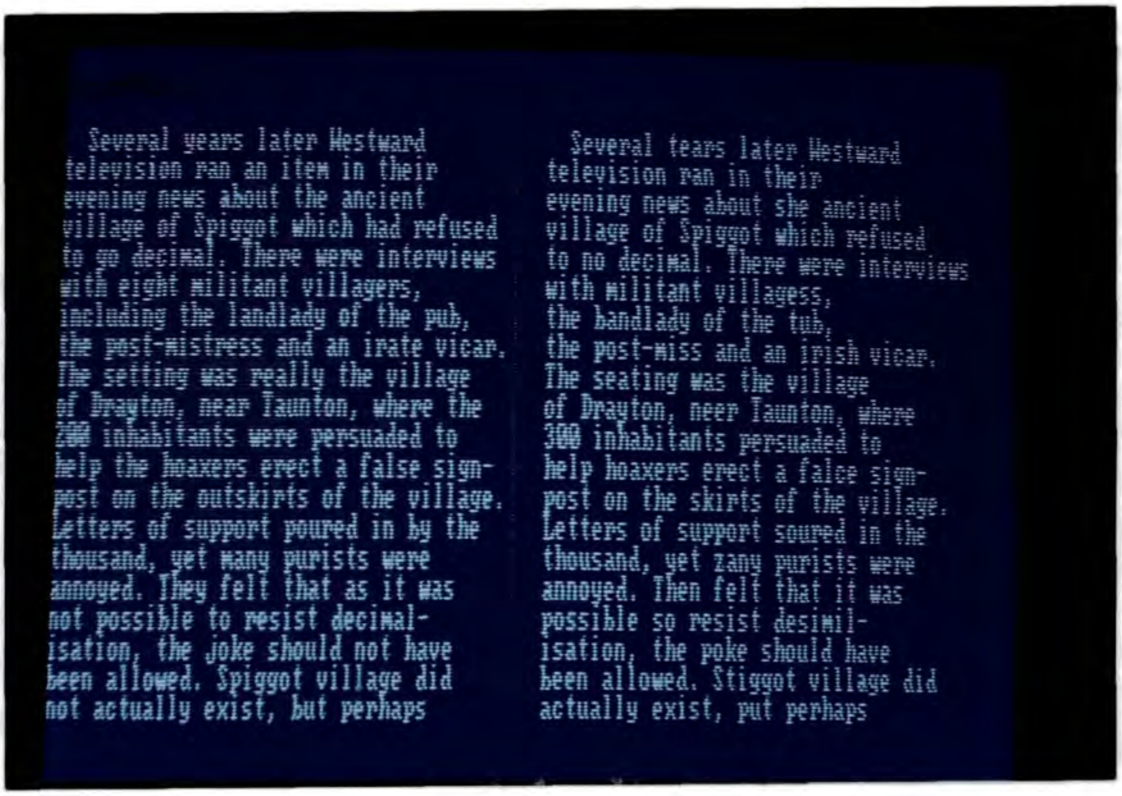


Figure 5.3: Example of Low Conspicuity condition in insertion mode.

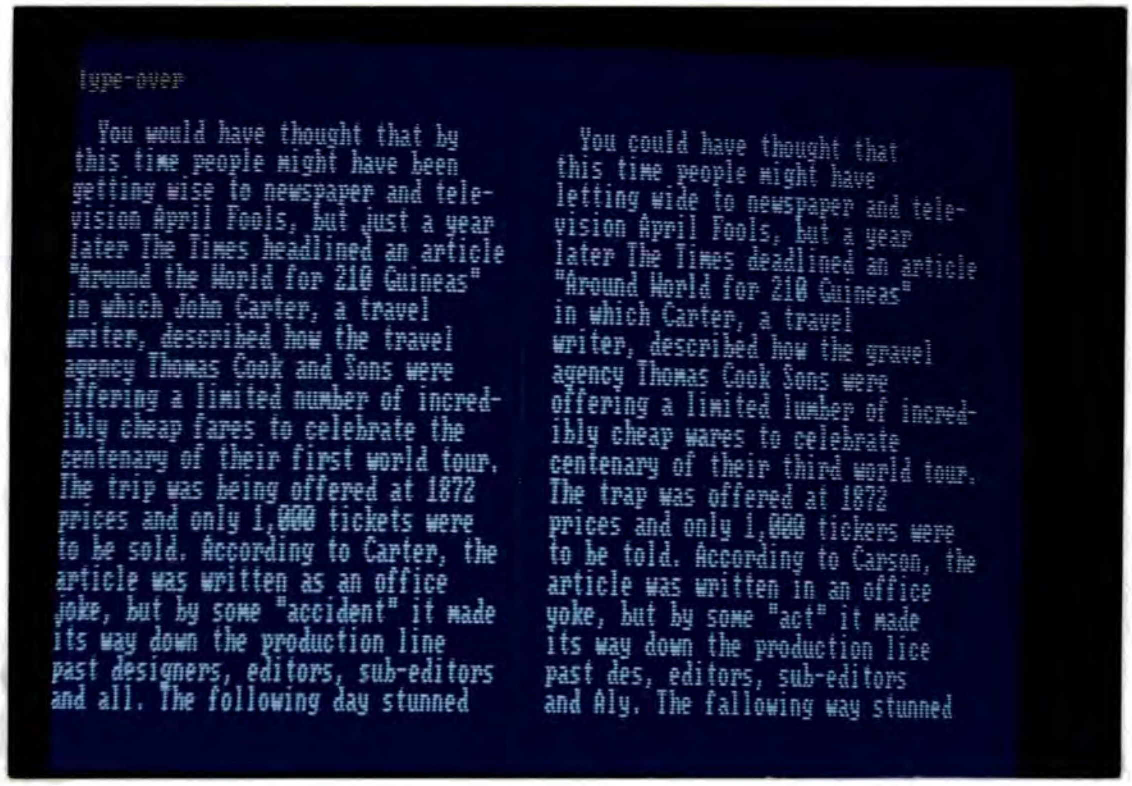


Figure 5.4: Example of Low/intermediate Conspicuity condition in type-over mode.

(2) **Low/intermediate Level:** This condition was as for the Low Level (1) condition, but with the difference that the words 'type-over' or 'insertion' were displayed in bright (greenish) yellow, and hence likely to be more conspicuous (see Fig. 5.4).

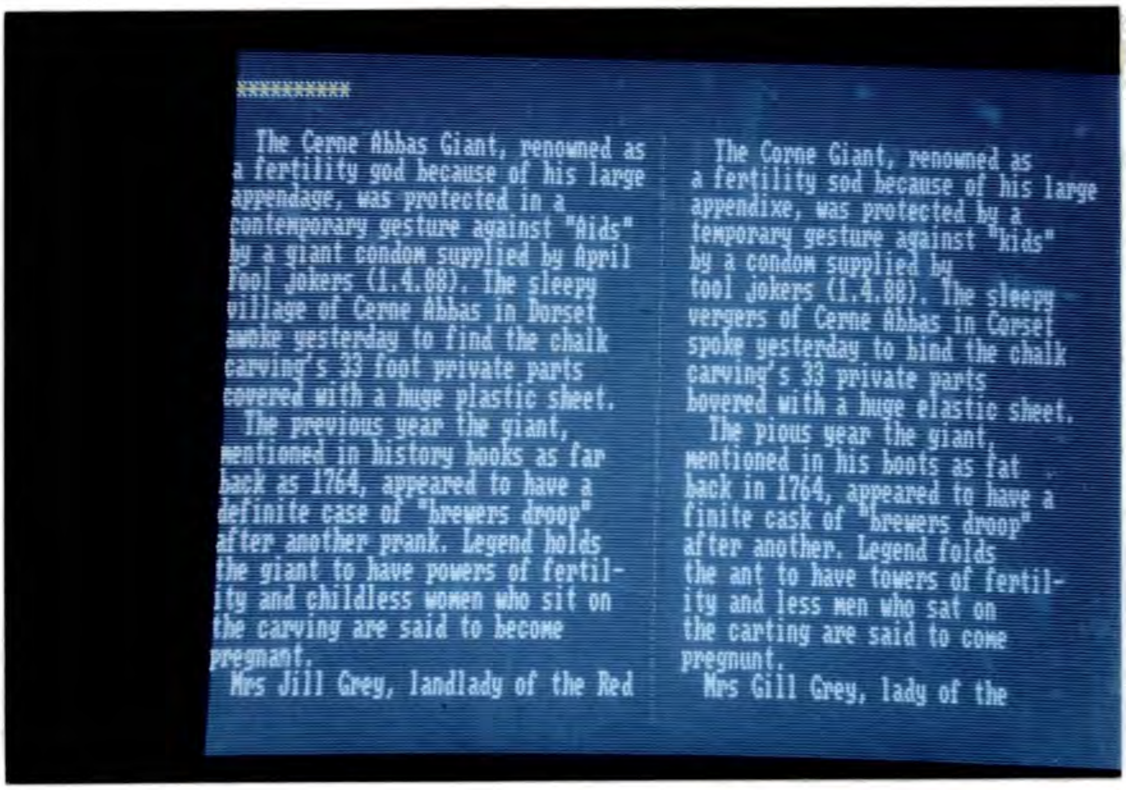


Figure 5.5: Example of High/intermediate Conspicuity condition in type-over mode.

(3) **High/intermediate Level:** Whilst in type-over mode, a row of seven asterisks (∗ ∗ ∗ ∗ ∗ ∗ ∗), coloured bright (greenish) yellow were present at the left hand side of the status bar (see Fig. 5.5). In insertion mode, no information was available (status line blank). It is assumed that being in any mode state would be detectable (particularly when subjects were editing near the top of the screen) without great difficulty and usually without a saccade and associated attentional interrupts.

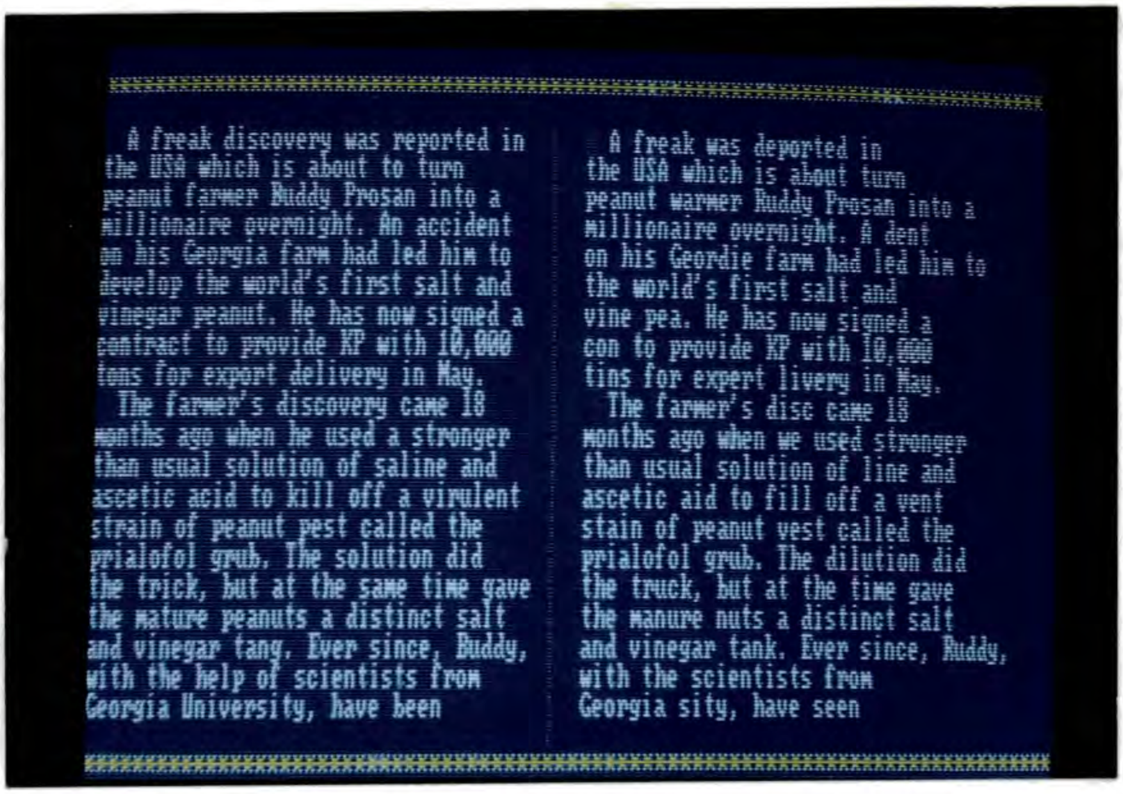


Figure 5.6: Example of High Conspicuity condition in type-over mode.

(4) **High Level:** In type-over mode, a complete row of yellow asterisks filled the status bar both at the top and bottom of the screen (see Fig. 5.6). No information was available in insertion mode. It is assumed that subjects would be less likely to require a saccade than in the above condition in order to check what mode they were presently in.

The information shown in Figure 5.7 was displayed above the keyboard and constantly present throughout the trials.

Type:		**Remember:**
CTRL-A	for type-over	∗∗∗∗∗∗∗∗∗∗∗ = type-over mode
CTRL-S	for insertion	
CTRL-D	for character deletion	
F1	to scroll screen down	
F4	to scroll screen up	

Figure 5.7: Keyboard prompts for subjects.

Each subject was administered all four passages and all four conditions. Presentation order of passages and presentation order of conspicuity level were counterbalanced between subjects.

An IBM PC computer was used with an 11″ IBM PC Display monitor. The monitor projected positive contrast green characters on black background. Subjects were seated with a typical viewing distance of 45 cm, although head movements were unrestricted. The vertical and horizontal viewing area of the screen therefore subtended about 18° and 26°, respectively, on the retina. Screen luminance was 55 cd/m^2. Ambient illumination was a constant 80 lux, measured horizontal to the monitor.

Subjects were simply required to correct the passages. They were instructed to try to maintain a constant viewing distance, although no physical restrictions on movements were actually imposed. Full instructions are included as Appendix 5.1.

The HIMS protocol analyser (as described in §4.4) was used to collect time-stamped keystroke records of all subjects, and enabled playback of the entire session for close examination. The HIMS monitor permitted accurate quantitative records of:

(a) overall times to complete tasks;
(b) pauses of more than 3 s before a mode change;
(c) pauses greater than 3 s at other times;
(d) pauses of more than 5 s before a mode change;
(e) pauses greater than 5 s at other times;
(f) incidence of attempts to change mode when already in the correct mode;

(g) number of other types of errors made (e.g. mis-typings); and
(h) errors remaining after task completion.
Subjective comments were also elicited.

5.4 RESULTS

Table 5.1 combines results from the performance parameters mentioned above.

Performance measure	Conspicuity condition			
	Low	Low/intermediate	High/intermediate	High
Overall times to completion (min)	11.6 (7.4)	9.5 (5.7)	9.3 (4.4)	7.8 (2.6)
Pauses > 3 *s* before mode change	7.6 (9.2)	5.1 (5.1)	5.6 (4.3)	5.4 (10.3)
Other > 3 *s* pauses	54.4 (43.3)	48.1 (41.4)	48.2 (33.2)	46.1 (32.4)
Pauses > 5 *s* before mode change	2.7 (3.4)	1.9 (2.0)	1.7 (2.0)	1.0 (1.6)
Other > 5 *s* pauses	22.7 (17.9)	20.6 (16.7)	21.4 (15.3)	17.4 (14.1)
Redundant mode changes	2.7 (2.0)	1.5 (1.4)	2.0 (1.9)	1.5 (1.8)
Other (typing) errors	21.9 (29.7)	14.0 (13.0)	8.9 (7.9)	13.7 (15.9)
Errors remaining after completion	2.0 (2.6)	1.0 (1.1)	2.2 (3.1)	1.9 (2.8)

Table 5.1: Mean results of all subjects from the HIMS monitoring analysis (standard deviations in parentheses).

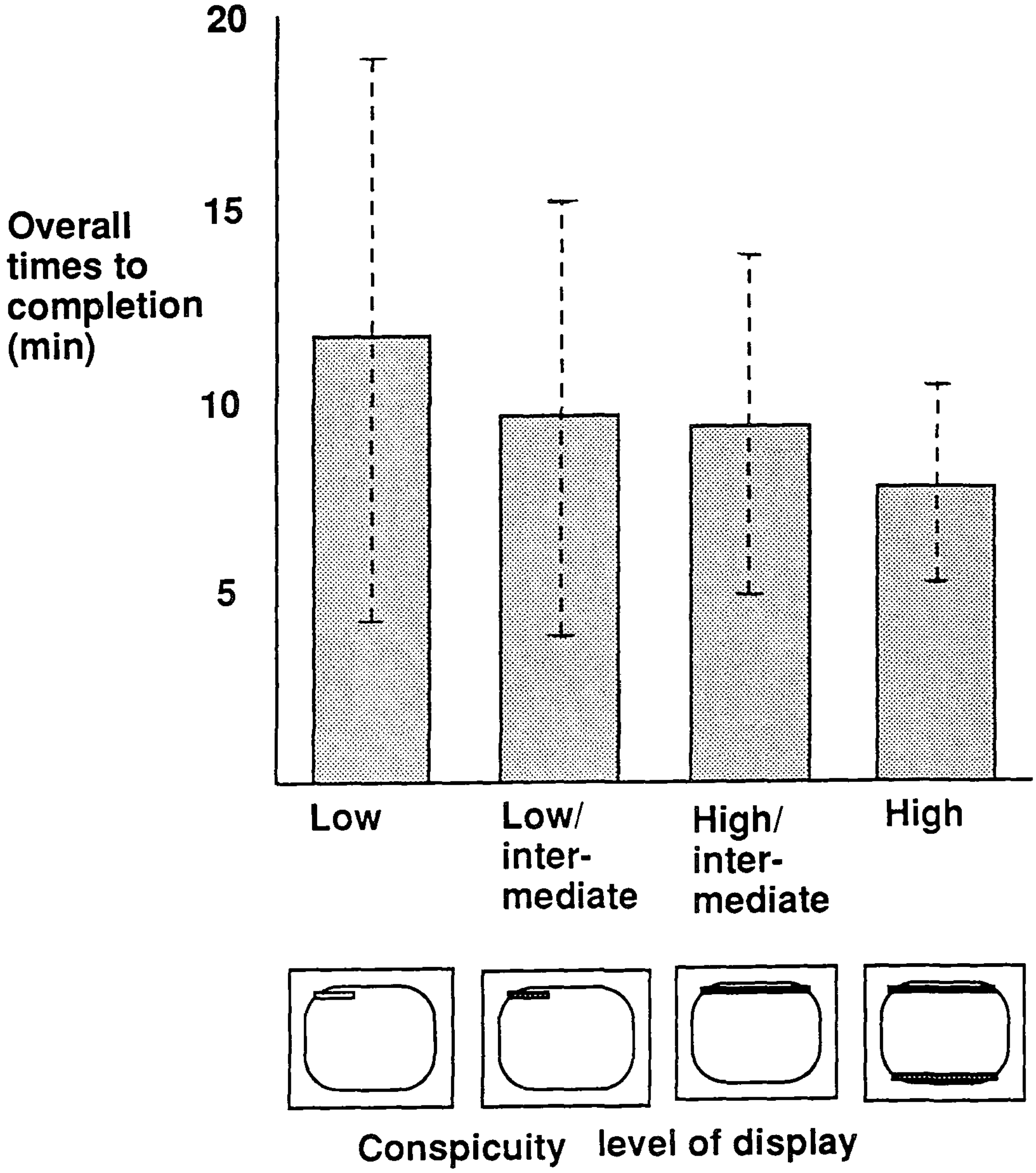


Figure 5.8: The linear trend across the four conspicuity conditions for the overall times (min) to complete the passages. Vertical lines denote standard deviations.

One-way repeated measures ANOVAs were conducted on the various performance measures. There was a significant difference ($F(3,45) = 5.524; p < 0.01$) between the overall times for completion of the passages. Figure 5.8 illustrates the linear decrease in overall task times as conspicuity of status information increases. A Newman-Keuls test showed the significant differences to lie between the Low (11.62 m) and the High Conspicuity (7.80 m) conditions ($q = 60.50; p < 0.01$), and between the Low Conspicuity condition and the Low/intermediate (9.55 m) and High/intermediate (9.27 m) conditions ($q = 33.09$ and 37.52, respectively; $p < 0.05$).

No other significant differences were obtained from analyses of: pauses of more than 3 s before a mode change ($F(3,45) < 1$); pauses greater than 3 s at other times ($F < 1$); pauses of more than 5 s before a mode change ($F = 2.445$); pauses greater than 5 s at other times ($F = 1.224$); incidence of attempts to change mode when already in the correct mode ($F = 1.477$); number of other types of errors made ($F = 2.213$); and errors remaining after task completion ($F < 1$).

Linear trends were also found between: (a) pauses of more than 5 s before a mode change (Fig. 5.9), where there were more such pauses in the Low Conspicuity condition ($\bar{X} = 1.94, sd = 3.02$) than within the High Conspicuity condition ($\bar{X} = 1.62, sd = 3.42$); and (b) pauses of more than 5 s at other times (Figure 5.10), where a similar pattern emerged between the Low Conspicuity condition ($\bar{X} = 22.69, sd = 17.87$) and the High Conspicuity condition ($\bar{X} = 17.37, sd = 14.08$).

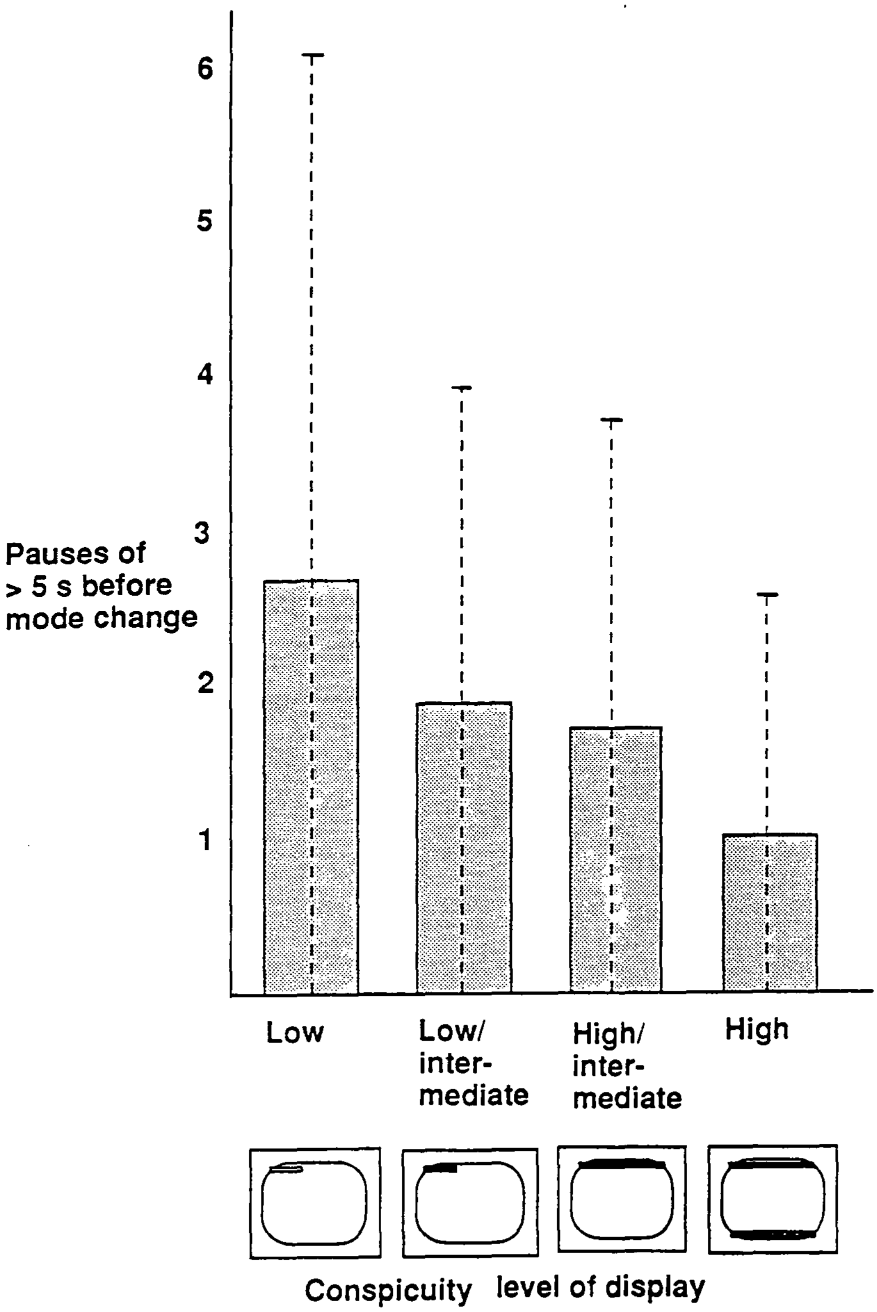


Figure 5.9: The trend across the four conspicuity conditions for pauses of more than 5 s before a mode change. Vertical lines denote standard deviations.

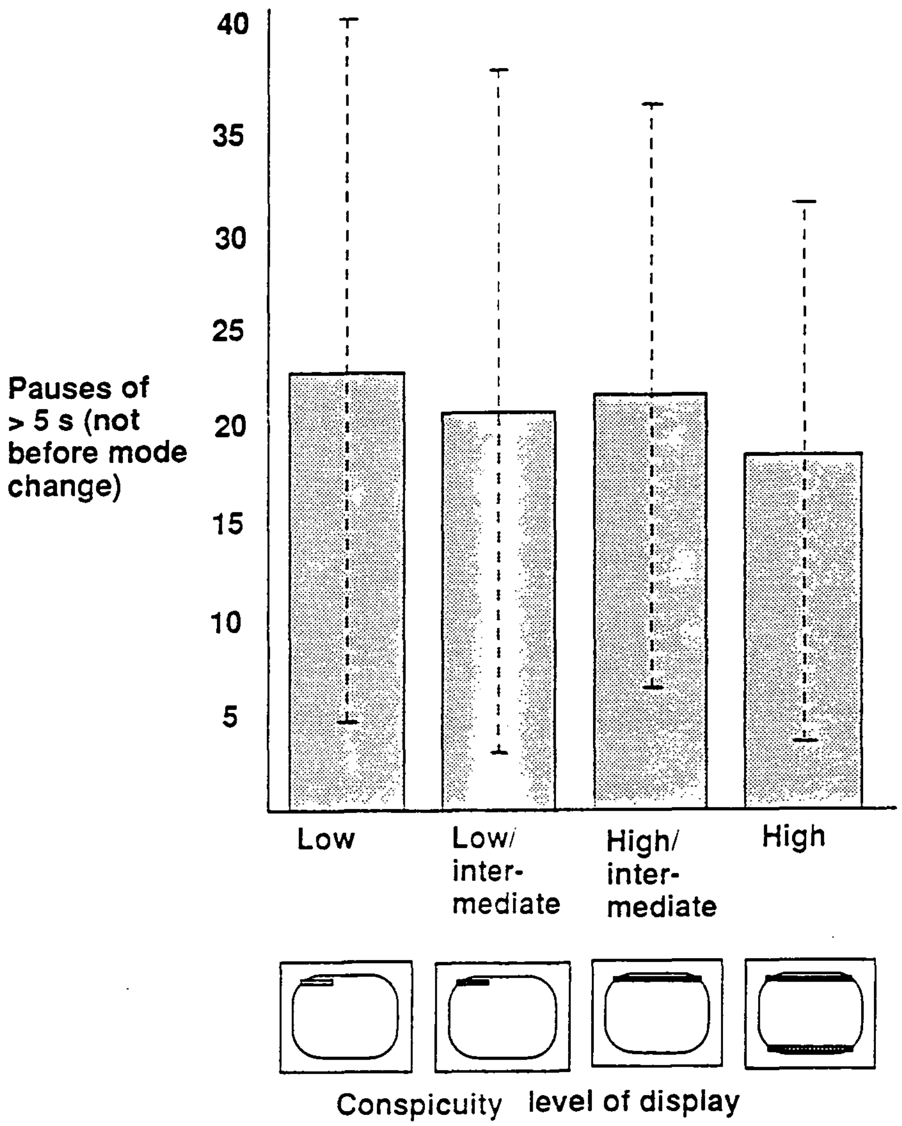


Figure 5.10: The trend across the four conspicuity conditions for pauses of more than 5 s (not before a mode change). Vertical lines denote standard deviations.

5.5 DISCUSSION

Results showed a steady trend in speed of performance improving through the low to high salience conditions, with performance speed being significantly greater where information was accessible via peripheral vision. Other variables such as error frequency, overall incidence of pausing, and number of mistakes made concerning which mode the user was actually in, all showed greater occurrence in the lower peripheral salience conditions. Although seven of the eight measures employed fell short of statistical significance, the consistency of trends in the expected direction across the measures said much for the benefits of conspicuous information presented within peripheral vision. The subjects used were regarded as representative of VDU users generally and possessed widely differing levels of keyboard experience. That there was only one difference exceeding the chance level is no doubt due to the wide between-subject variance found (i.e. between-subject F-ratio $= 13.144$; between-conditions F-ratio $= 5.524$). This wide range of abilities can also be observed in the large standard deviations depicted in Figures 5.8–5.10. Should more in the way of statistically significant differences have emerged, then there would have been a particularly strong case in favour of the conspicuity effect with a very broad spectrum of users.

There were also a wide range of completion times between subjects (e.g. 3.36 min to 25.12 min for the Low Conspicuity condition). If a screen condition can be regarded as generally beneficial, it should be beneficial for relative novices and experienced text editors alike; and a range of users were chosen as participants. Although mean completion times did not vary enormously across conditions, completion times were certainly diverse between subjects.

It is also worth commenting that whilst the pattern for pauses of five seconds or more before a mode change clearly descends in frequencies from the Low to the High Conspicuity conditions, no similar pattern emerged from pauses of five seconds at any other time, This strengthens the notion that overall performance differences (task completion times) rest on conspicuity differences.

What is apparent from the raw data is that the superiority of the high versus low salience screen is much more apparent for relative novices. A *post hoc* dichotomisation was made of subjects who had a relatively 'High' level of experience (more than two years typing or text editing) versus those falling into a 'Low' experience group (less than two years such experience). By subtracting the time to completion of the Low Conspicuity condition from that of the High Conspicuity display, performance time differences were obtainable for each of the eight subjects in the dichotomous groups. The mean of the Low experience group ($\bar{X} = 5.589$; $sd = 4.498$) was substantially greater than that of the High experience group ($\bar{X} = 1.459; sd = 4.340$). However, an independent *t*-test performed between these groups fell short of statistical significance ($t = 1.869; df = 14; p < 0.05$; two-tailed test). Briefly stated, the pattern suggests that inexperienced users seem to benefit most from the High Conspicuity condition although that trend may be due to chance alone.

Although no one error measure in itself could account for the overall performance (speed) superiority in the high salience condition, it is likely that these collectively (and possibly in addition to other specific measures not examined) could account for the difference. It was felt possible that performance in low salience conditions might be reflected in distinct pauses during typing keying immediately before a mode change as the user temporarily stopped and examined the mode indicator before pressing the relevant keys to change mode. It was possible to programme the HIMS to pick out pauses of greater than a set time just before a mode change. The expected trend was found in that there were more pauses of greater than three or five seconds (arbitrarily chosen criteria) before mode changes in the conditions with lower conspicuity, although the figures were not statistically significant. Similarly, there were progressively (but non-significantly) greater incidences of pauses (at all other times during the sessions) through from low to high salience conditions.

One or two criticisms may be levelled at this study. Firstly, as was mentioned of Experiment 5 (§4.11), counterbalancing the correct version on the right- versus left-hand screen side could have been included within the methodology. Secondly, the notion of conspicuity was a composite of several parameters including colour, size and position. As such, it was not the case—as should be so within a true experimental procedure—that only one variable was manipulated at a time. Also, 'meaningfulness' was an extra variable in that in the two lowest conspicuity conditions involving words rather than asterisks, subjects needed to read the information.

Subjective comments were revealing but also widely varied. Sixty-two percent of subjects said they preferred the High Conspicuity condition, twelve percent preferred the High/intermediate condition, whilst 25 percent preferred the Low Conspicuity condition. Of those who preferred the Low Conspicuity condition, one simply said that she preferred information to be written and that people grow accustomed to looking up at the screen on such tasks, anyway. Comments concerning the benefit of the High Conspicuity condition included: '...can pick asterisks up out of the corner of your eye better though...Don't have to stop and look'; 'Two rows of asterisks are better because there is a row nearer to where you are looking'; '...more accustomed to looking at the bottom of the screen...with Word Perfect'; '...couldn't help noticing stars...couldn't notice the words changing'; '...can associate asterisks with type-over...There are times when you are confused so something bold is the best one. Two lines of asterisks is most effective'; '... preferred yellow at top and bottom because instead of the eyes focusing on words to read to know whether you're in type-over or insert mode, you can concentrate on the text but still be visually aware of the yellow being there'; 'Short row of asterisks was no good... still had to look at top line to see what mode I was in'. One subject preferred the High/intermediate condition as she found the double asterisk line to be a distraction. Another individual felt that he relied most on memory of mode changes but added that this reliance was less so with the High Conspicuity condition. One subject (a biology student) phrased her opinion rather neatly: 'Two rows of asterisks are easier for me to see because they are there in the periphery of vision'.

5.5.1 Use of colour

One of the three parameters varied within the conspicuity dimension was colour. The contrast between the light yellow and dark blue was clearly much greater than between the maroon and blue. Due to precise quantification equipment not being available, it is not possible to state precise luminance contrasts. However, they might correspond with those used by Carter and Carter (1981) and which they term 'dark purplish red' and 'yellowish green'. The CIE Tristimulus X, Y, Z values for the former are 22.33, 8.54, and 10.75, respectively and for the latter are 27.64, 52.83, and 8.83, respectively. Luminance for the red was given as 4.711 cd/m^2 and for the yellowish green it was 6.843 cd/m^2. Research with PRESTEL screens suggests that luminance contrast between display items and background is perhaps the most important item affecting the legibility of a presentation (Sutherland, 1980). Generally, the greater are the contrasts, the more readable are the items on a VDU. Contrasts should not be so great, however, or employed so much, that they cause visual discomfort. An optimum luminance ratio of 8:1 to 10:1 for light characters on dark background and 1:8 to 1:12 for dark characters on a light background has been recommended (Ericsson, 1983). The symbol/background colour combinations in Table 5.2 have been suggested.

Character	Background	
	Use	Avoid
red	white, yellow, cyan	magenta, blue
yellow	magenta, red	white, cyan
green	yellow, white	cyan, blue
cyan	red, blue	green, yellow
blue	white, cyan, green	red
magenta	blue, white, cyan, green	red
white	magenta, red, green, blue	yellow

Table 5.2: Suggested character and background combinations for optimising screen presentations (after Bruce & Foster, 1984; Long, 1984).

5.5.2 Spatial position and conspicuity

A second aspect of conspicuity was spatial positioning of the information. Status lines are about evenly divided amongst text editors in terms of whether they appear at the top or bottom of the screen. From the basic results (Table 5.2) and from some of the subjective comments, it would seem that the condition in which mode information was displayed at the bottom of the screen (in addition to a top status line) was of particular benefit. This might be especially so for the relative novices who may have had more reliance on the keyboard than would a touch-typist. Obviously, people would be able to perceive information presented at a bottom status line more easily than at a top status line whilst they are generally looking at the keyboard. However, it is not suggested that this alone accounts for the superior performance differences of the Low experience group. It was not felt that the more experienced participants were in the touch-typist category, although there would likely be some degree of less need for checking exactly where keys were.

5.5.3 Size of status area

The third conspicuity factor involved was the area of screen actually covered by the status information. In Type-over mode this ranged from a block of seven characters to two complete rows of 80 characters each. Although it is intuitively obvious that the greater the area covered the greater is the likelihood of perceiving a change with peripheral vision, this must be weighed against the problems of unlimited screen space, visual clutter, etc. Considering that clarity (e.g. high contrast of colour against the background) is possibly more important than area covered, then perhaps a basic compromise suggestion would be to use something which is prominent on the bottom line but which would leave at least half of the row clear for icons and other items of information.

EXPERIMENT 7: STATUS INFORMATION AND EYE MOVEMENTS

5.6 INTRODUCTION

To strengthen the findings of Experiment 6, it was decided to gather incontrovertible evidence to support our working hypothesis that conspicuous mode information was actually reducing the need for saccades away from the place of attention (and the attendant need to relocate the point of current interest, etc.). A close analysis of eye movements was therefore required.

5.7 METHOD

Only two subjects were required to participate. The scleral search coil was again used. The procedure, viewing distance etc. were as described in Experiment 6. As the interest was purely in eye movements, the HIMS monitor was not involved. Presentation order was different for the subjects. Programme SAMDRK was used to collect and analyse the data. A 10 ms sampling rate was used.

5.8 RESULTS

Figures 5.11 and 5.12 show typical five second durations of eye movements on the Low conspicuity and High conspicuity screens respectively. Screen and keyboard areas are superposed upon the eye movements.

For every sampled interval, vertical (line number) and horizontal (column position) coordinates were obtained. From the calibration of the screen, it was possible to determine which range of coordinates corresponds to the areas of status information upon the screen. Thus, from the records it was possible to establish whenever the user looked at status information.

There were clear differences in terms of incidences of fixation on the mode information, with the two subjects (collectively) needing eye movements 22 times in the Low Conspicuity condition but only once in the High Conspicuity condition (see Figure 5.13).

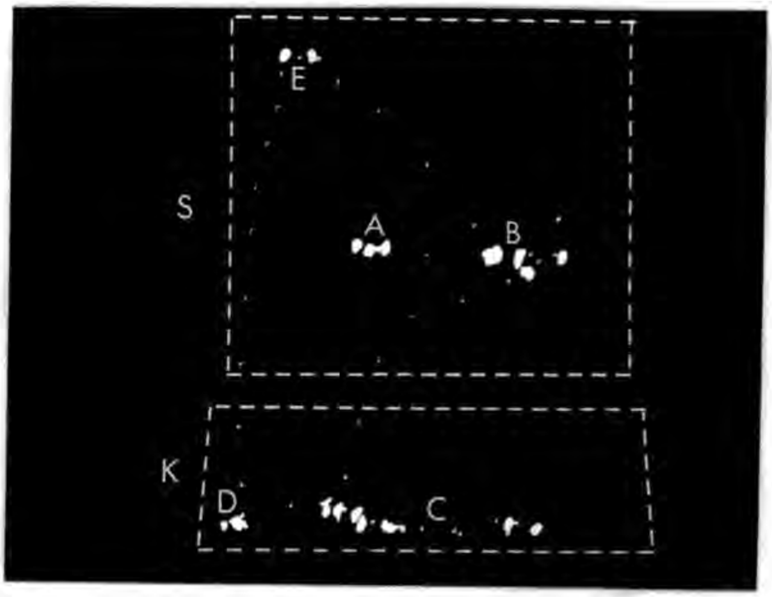


Figure 5.11: Typical pattern of eye movements from one subject (DS) over a five second period with mode information displayed in Low conspicuity (S = screen area, K = keyboard area).

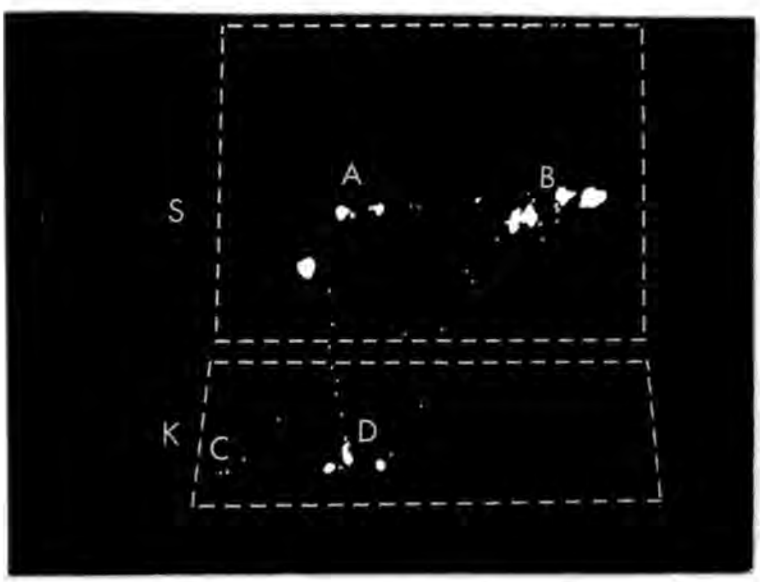


Figure 5.12: Typical pattern of eye movements from one subject (DS) over a five second period with mode information displayed in High conspicuity (S = screen area, K = keyboard area).

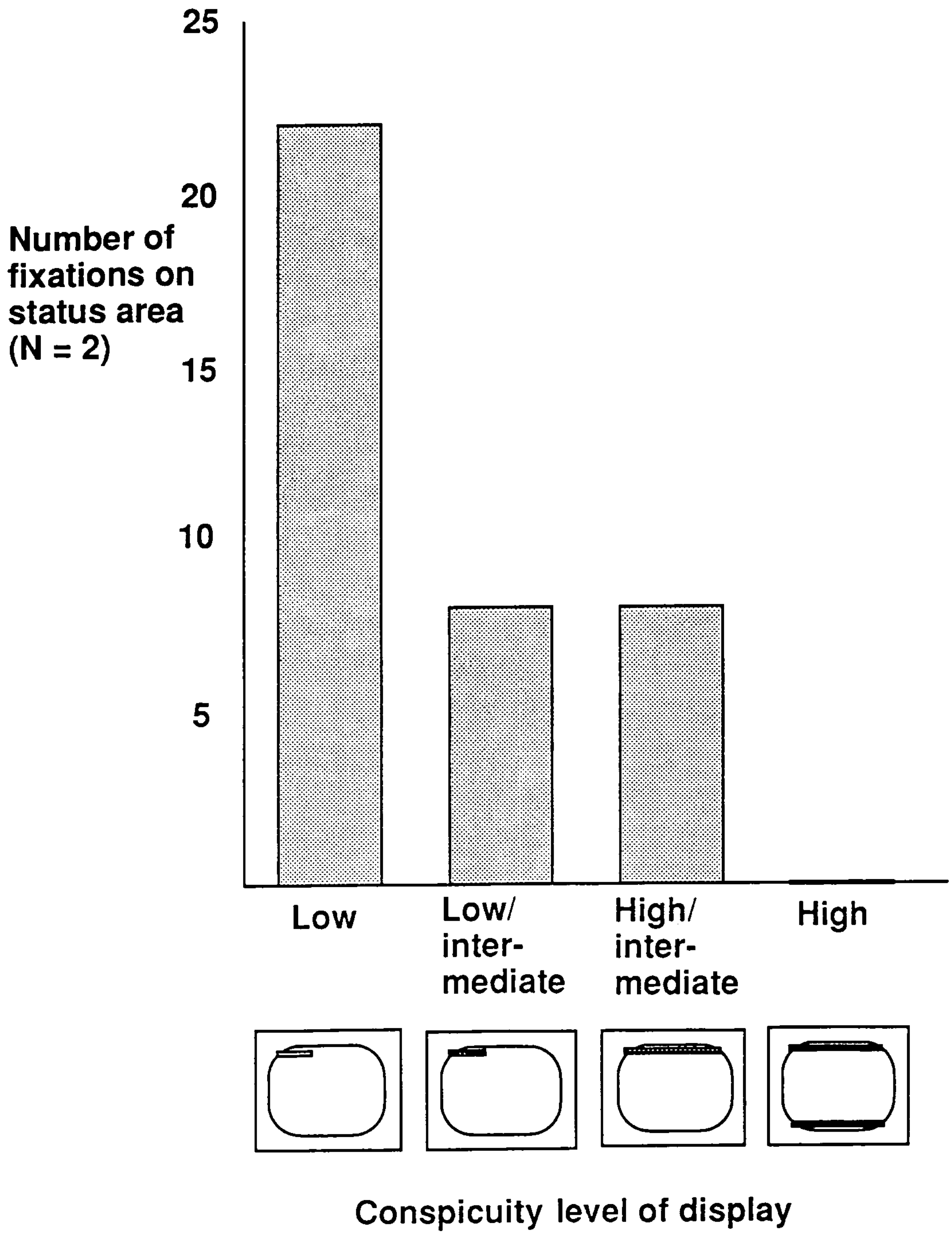


Figure 5.13: Histogram of incidence of fixations on status information across the four conditions (data combined from two subjects).

5.9 DISCUSSION AND CONCLUSIONS

Using the scleral search coil method of eye movement monitoring, it was possible to demonstrate precisely where an additional group of subjects (N=2) were directing their gaze about the screen. From this data, it was possible to determine how frequently subjects looked at the specific areas where mode information was displayed. For low to high salience conditions, frequencies of fixating upon the 'informative' screen areas diminished in an unexpectedly dramatic pattern but one entirely consistent with the hypothesised task analysis (22, 8, 8, and 0 occasions, respectively). The zero difference between the two intermediary conditions is understandable in terms of the very limited difference in conspicuity between these two (see Figs 5.4 and 5.5).

Figures 5.11 and 5.12 well illustrate the differences in eye movement patterns manifested over the Low and High Conspicuity conditions. In Figure 5.11 (Low Conspicuity), two masses of fixations (A and B) represent the subject working at about the mid-screen position and comparing the left-hand and right-hand sides of the text. The horizontal line of fixations at C represents the subject looking at the keyboard as he types. To the very left (D) there is a small area of fixation which is the position of the Control key (which is required for changing mode). The top area of fixations (E) represents the status area. Therefore, this figure clearly illustrates the subject (non-touch) typing at the keyboard, saccading back to the appropriate text position, reading and comparing versions, and deciding what needs to be done. Needing to change mode to insert or type-over text, he looks to the Control key and appropriate A or S key and changes mode.

In Figure 5.12 (High Conspicuity), the subject can be seen performing a similar change, but this time without recourse to consulting the status bar with a fixation. The subject reads and compares the two versions of prose (A and B), to detect via the large status bar in peripheral vision that the next correction requires him to change mode, saccades to the Control key (C) and keyboard (D), makes the necessary change and keystrokes for correction, and then saccades back to the screen looking at the next line of the correct version.

This study also demonstrates a further use of eye movement monitoring within visual search studies. In addition to providing precise data concerning saccade amplitudes, fixation durations, etc. (which were not required here), on this occasion the scleral coil technique served to substantiate a hypothesised task analysis of HCI.

Symbol blinking could also be a very effective status mode (or other) indicator, especially on areas of displays positioned at the periphery of the normal field of vision where the eye is most sensitive to movement and changes of intensity. Blink rates of

2–4 Hz are acceptable, but rates of 5–15 Hz should be avoided (Smith & Thomas, 1964). Up to five different blink rates can be distinguished (McCormick, 1964), but for coding purposes a maximum of two different rates is suggested (Long, 1984). This research area may prove amenable to the sort of peripheral presentation/eye movement monitoring approach described here.

In conclusion, it seems clear that the peripheral presentation of status information and other on-screen mode changes is at present grossly under-used, perhaps particularly for novices.

Chapter Six

Do Icons Facilitate Search?

6.1 OVERVIEW

Icons represent a fundamental aspect of the new wave of interfaces. For an increasing number of researchers, the initial enthusiasm is giving way to a more balanced view which asks *when and why* icons might be preferable to their counterparts. In this chapter, possible benefits of icons were examined for ease of locatability of information. The role of peripheral visual processing is of particular importance. Results of a visual search task showed icons to be more quickly located within an array than textual material. Precise scan paths were then recorded using a search coil follow-up study. Four methods of analysing the search paths (including transition matrices) suggest *greater systematicity* of search strategies for *complex* items.

6.2 INTRODUCTION

By way of definition, an icon is a small pictorial representation of some set of information in a system. This information is conveyed to users by drawing on their visual senses and knowledge of context. The high resolution of most WIMP system displays means that objects with which a user can interact can be displayed on the screen graphically.

Information represented on iconic interfaces may be of 'office-like' objects such as in-trays, out-trays, files, folders and wastepaper bins. Also, icons have been used to represent command operations, such as deleting, creating and printing files or sets of data. One reason for this increased popularity may stem from the fact that graphic symbols are often considered as a potentially universal means of communication which can convey certain types of information more directly and with more immediacy than can words. For instance, studies by Haber (e.g. 1970) which show that memory capacity for pictures may be unlimited suggest that information from visual images may also be among the most sensitive form of memorisable storage. However, this notion is questionable (e.g. Rogers, 1989*a*).

Baecker and Buxton (1987) raise the question: Is the icon a fundamental breakthrough heralding new dimensions in computer usability, or a fad, hiding loss of productivity behind a euphoric wave of apparent 'user friendliness'? Lodding (1983) argues the former, stressing that (a) because people find images 'natural', (b) because the human mind has powerful image memory and processing capabilities, (c) because icons can be easily learned and recognised, and (d) because images 'can possess more universality than text', iconic interfaces can 'reduce the learning curve in both time and effort, and facilitate user performance while reducing errors'. The advantages for international use of symbolic displays over those which are language-based are emphasised by Easterby (1970), whilst Gittens (1986) notes the ease with which graphical attributes of icons such as style and colour can be used to represent common properties of a collection of objects.

Manes (1985) argues the latter, asserting that icons may be confusing, wasteful of space, and totally ineffective in dealing with large numbers of similar commands, files or concepts. Gittens (1986) remarks on the difficulty of finding 'obvious pictographic equivalents' of computer system concepts, and of using icons to deal with the specification of large numbers of command parameters. Kolers (1969) remarks that the claims for 'immediacy' and 'directness' of the understanding of icon are exaggerated, and that recognising even realistic icons requires 'a great deal of perceptual learning, abstracting ability, and intelligence'. Although icons usually have a claimed benefit for novice users, some results suggest icon use as being more appropriate for experienced rather than novice users (Brems & Whitten, 1987). Echoing a theme of Carroll's (1982), those such as Hemenway (1982), Jervell and Olsen (1985), and Gittens (1986) emphasise the importance of the design of a set of interrelated icons to represent a sét of related commands, and of developing a metaphor to facilitate user comprehension of the system. Moreover, all express concern about the problems of overextending the metaphor, and of its breakdown when the logic of the metaphor leads to erroneous inferences about a computer system which the metaphor was intended to illuminate.

Baecker and Buxton (1987) further note that there is little experimental evidence documenting the advantages and disadvantages of icons and delimiting with precision when they can be used appropriately. Frequently cited literature (see, e.g. Kolers, 1969; Easterby, 1970) provides little more than valuable details of icon design principles. Hemenway (1982) comments that most of the stated advantages of icons are only speculations, with little empirical evidence to support them. Icons do, however, she acknowledges, have advantages because they are more visually distinctive than text, and because they can represent much information within a small space. So, although icons clearly are a valuable element of modern user interfaces, they are no panacea or cure for bad design.

6.2.1 Visual memory

One of the more obvious reasons why iconic-based interfaces may be easier to learn and remember is that users do not need to recall command labels to execute an operation. Rather, they need only recognise and select the appropriate icon from a set (whether menu, palette or other form of display). This has the effect of minimising the cognitive load on memory as compared with the requirements of needing to recall a name or abbreviated code. However, it could equally be argued that the same is true for labels or names that are visually displayed on the screen. Therefore, any advantage arising from the use of icons compared with other forms of display-based interfacing must be a result of the differences between the content of the displayed information. An important question follows of whether there is any difference between the amount of information which the user picks up and remembers from the different types of information displays and, moreover, how this affects the way in which a task is performed (Rogers, 1989*b*).

To summarise the present situation, it seems that for many computer users, the term 'icon' has become almost synonomous with 'good'. However, from the perspectives of cognitive ergonomics or vision science, serious doubts are cast on what is too often unquestionably hailed as a panacea. For example, it has been shown that in conditions where other factors are controlled, icons are no more memorable than equivalent verbal labels (Lansdale *et al.*, 1987). In what is possibly the only study to date which compares the locatability of icons versus words, Lansdale *et al.* (1989; discussed in detail in the next section) demonstrated that icons are the more easily searched for items. However, there are several explanations, of both cognitive and visual natures, for the basis of this. On the cognitive side, differences in the *articulatory distance* (the extent to which the link between what is being depicted and the underlying referent can be inferred, or, the gap between the meaning of an expression and its form) is just one. As the distance is less for icons, less time is required to transform an intention into an actual selection (Hutchins *et al.*, 1986). Alternatively, it may be the case that icons facilitate a *parallel visual search pattern* whilst verbal material encourages a serial search (Treisman & Southern, 1985; Muthig & Reinhardt, 1987). A further possibility involves the observed superiority of *global features* (e.g. shape or closure) versus local features (e.g. lines and structures within figures) in visual perception (Arend *et al.*, 1987). Finally, and in relation to some of the material outlined in §1.8, preattentive processes locate and isolate objects in the visual field. This report concentrates on the visual aspects of these differences, but leaves the notion of preattentive processes for further exploration by other researchers.

6.2.2 Visual search and icons

Many established views on icons are criticised by Lansdale *et al.* (1989). For instance, claims that icons are more readily understandable seem over-optimistic, with many studies reflecting the difficulties of creating icon sets which are comprehensible but not confusable (e.g. Bewley *et al.*, 1983). This is not particularly surprising as the notion that icons are helpful because of their linguistic neutrality may say no more than that they are equally obscure to everyone. It has been demonstrated that in conditions where other factors are controlled, icons are no more memorable than equivalent verbal labels (Lansdale *et al.*, 1987). Finally, studies are emerging in which performance at iconic interfaces are less successfully than with other styles, depending upon the task in hand and the particular implementations (e.g. Whiteside *et al.*, 1985). Thus, it is generally not possible to locate literature indicating that visual modes of communication do actually release additional cognitive resources. Where such a claim is made, it is usually because the circumstances differ between the verbal and iconic methods in favour of the latter. Nevertheless, one area where we can expect verbal and visual methods to differ is in the perception of visual material. Lansdale *et al.* (1989) have recently reported on two experiments undertaken in this area. In these, subjects were required to search a display for specified targets which were shown either as verbal labels or as icons. The speed and accuracy with which these targets could be located were measured.

Results clearly show that the iconic condition resulted in significantly faster search times than for the verbal condition with no significant increase in errors where the wrong target was hit. 'The simplest explanation for this increase in speed would seem to be that many of the icons are identifiable by the subjects in their peripheral vision. Consequently, the subjects are able to identify and fixate directly upon potential targets and search time is thereby reduced' (Lansdale *et al.*, 1989, p. 426). This interpretation is consistent both with what subjects informally reported and with the search strategies adopted by them. Subjects commonly noted that the iconic targets were readily identifiable in peripheral vision, but this was rarely so for the verbal targets.

This team summarise by saying that icons in these experiments confer a benefit in terms of speed of perception and scanning times.

> These initial experiments suggest that iconic displays result in more missed targets compared to verbal material which has been scanned serially... It is not difficult to imagine why iconic information should result in faster scanning times. Text involves a large number of lines closely spaced together. In other words it has a high spatial frequency which would be harder to resolve in peripheral vision. In these experiments, there was general agreement between

the subjects that about all that was resolvable in peripheral vision was the overall size and shape of the words, and it is interesting to note that subjects occasionally commented upon using these shapes when the target pair of labels produced a distinctive overall shape. The icons, on the other hand, were much more variable in terms of spatial frequencies, having broader as well as more detailed features. They also varied in terms of their overall density. There are therefore many more cues available in peripheral vision to the subject in the iconic condition. Whereas it was rarely possible in the verbal condition to identify a target in peripheral vision (thereby encouraging an orderly search strategy) it seems likely that on many occasions the iconic information available in peripheral vision was sufficient for recognition, hence directly accessing the target without a search. (Lansdale *et al.*, 1989, p. 428)

6.2.3 Aspects of icons

Despite the attention that has been paid to iconic interfaces, little work has been done on how to construct icons and iconic interfaces optimally. Such questions as 'Which visual properties will make an icon a "good" one?', and 'When, where, why, and in which respect should there be an advantage when information is portrayed by icons?', are largely unresolved (Korfhage & Korfhage, 1984). Sound empirical evidence is thus lacking as to whether icons will actually prove to be better in menu selection than word commands or text, as well as which characteristics of those icons improve menu selection (Arend *et al.*, 1987).

However, studies have been reported in which menu selection by icons and by word commands was investigated. In one study, menu selection for text-only menus was compared to menu selection where the text was supplemented by simple graphics. It was found that the addition of graphics reduced error rates considerably (about 50 percent) in comparison with the text-only condition but did not affect response times. To account for these results, Muter and Mayson (1986) suggested that graphical symbols provide more relevant information and allow access to meaning more efficiently. Although these results indicate that graphics may have a positive effect in menu selection, the study provides no evidence as to whether this was due to the combination of text and graphics, or to the visual characteristics of the simple illustrations used.

In another study, the effect of presentation mode (word commands versus icons) on menu selection was specifically addressed. In their first experiment, Wandmacher and Müller (1987) used a 'search-and-select' paradigm. On each trial, a task description was first presented to the subject and then either a randomised menu of items or a menu of word commands was displayed. Subjects were required to search for the correct menu item and to select it by pressing a corresponding number on the keyboard. A recognition paradigm was used in a second experiment. On each trail, only one item was displayed and had to be responded to by pressing a prelearned number on the keyboard. In contrast to Muter and Mayson, these authors found no effect of presentation mode on error rates but on reaction time. Wandmacher and Müller suggested that the articulatory distance is smaller for icons than for word commands. Therefore, for icons, less time is required to transform an intention into an actual selection. Both of these studies claim that icons improve menu selection because they provide a more direct access to the meaning of the commands involved. This does appear reasonable, but articulatory distance may not be the sole important factor in menu selection. When menus have to be searched for appropriate options, the search process may be assumed as also being heavily dependent on the stimulus characteristics of the items presented (Schwartz & Norman, 1986). However, neither the Muter and Mayson study nor the Wandmacher and Müller study explicitly controlled for the visual characteristics of the icons used for menu selection (Arend *et al.*, 1987).

A well-studied phenomenon in visual perception research which may be relevant is the observed superiority of global features in visual perception. Global superiority relates to the fact that global features of figures (e.g. shape, colour, size, closure, free ends of a line) can be selected and responded to much faster than local features (e.g. lines and structures within figures). Evidence comes from identification and search experiments, where visual similarity between figures was found to impair identification (e.g. Mohr, 1984). Furthermore, separable (global) features were found found not only to be searched very quickly, but in a time fairly independent of search size (e.g. Treisman & Southern, 1985). Arend *et al.* note that if the global superiority effect obtained in search and identification tasks also holds for more complex materials such as icons, then icons which differ from each other in global features should be searched and selected faster than icons which differ from each other in local features. Maximising icon distinctiveness by the use of global features, however, will enlarge the articulatory distance between an icon and its referent, since local features are required to make an icon more representational. Icons containing distinct global features will thus be necessarily more abstract and will provide only weaker retrieval cues than icons in which the respective options are portrayed primarily by means of local features.

In order to isolate the relative effects of icon distinctiveness and icon representativeness on menu selection, Arend *et al.* (1987) designed an experiment in which two icon sets (one constructed to differ with respect to global features, and one with respect to local features) were compared with a word command set. It was reasoned that if the global superiority effect also held for icons (which are more complex figures than the ones typically used in visual perception studies), then search times for icons differing from each other in global features should be much faster than search times for word commands (owing to a smaller articulatory distance and the global feature precedence). Concerning icons differing from each other in local features, the presence of local features can be expected to slow down the speed of the search (according to global feature precedence). Search time should therefore be longer than for icons differing in global features. Conversely, a smaller articulatory distance and a higher degree of representativeness can be obtained for icons by the use of local features than by the use of global features. Icons differing in local features can therefore be expected to provide stronger retrieval cues and a more direct access to the meaning of the commands they portray. In turn, this can be expected to enhance the identification process, especially where larger menus need to be searched. In order to obtain data on which this topic could be resolved and to obtain reaction time functions for each of the command sets used, Arend *et al.*'s experiment was designed to incorporate menus of different size. For icons differing in global features the slope of the reaction time function was anticipated as small, since parallel search was expected. On the other hand, word commands would be expected to be searched serially.

The search-and-select paradigm (Wandmacher & Müller, 1987) was again used to investigate which visual characteristics if icons are relevant for menu selection. Two sets of icons (abstract icons and representational icons) were compared to a word command set. For abstract icons, global features were employed to maximise their visual distinctiveness. For the representational icons, local features were employed to ensure a high degree of representativeness and a small articulatory distance.

The results showed that when menu options are presented in different modes, the search-and-select process is altered considerably. Options depicted as word commands can be searched and selected only very slowly. Additionally, and consistent with results of other studies (e.g. Perlman, 1984; Muthig & Reinhardt, 1987), selection time for word menu options was found to be heavily dependent on menu size, which would indicate a serial search for this sort of material.

However, Arend *et al.* also remark that simply depicting menu options as icons does not automatically improve speed and accuracy of the search-and-select process. With representational icons varying in local features, it was not generally found that menu selection was more efficient than with word commands. Nevertheless, representational icons seemed to be more efficient than word commands for menu selection where menus exceeded a critical size of about ten simultaneously presented options, or at least they were if some of the icons contained more salient (global) features.

Therefore, the advantage usually claimed for representational icons of greater similarity to their respective referents and smaller articulatory distance appears to be overridden by the greater mental effort needed to search for the required option on a local feature base. On the other hand, representational icons appear to offer some advantage if menu selection is not completely dependent on visual search processes, but also requires memory search processes, for larger menus. This is consistent with the claim that representational icons provide stronger retrieval cues with respect to their respective referents.

The results obtained for the abstract icon set do, however, query the assumption that icons should be made representational, even when options need to be chosen from large menus. Although the abstract icons were designed to be maximally dissimilar to each other (according to different global features) and to provide only weak retrieval cues with respect to the options they portray, it was found that they were far superior with respect to both speed and accuracy of the search process. As search times for abstract icons were found to be relatively independent of menu size, it was concluded that icons containing global features can be searched in parallel. Furthermore, abstract icons were also not more susceptible to errors than representational icons, which indicates that visual distinctiveness may prove to be an even more efficient retrieval cue than functional representativeness.

Arend *et al.* regarded their results as generally indicating that basic research in visual search can be successfully applied to the design of iconic interfaces. They highlight at least two factors which must be separated when an advantage of icons over word commands is claimed for menu selection. Firstly, icons are graphical or pictorial signs which are related to their respective referents by way of a similar relation that holds for at least one of their constituents. Because of this, it is possible that icons have a smaller articulatory distance to their respective meanings than a set of word commands. This might account for why icons can generally be identified faster than word commands (e.g. Wandmacher & Müller, 1987) and why there are less prone to error (e.g. Muter & Mayson, 1986). However, at least for icons appearing on menu-driven interfaces, 'visual quality' or 'item distinctiveness' (Schwartz & Norman, 1986) may be of more importance. Arend's results provide evidence that 'visual quality' can be made explicit in terms of global/local feature theory, from which empirically testable predictions for the speed of the selection process can be derived. The considerable advantage found for icons containing global features suggests that even greater articulatory distances can be overridden by greater distinctiveness of the respective icons.

However, as Arend *et al.* conclude, it might be argued that the smaller articulatory distance of representational icons will pay off in online computer sessions, since these icons provide more information about the scope, the action components involved, and the consequences of the invoked menu options. Therefore, representational icons should score better in learnability, usability, and understandability of application systems than abstract icons. This idea could be tested by implementing the command sets used by Arend *et al.* on a prototype editor and testing them in online editing studies. Arend *et al.* do suppose, nevertheless, given the large performance differences found in their experiment, that an advantage for the abstract icon set would also be apparent in real editing tasks.

6.2.4 Experimental rationale

The reasoning behind Experiment 8 was two-fold. Firstly, it was designed to further test the notion of peripheral processing accounting for the icon locatability supremecy. Secondly, there was the intention to develop an experimental design which would be suitable for using eye movement recordings to test for the use of peripheral vision (e.g. items detected with only a single fixation).

6.2.5 The pilot study (Stone, 1990)

In a pilot study, Stone (1990) devised sets of 20 'simple' words (four letters long), 20 'complex' words (six letters long), 20 'simple' icons and 20 corresponding 'complex' icons (the latter essentially containing more detail than the former). To test this distinction, two individuals were shown each pair and asked to indicate the more complex of the pair. The agreement was total. Slides were prepared containing either 20 icons (mixed simple and complex) or 20 words (mixed simple and complex) arranged around the periphery and a central referent item which was either an icon (with peripheral icons) or word (with peripheral words). The central item was identical to one of the 20 surrounding items (the target). Basically, the study involved showing these slides to 24 subjects and measuring search times.

Mean search times (seconds) for each type of presentation was as follows: simple words = 2.508 (sd = 0.404); complex words = 2.964 (0.564); simple icons = 1.936 (0.383); and complex icons = 2.293 (0.480). Mean response times for icons (2.115; sd = 0.407) proved to be significantly faster ($p < 0.00001$) than for words (2.736; sd = 0.424). Mean response time for simple items (2.222; sd = 0.358) proved to be significantly faster ($p < 0.00001$) than for complex items (2.629; sd = 0.441). There was also a small interaction between the two variables.

EXPERIMENT 8: A VISUAL SEARCH COMPARISON OF ICONS AND WORDS

Due to the multivariate nature of icons, it is not an easy task to either choose a variable for experimental manipulation or to tightly control other variables. Previously chosen experimental variables include the abstract/concrete distinction (Stammers *et al.*, 1989), high/low imagery (Rogers & Oborne, 1987), and memorability (Lansdale, 1988). A hitherto unresearched variable within icon studies, the simple versus complex dimension, was chosen. It is acknowledged that this distinction must largely be an arbitrary one.

It was hypothesised that search times for icons would be significantly faster than for words. Again due to the possible peripheral processing/spatial frequency advantage, it was also hypothesised that simple items (whether icons or words) would be significantly more quickly located than complex items.

6.3 METHOD

Twenty-four (7M:17F) subjects (mean age = 26.3; sd = 9.2) participated. Materials were largely as those used by Stone (1990) and consisted of 80 slides containing a central referent icon or word surrounded by 20 peripheral icons or words. The central item in each case was chosen from a set of either: (a) 20 'simple' words (words of four (upper case) characters) (SW; see Fig. 6.1); (b) 20 'complex' words (six characters) (CW; see Fig. 6.2); (c) 20 'simple' icons (icons of relatively simple visual structure) (SI; see Fig. 6.3); or (d) 20 icons of relatively complex pattern (CI; see Fig. 6.4). In each case, all items came from the same set. Four- and six-letter words were printed in different font sizes so as to occupy the same overall area. There was an equivalent simple and complex version of each icon. All icons were derived from icons actually in computer usage. Only one peripheral item was identical to the central item. This was the target. The position of target items was balanced across all displays so as to occur only once in each of 20 possible positions.

The materials differed in three ways from those of Stone (1990). Firstly, simple *and* complex words or simple *and* complex icons did not appear in the same array. Arrays consisted entirely of one category of item. Secondly, as many of the errors of identification had occurred between the paintbrush and the pencil icon, the paintbrush was modified so as to reduce the visual similarity. Thirdly, Stone's positioning of items around the periphery had involved more items along the horizontal axis than along the vertical axis, thereby creating a rectangular rather than a square array. Consequently, targets on the horizontal sides would be situated slightly further away from the central referent item than would their counterparts on the vertical sides and, although any differential effects on search times were hopefully counterbalanced out, it is best to avoid this discrepancy which could affect search times.

Presentation order was randomised between subjects. Slides were presented via a projector linked to an electronic timer and morse key. Subjects were simply required to locate the target item and press the key when they had done so. To verify their success, subjects were required to verbally indicate the target's position (e.g. 'Ten o' clock' or 'top left' or 'north west').

Within the pilot study less than two percent of all trials had produced an error and within half of these errors, as noted above, the confusion was between two particular icons. Discounting these two items, errors were not disproportionately spread across the representations. Because of this low error rate and because the troublesome similarity had been reduced, it was not felt that a speed/accuracy tradeoff would be present and hence could not influence results. Therefore, it was decided that in Experiment 8 no particular treatment of possible errors by omission or otherwise was warranted.

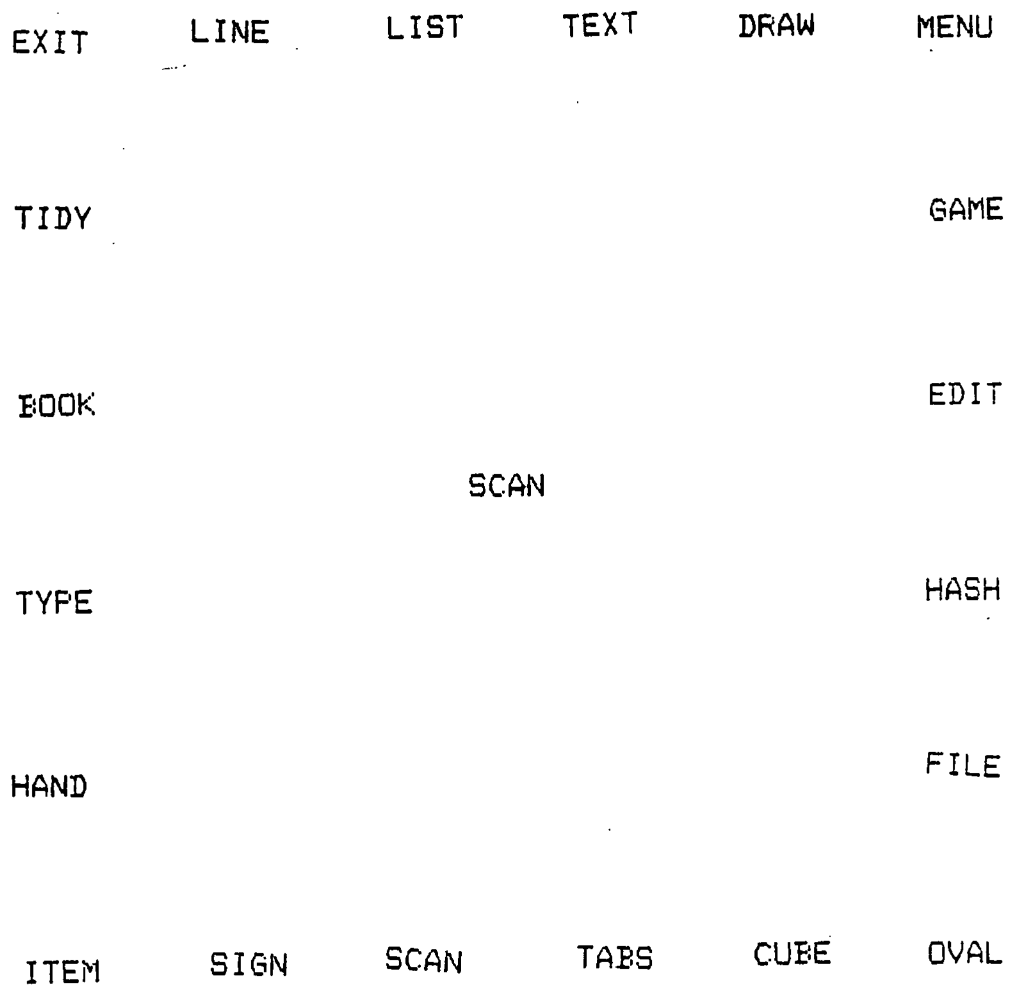


Figure 6.1: Example of a simple word presentation.

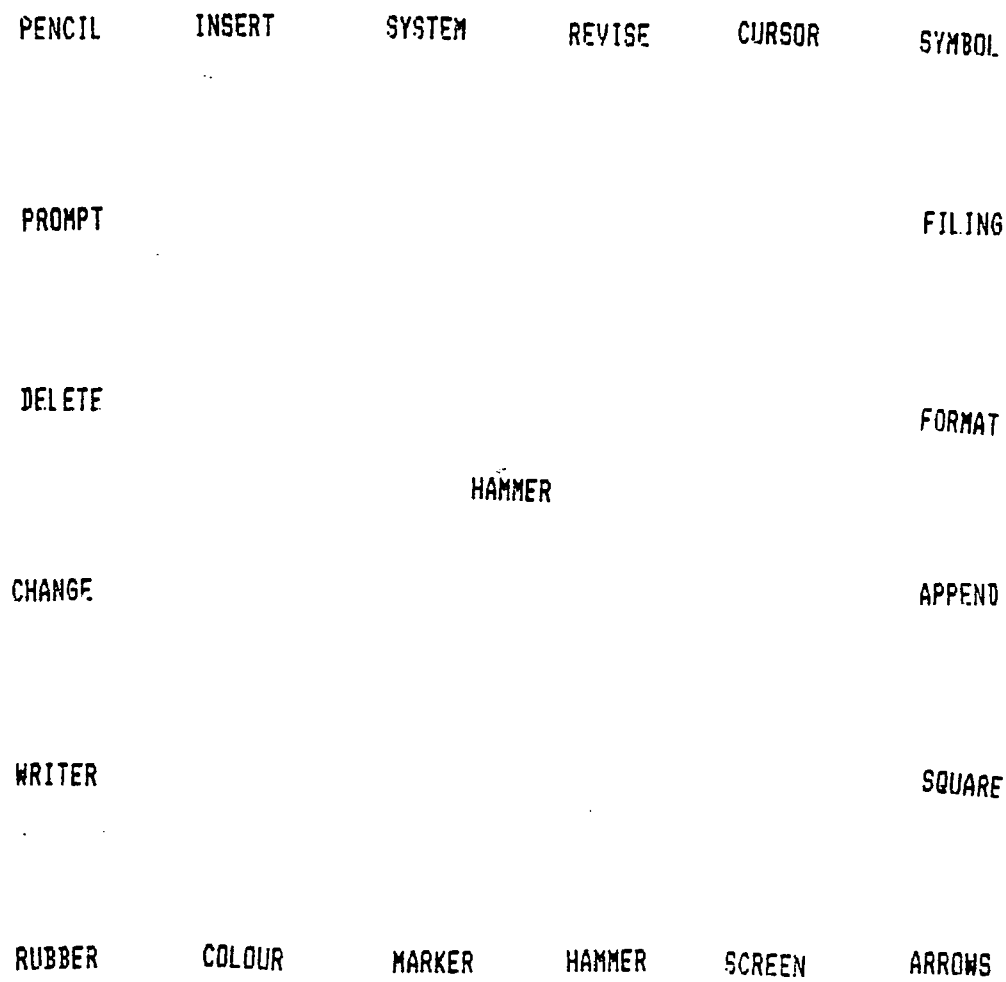


Figure 6.2: Example of a complex word presentation.

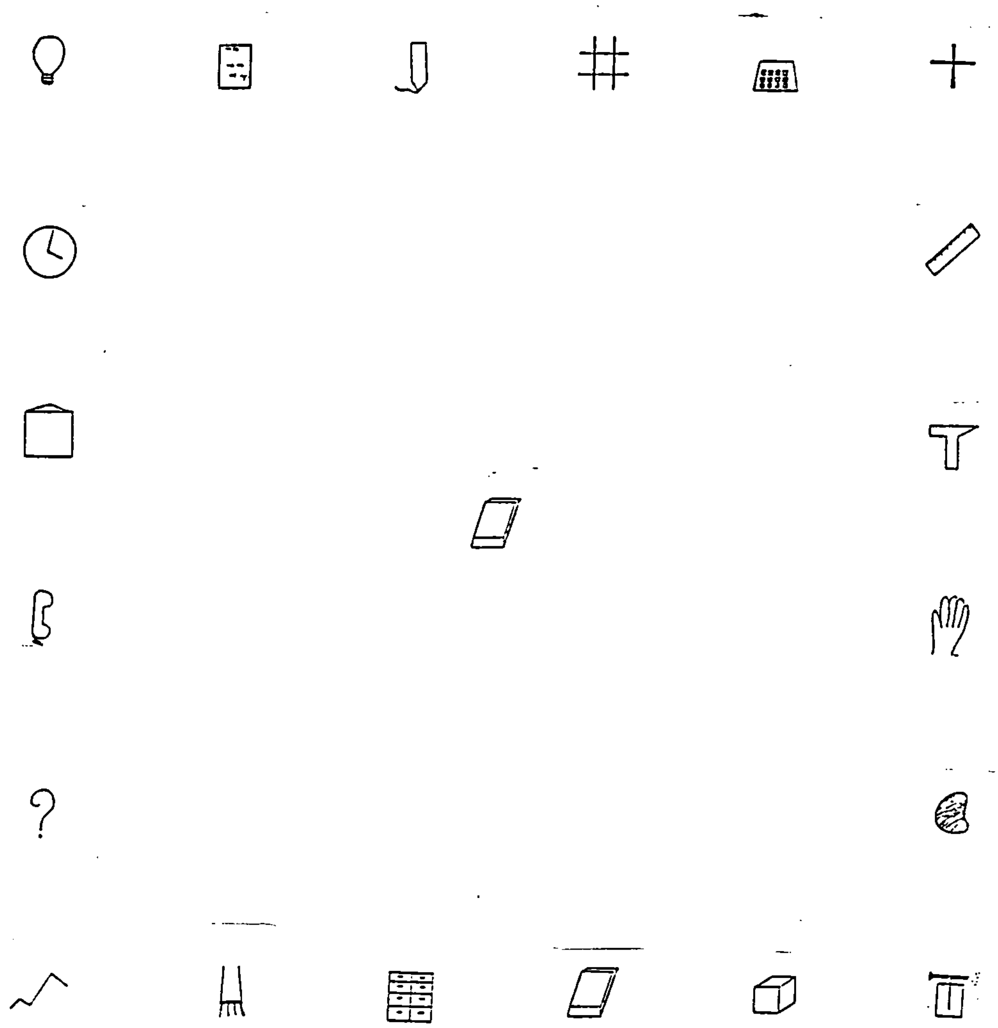


Figure 6.3: Example of a simple icon presentation.

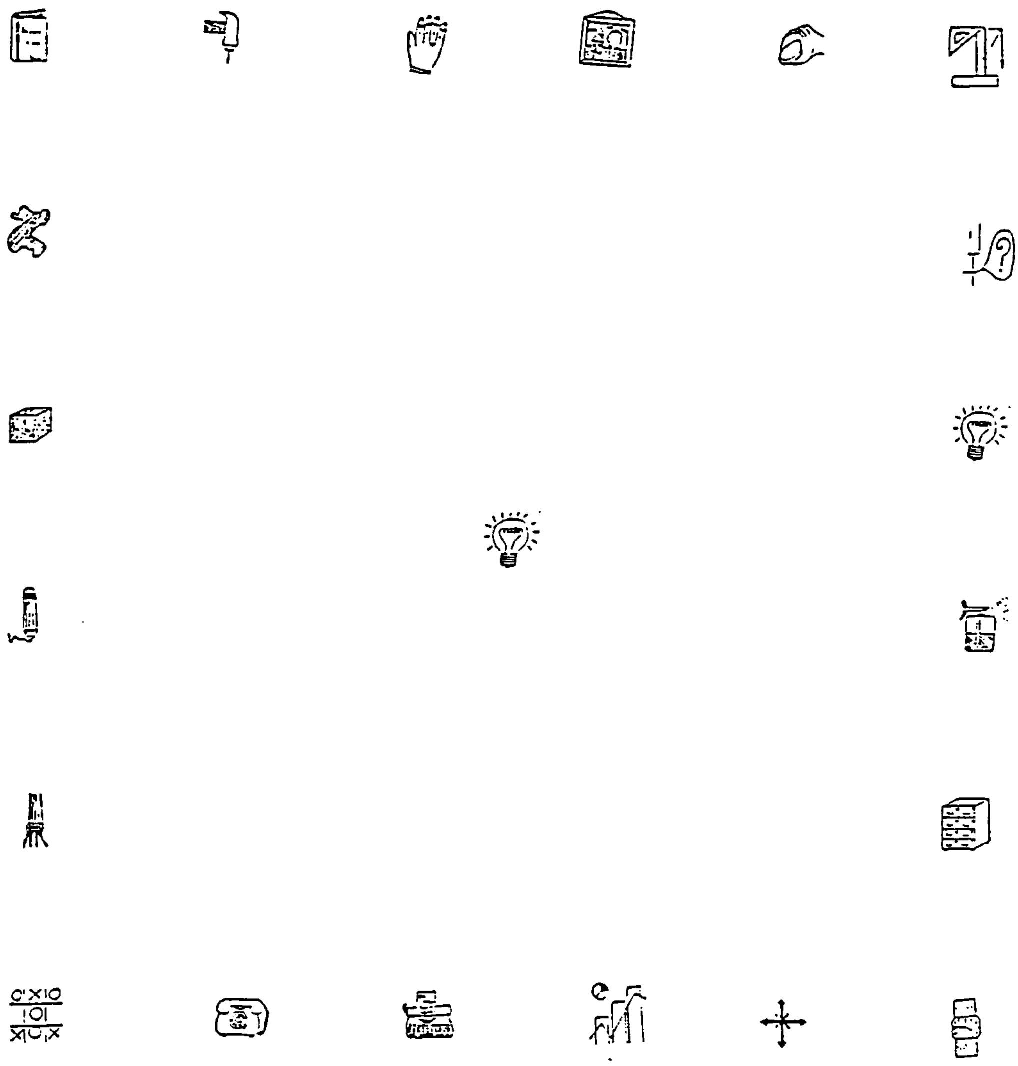


Figure 6.4: Example of a complex icon presentation.

6.4 RESULTS

Mean search times for the four target types employed are shown in Table 6.1.

	Icons	Words	Mean
Simple	2.035 (0.496)	2.821 (0.388)	2.428
Complex	2.739 (0.654)	3.171 (0.293)	2.955
Mean	2.387	2.996	

Table 6.1: Mean search times (s) of the four target types employed (standard deviations in parentheses).

It was found that mean search times for icons was faster ($F(1,23) = 33.033$; $p < 0.001$) than for words, and that search times were faster ($F(1,23) = 65.457; p < 0.001$) for simple rather than complex targets (see Table 6.2). There was also a significant interaction between these variables ($F(1,23) = 7.078; p < 0.05$).

Factor	SS	MS	F	df	p
Complexity	133.183	133.183	65.457	1	0.001
Type	178.059	178.059	33.033	1	0.001
Trial	121.811	6.411	3.262	19	0.001
C x Type	14.939	14.939	7.078	1	0.05
C x Trial	129.481	6.815	3.524	19	0.001
Type x Trial	121.383	6.389	2.796	19	0.001
C x T x T	149.184	7.852	3.800	19	0.001

Table 6.2: Results from the three-way ANOVA.

The interaction between the two variables is shown in Figure 6.5.

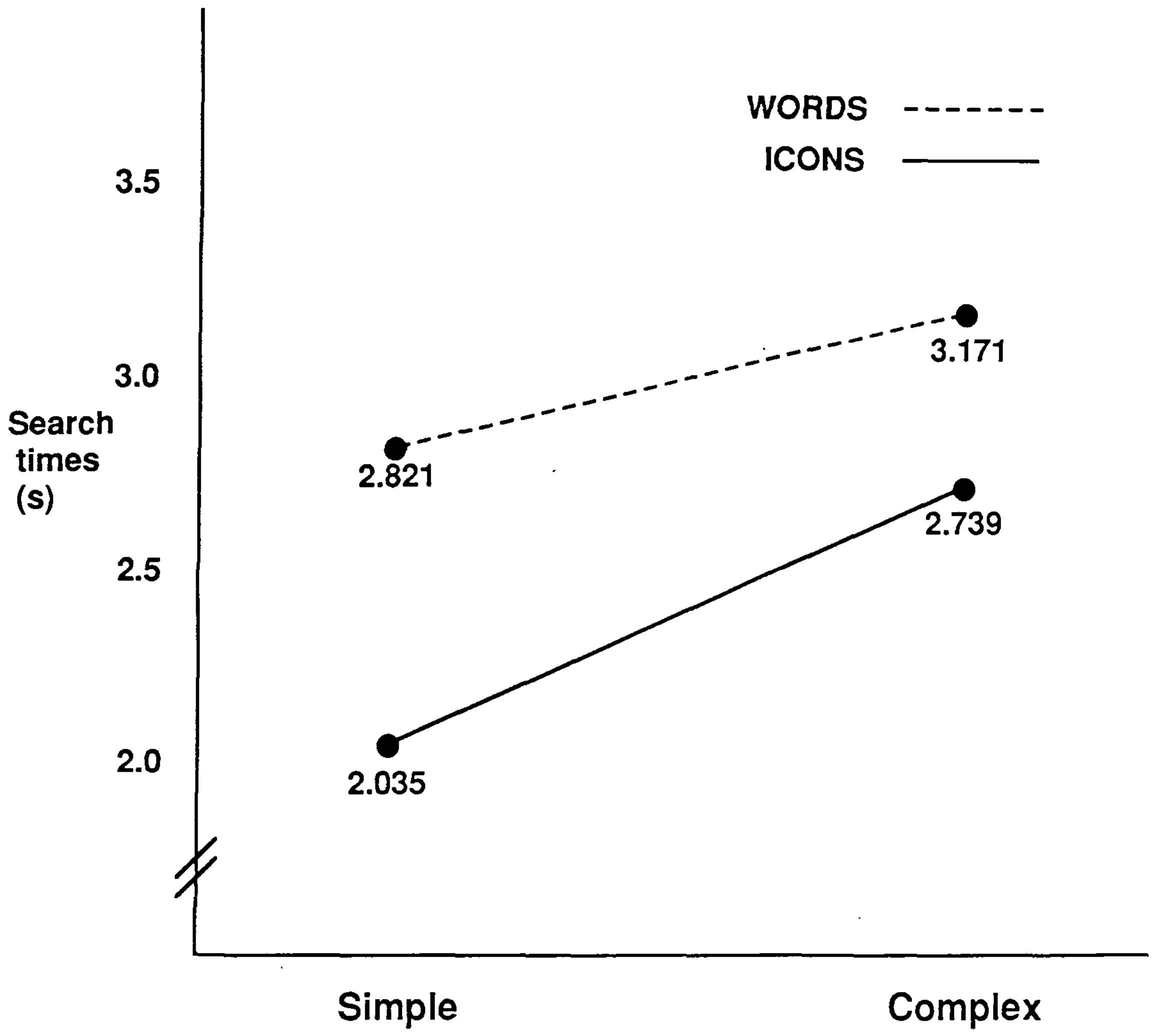


Figure 6.5: The interaction between the simple/complex and words/icons variables.

6.5 DISCUSSION

6.5.1 Subjective comments

When asked to choose whether icons or verbal labels were the easier to locate, seven of the 24 participants (29%) said that words were the easier. A number of subjects also noted that the 'cross' icon was probably the easiest item to find.

Other comments elicited included: 'Things with more shape are easier'; 'If I don't see it the first time, then I'm lost... those in the corners are difficult' 'I'm left handed so I hate anything on the right... icons are easier... they have a different outline'; 'Easier to keep words in mind'; 'Words slightly easier, but it depends on the shape of the icon'; 'Icons are easier... You have to read the words'; 'Icons are easier... less cognitive associations'; 'Cognitive factors do come into it'; 'I omit the corners when I'm scanning—then I have to go back'; 'I found it more difficult when in corners because your eyes miss corners'; 'Surely everyone starts at the top left and reads across or around'; 'I give a quick scan first and then search properly'.

6.5.2 General conclusions

It is worth noting that words were displayed entirely in upper case, and therefore *word shape* provided no visual information. That might not be the situation in dealing with some commercially available icon-driven systems.

The main conclusion from this study was that search times were greater for the word condition than for the icon condition, and greater when the item was complex than when it was simple. Both experimental hypotheses were therefore supported and these findings substantiate those of the obtained within the pilot study. These data also confirmed Lansdale *et al.*'s (1989) findings and extended the pattern to two (albeit rather arbitrary) levels of complexity. The principle explanation offered for these results assumes differences in peripheral visual processing. When the item was simple or when it was an icon, the target was more easily resolved in peripheral vision and hence performance is enhanced. One reason offered for this involves spatial frequencies. Text involves a large number of lines more closely spaced together; i.e. it has a higher spatial frequency which would be harder to resolve in peripheral vision (Lansdale *et al.*, 1989; Wilson *et al.*, 1990). It seems likely that on many occasions the iconic information available in peripheral vision was sufficient for recognition, hence directly accessed without a search.

In addition, these variables showed a significant interaction suggesting that both complex and simple representations were affected by whether the representations were words or icons.

EXPERIMENT 9: THE ROLE OF PERIPHERAL VISION IN THE ICON EASE OF LOCATABILITY PHENOMENON

6.6 INTRODUCTION

To investigate the basis for the differences observed within Experiment 8, eye movement records using the scleral search coil method were taken. Data were collected from a further six subjects. It was then possible to examine data such as fixation duration and scan paths.

6.7 METHOD

Six subjects participated. Other details are as in Experiment 8. The same four sets of stimuli were used.

6.8 RESULTS

6.8.1 Search paths

Figures 6.6 to 6.8 show examples of search paths from one particular subject (FN). Figure 6.6 shows a systematic, clockwise search with the target located at about 2 o'clock.

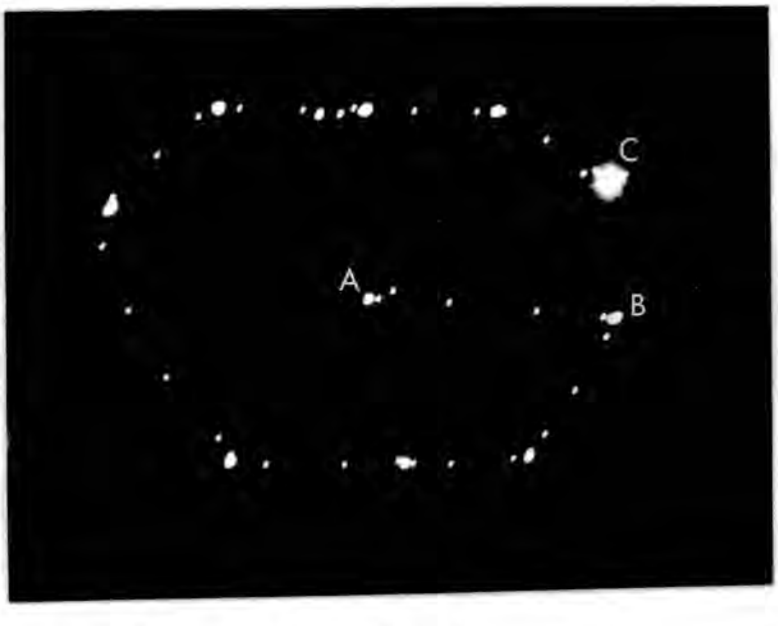


Figure 6.6: A systematic search path.

Figure 6.7 shows a fairly systematic search which demonstrates a 'miss' at three o' clock.

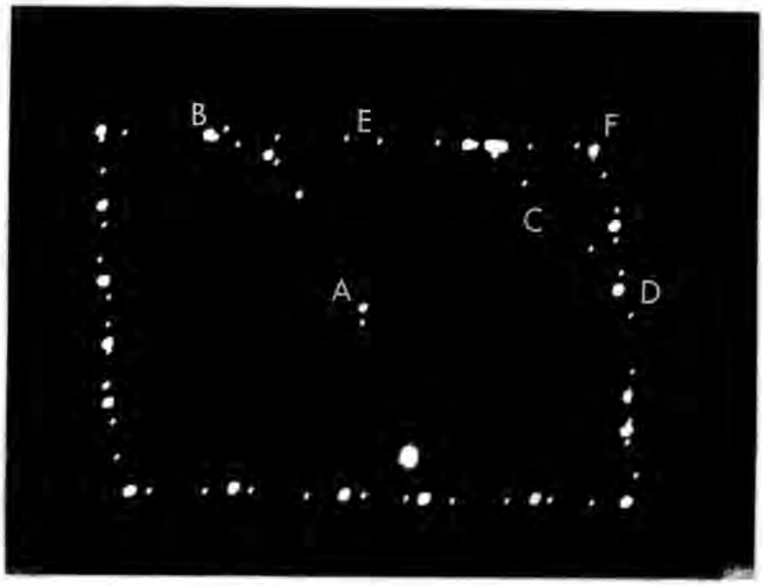


Figure 6.7: A search path demonstrating a 'miss'.

Figure 6.8 shows an example of an unsystematic search. These three search paths are described further in §6.9.1.

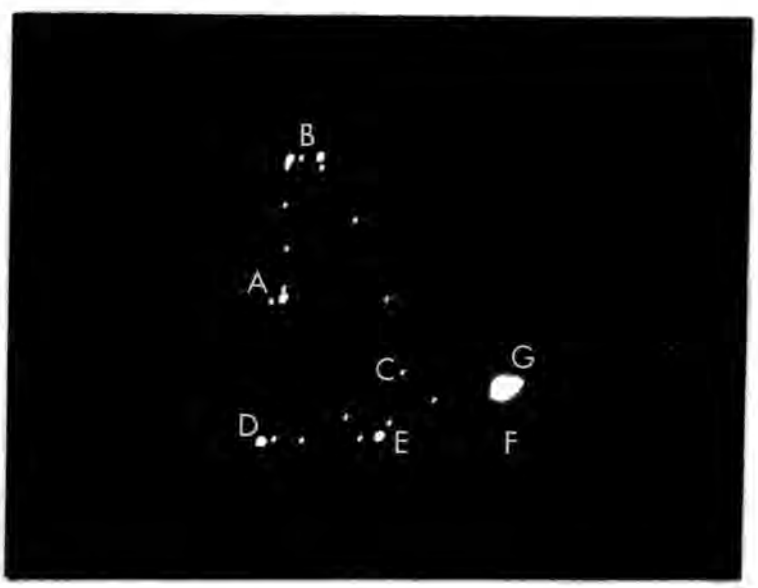


Figure 6.8: An example of an unsystematic search path.

6.8.2 Fixation duration

Table 6.3 shows the pattern of fixation durations between the four conditions. These results are obtained by summating all (several hundred) fixation durations recorded for all six subjects. Data files were manually edited to remove lines containing initial fixations on the referent item and 'sector information' automatically detailed by the programme, and then combined into one file for $SPSS_X$ analysis. Fixation durations were longest for CI ($\bar{X}$ = 225.6) and shortest for SW ($\bar{X}$ = 204.4). Raw data including mean, median and mode fixation results from individual subjects are included as Appendix 6.2. Although the complex icons had longer fixation durations than the three other presentation forms used, a one-way ANOVA did not demonstrate the differences to be statistically significant ($F = 2.045; df = 3$). However, four measures employed suggested differences in eye movement/search efficiency between the different display types.

	Mean	sd	Median	Mode
Simple icons	219.2	104.2	200	170
Complex icons	225.6	122.2	200	170
Simple words	204.4	66.1	190	180
Complex words	206.8	88.8	200	170

Table 6.3: Collective data (ms) on fixation durations from all intervals sampled for all six subjects.

6.8.3 Saccade amplitudes

Table 6.4 shows the means of saccade amplitudes (arbitrary scaling) recorded from the six subjects over the four conditions. Scale units are arbitrary units as defined when setting up the programme parameters as the icon/word items were not measurable by the likes of alphanumeric character units. Saccade amplitudes were greater for icons than for words, and greater for simple items than for complex items, although these differences were not statistically significant.

	Icons	Words
Simple	7.450 (1.887)	6.067 (1.521)
Complex	6.600 (2.020)	5.633 (1.556)

Table 6.4: Means of saccade amplitudes (arbitrary units; standard deviations in parentheses) for the four target types.

Figure 6.9 shows mean saccade amplitudes for the six subjects on the four conditions.

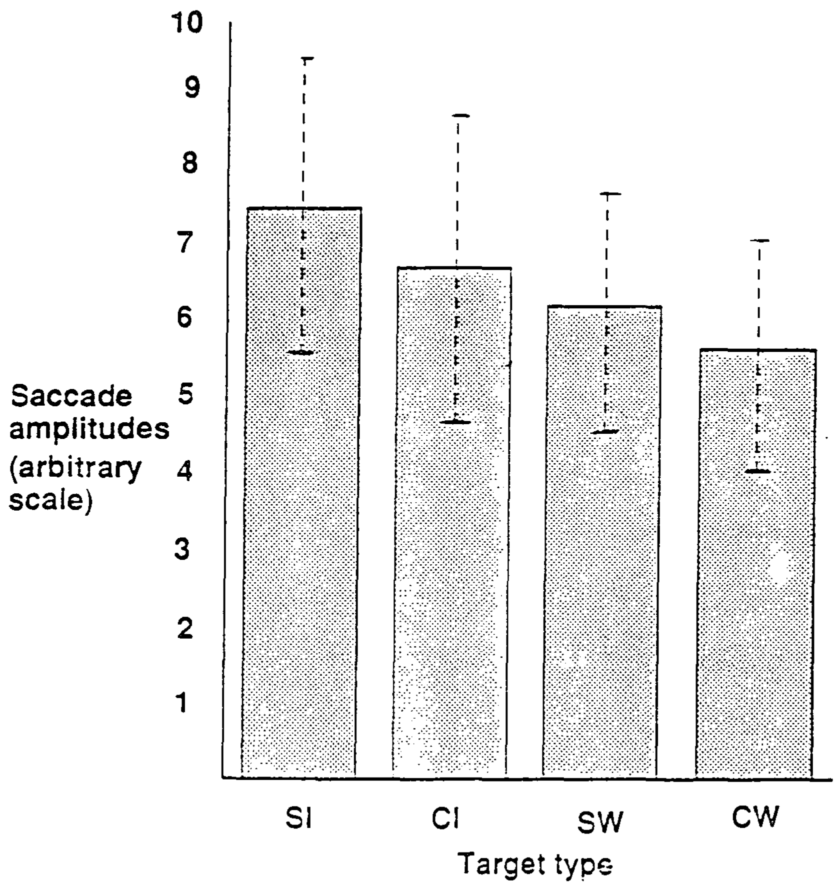


Figure 6.9: Histogram depicting mean saccade amplitudes (arbitrary units) for the six subjects (vertical lines indicate standard deviations).

6.8.4 First saccades to target

Looking at instances where the first saccade made successfully located the target (see Table 6.5), there was a highly significant greater frequency ($\chi^2 = 7.258; p < 0.01$) of saccades successfully fixating on icon targets (23) as opposed to word targets (8). There was also a significantly greater frequency ($\chi^2 = 3.903; p < 0.05$) of first fixations being on simple targets (21) versus complex targets (10).

	Icons	Words	Total
Simple	15	6	21
Complex	8	2	10
Total	23	8	31

Table 6.5: Breakdown of first saccades to target.

Again looking at cases where the first saccade made successfully located the target (see Table 6.6), there were less instances of items being located in the corner than elsewhere, although this difference was not significant ($\chi^2 = 2.502$; NS).

	Corner	Other
Actual	3.0	28.0
Expected	(6.2)	(24.8)

Table 6.6: Instances of first saccades made to corner and non-corner locations.

6.8.5 Items fixated before target

Secondly, there was a measure of number of icons or words fixated before the target was located (see Table 6.7). This did not include the initial (central) fixation unless, as often did occur, it was 'checked' during the search perhaps to compare it with a suspected target. Some of these fixations on non-target items were repeated a second (or third) time round as the subject missed the target and re-searched the array again. These could be calculated with or without multiple fixations (instances where two or more successive fixations were recorded on the same general area). Differences between the display types were highly significant ($F\ (3,15) = 13.275; p < 0.001$) including multiple fixations and excluding them ($F(3,15) = 4.383; p < 0.05$). Tukey's

WSD analyses revealed the highly significant differences ($p < 0.01$) to lie between the SI and CW, and between the SI and CI comparisons, whilst the difference between the SI and SW targets was also significant ($p < 0.05$). Furthermore, there was a highly significant ($p < 0.01$) difference between subjects in the analysis where multiple fixations were included.

	With mult. fixs.		Without mult. fixs.	
Target type	$\bar{X}$	*sd*	$\bar{X}$	*sd*
Simple icons	5.6	1.5	4.6	0.9
Complex icons	10.1	1.7	7.2	1.0
Simple words	8.4	2.0	6.1	1.0
Complex words	10.6	3.0	7.2	1.2

Table 6.7: Number of different positions fixated before the target was detected within the four conditions.

In the former analysis, 45 percent of the variance was accounted for by target type and 25 percent of the variance by subject differences. In the latter, stricter, analysis, 32 percent of the variance was accounted for by the target type (ω^2; Vaughan & Corballis, 1969).

6.8.6 Misses

Another measure of differences in search efficiency was termed 'misses'. This was where eye movement recordings determined that subjects had, in fact, fixated on the target, yet apparently failed to perceive it. Instances where two or more successive fixations were recorded on the same general area (multiple fixations) were not included. Table 6.8 shows results where side targets and corner targets were analysed seperately.

Target type	Sides	Corners	Total
Simple icons	21	4	25
Complex icons	38	16	54
Simple words	26	15	41
Complex words	42	19	61
Total	127	54	181

Table 6.8: Frequencies of 'misses' between the target types.

There was no significant difference in frequencies of misses between icons and words ($\chi^2 = 2.923$). However, there were highly significant frequency differences between complex targets and simple targets ($\chi^2 = 13.265; p < 0.001$), between corner and side targets ($\chi^2 = 10.940$; $p < 0.001$), between level of complexity and target type (icons or words) ($\chi^2 = 16.635; p < 0.001$), and between level of complexity and position of word targets ($\chi^2 = 15.255; p < 0.001$) and of icon targets ($\chi^2 = 13.864; p < 0.001$). Table 6.9 shows actual and expected frequencies of misses with corner and other target locations.

	Corner	Other
Actual	127	54
Expected	(144.8)	(36.2)

Table 6.9: Actual and expected frequencies of misses with corner and other target locations.

6.8.7 Transition matrices

Fourthly, transition matrices of successive saccades were produced from these data. Saccade direction was analysed by computer programme into one of eight 'compass points' from the source. It was possible to determine the frequency of change in direction from one saccade to the next. It was hypothesised that, due to the perpendicular arrangement of the arrays, a more systematic search would largely consist of vertical or horizontal saccades with the occasional change through 90^o when scanning the array's corners. On the other hand, an unsystematic search strategy would be expected to show relatively more oblique changes in saccade directions. The frequencies of oblique direction changes depicted in Table 6.10 are expressed as a ratio of one. The higher the ratio, the greater is the percentage of oblique saccades and the less systematic the search is held to be. In this table, '$\bar{X}$' refers to the mean of the mean of the six subjects, whilst 'combined' refers to the mean of each saccade for all subjects. In data from five out of six subjects, these entries show greater systematicity in searching for icons rather than words.

	MS	AH	DS	JF	FN	RW	$\bar{X}$	Combined
Simple icons	0.256	0.357	0.245	0.438	0.205	0.196	0.283	0.291
Complex icons	0.162	0.244	0.212	0.219	0.171	0.105	0.186	0.194
Simple words	0.168	0.268	0.189	0.226	0.163	0.060	0.179	0.184
Complex words	0.135	0.222	0.140	0.168	0.204	0.056	0.154	0.169

Table 6.10: Transition matrices of successive saccades for the six subjects on the four target types.

A one-way repeated measures ANOVA (with four levels: CW, SW, CI, SI) gives a highly significant result ($F(3,15) = 11.99, p < 0.001$).

The mean of conditions CW and SW is 0.167 and the mean of conditions CI and SI is 0.234. A second repeated measures ANOVA with one factor (with two levels; words, icons) shows that these two means are significantly different ($F(1,5) = 15.32, p < 0.05$).

The mean of conditions CW and CI is 0.170 and of conditions SW and SI is 0.231. A third repeated measures ANOVA (with two levels; complex, simple) shows that the difference is also significant ($F(1,5) = 10.09, p < 0.05$).

Due to the arrangement of the stimulus array, participants only needed to employ E, S, W and N saccades in order to cover it. So, 4×4 transition matrices were derived. To obtain meaningful proportions in a 4×4 transition matrix, it is necessary that a substantially greater number of saccades be available than that provided by a single observer. This difficulty prevented the analysis of each individual matrix, and it was necessary to use a combined transition matrix in each parallel condition, analysing all saccades as combined across subjects. The four 4×4 transition matrices are depicted in Table 6.11.

It was also worth examining whether the lack of independence was of perceptual and cognitive importance, and if it was related to the experimental conditions. As the matrix in Table 6.12 shows, previous and successive saccade directions can be 'Identical' (I in the matrix) or 'Opposite'(O), and can be compatible with a 'Clockwise' searching of the array (C) or compatible with an 'Anti-clockwise' searching (A).

	CW				SW				CI				SI			
	E	S	W	N	E	S	W	N	E	S	W	N	E	S	W	N
E	55	13	21	11	55	17	21	7	43	23	21	13	39	34	18	9
S	8	53	23	16	10	53	21	16	14	41	24	21	15	52	18	15
W	20	10	58	12	20	7	54	19	18	17	51	14	15	11	52	22
N	12	21	12	55	17	20	11	52	18	25	13	44	19	13	17	51

Table 6.11: 4 × 4 transition matrices results (percentages).

	E	S	W	N
E	I	C	O	A
S	A	I	C	O
W	O	A	I	C
N	C	O	A	I

Table 6.12: Categorisation of each possible pair of successive saccades in a 4 × 4 transition matrix.

The mean percentage in each of these four categories of the experimental conditions are in Table 6.13.

	I	O	C	A
Complex words	55	20	15	10
Simple words	54	19	18	9
Complex icons	45	21	20	14
Simple icons	49	15	23	13

Table 6.13: Proportion of each type of pairs of successive saccades in the four experimental conditions. Identical, Opposite, Clockwise and Anti-clockwise are the types (according to Table 6.12).

Computing the means, two new tables were produced which compare words versus icons and complex versus simple items (combined in Table 6.14).

	I	O	C	A
W	54	20	16	10
I	47	18	21	14
W−I	7	2	−5	−4
C	50	20	18	12
S	52	17	20	11
C−S	−2	3	−2	1

Table 6.14: Proportion of each type of pairs of successive saccades comparing words versus icons and complex versus simple items.

6.9 DISCUSSION

6.9.1 Search paths

Figures 6.6 to 6.8 provided examples of search paths shown by subjects. Figure 6.6 shows a systematic, although long, clockwise search. The subject makes a saccade from the central referent (A) firstly to B at 3 o'clock and sweeps round until locating the target at 2 o'clock (C). Of particular note here is the 'corner effect'; the four items from all corners have been omitted from the search due to the subject 'cutting corners'. Clearly, had the target been in any of these areas, it would have been missed. It might also be added that such scan paths show no evidence of a consistent visual lobe pattern.

Figure 6.7 demonstrates a miss and the commencement of a re-search cycle. Here the subject firstly saccades from A to about 11 o'clock (B) and into what looks like the start of an anti-clockwise search, but sharply turns into a clockwise search cutting the top-right corner (C) at 2 o'clock and then missing the target at 3 o'clock (D). The subject continues, but more carefully searching corner items also, and hence producing a characteristic square/rectangular path. The search path is rejoined on the top line (E) before the subject searches the previously-missed top-right corner (F), and follows the items arranged down the right-hand side to locate the target at D.

Figure 6.8 illustrates an unsystematic search shown during a complex word search. Replaying the computer-analysed sequence on a display screen makes what occurred clearer. Starting from the central referent (A), the subject saccades to 12 o'clock (B) before traversing a blank area of the display, seems to fixate at a blank point (C), perhaps momentarily wondering whether to adopt a strategy. The search veers off to D before almost reversing on itself anti-clockwise to E. The subject then omits corner F before locating the target at G. Due to the relatively high visual similarity of complex words to one another, it is difficult to interpret this erratic path as due to peripheral vision guiding the eye (incorrectly) to non-targets which were similar to the referent item.

6.9.2 Fixation durations and general comment

Although fixation durations were longer for icons than for words, these differences were not statistically significant. The fixation durations, however, were generally shorter than might have been expected. For instance, Gould and Dill's (1969) visual search study of pattern discrimination yielded a mean fixation duration of 323 ms (sd = 55). It is likely that the easier is the task, the shorter will be the fixation duration. The more information which needs to be perceived in one fixation, the longer will be necessary for a fixation. Clearly, little fine detail needed to be taken in during each fixation on a cell in order for the perceptual system to decide that the search had failed and to look elsewhere. Perhaps the abstract dot patterns used by Gould and Dill required longer processing time than the more meaningful icons or words used in this study.

Four different measures of scan paths showed what are possibly differences in search efficiency between the four different target types. The data seem to suggest that more systematic strategies are used on complex words with progressively less efficient strategies been shown with simple words and complex icons through to a fairly inferior strategy shown for simple icons. Scan path analysis of the number of cells fixated before target location suggests inferior strategies being employed on complex words with progressively more efficient strategies with simple words and complex icons to a superior strategy shown for simple icons. In addition, it was observed that there were approximately equal numbers of clockwise and anti-clockwise searches as well as those where the scan path was erratic or too brief to interpret.

6.9.3 First saccades to target

One of the things which proved useful was an analysis of first saccades to target. Looking at instances where the very first saccade made located the target, there was a greater frequency of successful fixations after the first saccade on icon targets as opposed to word targets. There was also a greater frequency of successful first fixations being on simple targets as opposed to complex targets.

Although there were less instances of fixations falling correctly where targets were in corner, as opposed to non-corner positions, this difference was not beyond chance expectation.

6.9.4 Items fixated before target

Secondly, there was a measure of number of items fixated before the target was located. The most salient finding here was that there were less different targets fixated before a simple icon was detected than before a complex word was detected.

6.9.5 Misses

A third measure was termed 'misses'. This was where eye movement recordings determined that subjects had, in fact, fixated on the target, yet apparently failed to perceive it. For instance, there were more misses on complex targets than simple targets.

This phenomenon is akin to the problem of human performance error seen in chest x-ray searches by radiologists (e.g. Nodine & Kundel, 1990). Eye movement research has shown that most missed lung tumours are the result of faulty interpretation, not faulty search. Over 70 percent of the missed tumours are fixated and most of these are processed by fixation durations long enough to recognise the abnormality, but rejected as true tumours.

One of the many significant differences revealed with this method was that of corners and side targets. Whereas 127 misses related to sides positions (expected frequency = 144.8), 54 related to corners (expected frequency = 36.2). This may be explained in terms of peripheral vision. Given a situation where the attentional field consists only of the stationary field and the eye field (see §5.2.1), targets situated at or near the corners of a square arrangement will be situated more into the periphery of vision and therefore their peripheral conspicuity would be lower. In order for these targets to have a similar conspicuity level they would need to be larger than those within the sides. This difference could of course have been avoided by positioning targets and non-targets in a circular arrangement, equidistant from a central point. However, in this particular study (compared with Expt. 11), a relatively 'real' scenario was been strived for, complete with the typically square or rectangular arrangement of VDU displayed items. What is important is that such a 'corner effect' is likely to occur in real-life VDU tasks, although in general it passes without acknowledgement.

6.9.6 Transition matrices

Possibly more productive was a transition matrix analysis. The frequency of change of direction was determined by use of transition matrices of successive saccades. Due to the perpendicular arrangement, a more systematic search would consist largely of vertical or horizontal saccades. On the other hand, an unsystematic or inefficient search strategy would be expected to show relatively more oblique changes in saccade directions.

In data from five out of six subjects, the matrices in Table 6.5 show greater asystematicity (greater proportion of diagonal saccades) in searching for icons rather than words. For instance, with subject JF, some 44 percent of saccades were diagonal in the simple icon condition, whilst only some 17 percent of saccade direction change were diagonal whilst searching for complex word arrays. In the vast majority of cases (11 out of 12 comparisons), words showed less diagonal saccades than did icons. Also, as might be expected, simple items showed more diagonal saccades than complex items. It is suggested that peripheral vision is used more in locating simple items. Where peripheral vision cannot be used, users are compelled to employ a more systematic search.

The four 4 × 4 transition matrices are depicted in Table 6.9. Each entry within Table 6.9 is a percentage and its meaning is consistent with the rationale outlined in §3.8.1. For example, in condition CW, the probability of a saccade of type E after a saccade of type E is 0.55, the probability of a change of direction after a saccade of type E to one of type S is 0.13, that to type W is 0.21, and of that to type N is 0.11. In these transition matrices, only pairs of successive non-diagonal saccades have been taken into account.

The first conclusion from the 4 × 4 transition matrices is that there is not independence between previous and successive saccade directions. Independence would require all the rows in each matrix to be approximately equal and, clearly, this is not the case.

It was therefore worth examining whether the lack of independence was of perceptual and cognitive importance, and if it was related to the experimental conditions (see Table 6.10). Some elements of the transition matrices have a similar meaning. Previous and successive saccade directions can be 'Identical' (I in the matrix) or 'Opposite' (O), and can be compatible with a 'Clockwise' searching of the array (C) or compatible with an 'Anti-clockwise' searching (A).

The mean percentage in each of these four categories of the experimental conditions are in Table 6.12. In interpreting these latter tables, it is important to bear in mind that only non-diagonal saccades have been used in the construction of the 4×4 reduced transition matrices. These transition matrices are made from saccades type E, S, W and N. The twenty items to be searched for are regarded as the four sides of a square, and therefore these four types of saccades would be the more 'natural'—should a systematic search strategy be in use. So, it would be expected that the transition matrix analysis would provide information on the main features of the systematic component of the search; again, should such a search be involved.

Computing the means, two new tables were produced which compare words versus icons and complex versus simple items (combined in Table 6.13). We can summarise these results in terms of the proportion of diagonal saccades and in terms of transition matrices. Concerning the former: It is proposed that if a target is detectable in peripheral vision, then there is a less systematic search.

Concerning the transition matrices: Firstly, as Table 6.13 shows, figures in column A are smaller than corresponding figures in column C. This suggests that 'clockwise' searches were more typical than 'anti-clockwise' searches. If the C percentage is compared with the A percentage in each of the rows of the 4×4 matrices of Table 7, it is apparent that only in one of the comparisons out of sixteen does the A percentage exceed the C percentage.

Secondly, as the $C - S$ row of Table 6.14 shows, no clear differences transpired between the difficulty levels: Differences between complex and simple items are not large. Consequently, the relation between a previous saccade direction and a subsequent saccade direction is similar in complex and simple tasks.

Thirdly, the effects of icons versus words seem more important. The differences are now larger than before. When the target is an icon, the I percentage is lower, and both the C and A percentages are higher than when the target is a word. This pattern is to be expected if searching for an icon is conducted with longer saccade amplitudes. If this were the case, then the search would require more changes in saccade direction and less same-direction saccades in order to cover any side of the square arrangement. A *post hoc* two-way repeated measures ANOVA was conducted on the mean of non-diagonal saccade amplitudes comparing words versus icons and complex versus simple items. This showed that there was, in fact, a significant effect on the factor words-versus-icons ($F(1,5) = 12.50; p < 0.05$) but not on the factor complex-versus-simple ($F(1,5) = 5.07; NS$). Searching for icons may therefore be regarded as involving a more efficient pattern of eye movements.

6.10 CONCLUSIONS

It was concluded within Experiment 8 that search times were greater for the word condition than for the icon condition, and greater when the item was complex than when it was simple. It seems sensible to assume that the more useful the peripheral vision is—in other words, the more easily resolvable the target is in the periphery—the more the saccades used in the task will be guided by peripheral vision and the less by centrally determined scanning strategies. For instance, if the target is resolvable in peripheral vision, it is an appropriate strategy to rely on peripheral information and to search only those positions where peripheral vision detected greater similarity with the target. In such a case, the observer clearly does not produce a systematic search. On the contrary, if peripheral vision is not particularly informative, the observer can readily rely on a systematic search and so avoid redundant searching. Therefore, this assumption predicts a high level in the proportion of diagonal saccades in conditions where search time is low. The present results neatly fit this prediction. It is worth noting that in Experiment 2, a high proportion of diagonal saccades was related to low efficiency, whilst the opposite holds in this study. The inconsistency may simply lie in whether or not peripheral vision is important. As ever, such incongruencies illustrates the problems faced in attempts to relate psychological strategies to efficiency.

The structure of the systematic component of the search, as indicated by the transition matrix, shows greater differences between words and icons than between simple and complex items. The principle difference between words and icons is due to the presence of longer saccades in the icon condition.

The data from five out of the six subjects show greater systematicity in searching for icons rather than words. Also, simple items showed more oblique changes than complex items. It is suggested that peripheral vision is used more in locating simple items.

Scan path analysis of the number of cells fixated before target location suggests inferior strategies being employed on complex words with progressively more efficient strategies with simple words and complex icons to a superior strategy shown for simple icons. In addition, it was observed that there was approximately equal numbers of clockwise and anti-clockwise searches as well as those where the scan path was erratic or too brief to interpret.

The fact that average search times were shorter for icons may seem at odds with the suggestion that icons may be characterised by a *less* systematic search. The analogy which we would propose is that of the difference between searching on a page of a book for a reference date (e.g., '1959') and a Japanese name in its untranslated form. Although people are likely to be less systematic in searching for the Japanese character (they are likely to randomly move their eyes over the page waiting for the name to *pop-out*, search time is likely to be shorter than for the year date.

Where peripheral vision cannot be used, users are compelled to employ a more systematic search. In these situations we therefore see lower values within a transition matrix.

One final brief point concerns the recent advent of dynamic icons. Static icons are sufficiently 'messy' materials to deal with due to their irregularity of shape, problems in measuring area covered, differences in meaningfulness, etc. However, the new breed of dynamic icons would no doubt be even more difficult to standardise or quantify; e.g. in terms of spatial frequency.

In conclusion, we would propose that if a target is detectable in peripheral vision, then (a) there are fewer different positions fixated before the target is detected; (b) there are more cases where the first saccade successfully locates the target; and (c) there is a less systematic search (as revealed by a transition matrix). It therefore seemed a logical progression to conduct a further visual search study to investigate the role of spatial frequencies in icon versus word locatability (Chapter 7).

Chapter Seven

A Spatial Frequency Investigation

7.1 OVERVIEW

This chapter describes a relatively 'pure' study of search with particular regard for the possible role of peripheral processing. As search times for icons seem reliably shorter than for words, it is important to know what factors underlie this. A visual search study was carried out using spatial frequency grids to examine the role of variables such as cycle frequency, high/low contrast, and high/low similarity of distractors. An eye movement analysis of fixation durations and scan paths was then conducted to determine if this could account for the observed differences.

7.2 SPATIAL FREQUENCY REPRESENTATION

Chapter 6 compared the locatability of computer-type icons and more conventional verbal labels. Design of visual display symbology draws upon tradition, general recommendations (e.g. McCormick & Sanders, 1982), and experimental comparisons between sets (e.g. Stammers *et al.*, 1989). Ergonomists do not possess an effective descriptive metric for comparing visual stimuli. Such descriptive methods would allow quantitative comparisons of different symbols and offer predictions concerning discriminability (Marshak & Osarczuk, 1987).

In describing or thinking about objects, one naturally refers to areas of light and dark at particular locations in space. *Spatial frequency analysis* is an alternative descriptive way of specifying the visual scene. The spatial frequency representation is a *transform* of the visual image into a different, mathematically equivalent representation. The fundamental principle underlying the notion of spatial frequencies was posited by the French mathematician, J. B. Fourier:

> *Any function that repeats itself over and over can be synthesised as the sum of a series of sinusoids.*

The mathematical procedure by which functions are approximated by sums of sinusoids (e.g. a sine wave) is termed a *Fourier transform*. The classic text on this issue is that of Brigham (1974).

A *sine wave* is a continuous waveform that oscilates in a smooth and regular manner. It is characterised by the amplitude, wavelength (the time between corresponding points in successive cycles) and phase (the difference in timing between the two waves). One generally refers to the frequency of spatial sinusoids in terms of the number of cycles per degree of visual angle. If only cycles per mm were specified, it would be necessary to specify the viewing distance. Referring all measure to visual angle completely specifies the size of the retinal image.

It is worth emphasising that the spatial frequency (f in the equation of a sine) is *inversely* related to the wavelength, or size of the wave. A high spatial frequency (large f) has a small wavelength, and there are many cycles per degree. Thus, a high frequency pattern is fine and detailed, containing many waves in a small area, whilst a low frequency pattern corresponds to long, smooth, drawn-out waves (Levine & Shefner, 1981).

'Spatial frequency gratings' (see §1.6.1), as used in experimental work on vision, are defined by (a) the actual spatial frequency or number of cycles per unit area or per degree of visual angle, (b) contrast, and (c) orientation. To demonstrate that the visual system is differentially sensitive to different spatial frequencies, Blakemore and Campbell (1969) conducted an adaptation study. Their reasoning was that if there are seperate neural networks or 'channels' in the visual system that encode each spatial frequency seperately, then an individual should be able to adapt selectively and so reduce the sensitivity of a particular channel. After adaptation to a specific spatial frequency, which should reduce sensitivity to that frequency, participants were briefly shown several gratings which were of very low contrast (i.e. close shades of grey). It was found that after adaptation, more contrast was required in order to descriminate the spatial frequency from a uniform grey of the same average luminance. When tested on spatial frequencies that differed from the adaptation grating, however, little loss of sensitivity was shown.

This suggested that a neural pathway is responsible for coding spatial frequencies of one cycle per degree, another for coding two cycles per degree and so on. When a channel is adapted so that its sensitivity is lowered to subsequent stimulation, then a grating of that frequency needs to contain more contrasting light and dark areas for it to be detected. It also appears that there are channels specifically for each spatial frequency at each *orientation* (Haber & Hershenson, 1980).

EXPERIMENT 10:
SPATIAL FREQUENCY AND PERIPHERAL VISION

In the previous chapter, it was suggested that icons might be more easily searched for due to their lower spatial frequencies compared with words. The present study was designed to explore some aspects of spatial frequencies in a preliminary attempt to identify what variables underlie the supposed spatial frequency superiority of icons. Due to the difficulty of controlling for the multivariate nature of icons, it was decided to use 'pure' spatial frequency grids rather than 'real' icons as experimental materials. Also, due to the 'corner effect' suggested in Experiment 9, it was decided to employ circular rather than square arrangements.

It was hypothesised that visual search would be quicker for items of lower spatial frequency. Also, on little more than intuitive grounds, it was hypothesised that search times would be quicker where contrast was higher and where distractors (non-targets) were less similar to the target.

7.3 METHOD

Forty-eight paid subjects (28M:20F) participated. Ages ranged from 18 to 58 ($\bar{X}$ = 25.3; sd = 9.5).

Each presentation contained a central reference grid surrounded by eight grids in a circular arrangement equidistant from the centre. The peripheral grids consisted of one target and seven distractors. Due to possible problems from lateral masking (Bouma, 1978), it was decided to restrict the number of items to eight.

There were three variables examined: (a) number of cycles; (b) high versus low contrast; and (c) high versus low similarity of non-targets.

Arrays were prepared with MacDraw and laser printed ready for photographic slide preparation.

7.3.1 Number of cycles

Number of black/white stripes within each target varied from 1, through 1.5 (one black stripe, one white, one black), 2, 2.5, 3, 3.5, 4, 4.5, to 5. Thus, there were nine different cycles involved. Overall luminance of the (non-half-cycling) targets was about 45 cd/m^2; comprised of the combined luminance of the pure black area (25 cd/m^2) and the pure white area (62 cd/m^2).

7.3.2 High versus low contrast

Bands were either (a) pure black and pure white or (b) dark grey and light grey. The two greys (54 cd/m^2 and 31 cd/m^2) had a combined luminance equivalent to the black/white gratings.

7.3.3 High versus low similarity

This referred to the non-targets. The low similarity arrangement had no distractors of spatial frequency close to that of the target, whereas the high distractor arrangement did. The high similarity arrangement with a target of 2 cycles, for example, would have non-targets of 1, 1.5 (two) 2.5 (two), 3, and 3.5. As another example, the low similarity arrangement with a target of 5 cycles would have non-targets of 1 (two), 1.5 (two) 2, 2.5, and 3 cycles. Figures 7.1 and 7.2 show examples of the arrangements used. Appendix 7.1 details the systematic organisation of non-targets in high and low similarity conditions.

In addition, targets were presented at each of eight positions of the compass. There were thus a total of $9 \times 2 \times 2 \times 8$; i.e. 288 arrangements. As this was considered too large a number to show to subjects without limiting the number of volunteer participants, it was decided to use only a balanced half of the available arrangements with each subject. Thus, each participant viewed a total of 144 arrangements.

Search arrays were presented to subjects via slides. The presentation order for each subject was varied. Slides were displayed from a projection tachistoscope. Subjects were simply required to search for the grid corresponding to the central grid and to press a key when found (complete instructions to subjects are contained within Appendix 7.2). Timing (ms) was automatically recorded with a task-designed timing programme running from a BBC micro-computer connected to the morse-key and projection tachistoscope. The interval between a key-press and the next slide appearing was a constant four seconds.

Viewing distance (held constant by use of a chin rest) from subjects' eyes to screen was 90 cm. Screen to lens distance was 100 cm. An f = 85 mm lens was used. The images (projected onto a rear-projection screen) measured 27 cm from top grid to bottom grid and from right grid to left grid. Each grid measured 9 cm^2. Each grid subtended a visual angle of about 1.9° whilst the entire display subtended about 16.7°. Therefore, the grids of lowest spatial frequency consisted of approximately 0.5 cycles/degree and those of the highest spatial frequency were comprised of 2.5 cycles/degree. Screen luminance was 46 cd/m^2. Ambient illumination was 16 lux.

Errors were recorded by the experimenter.

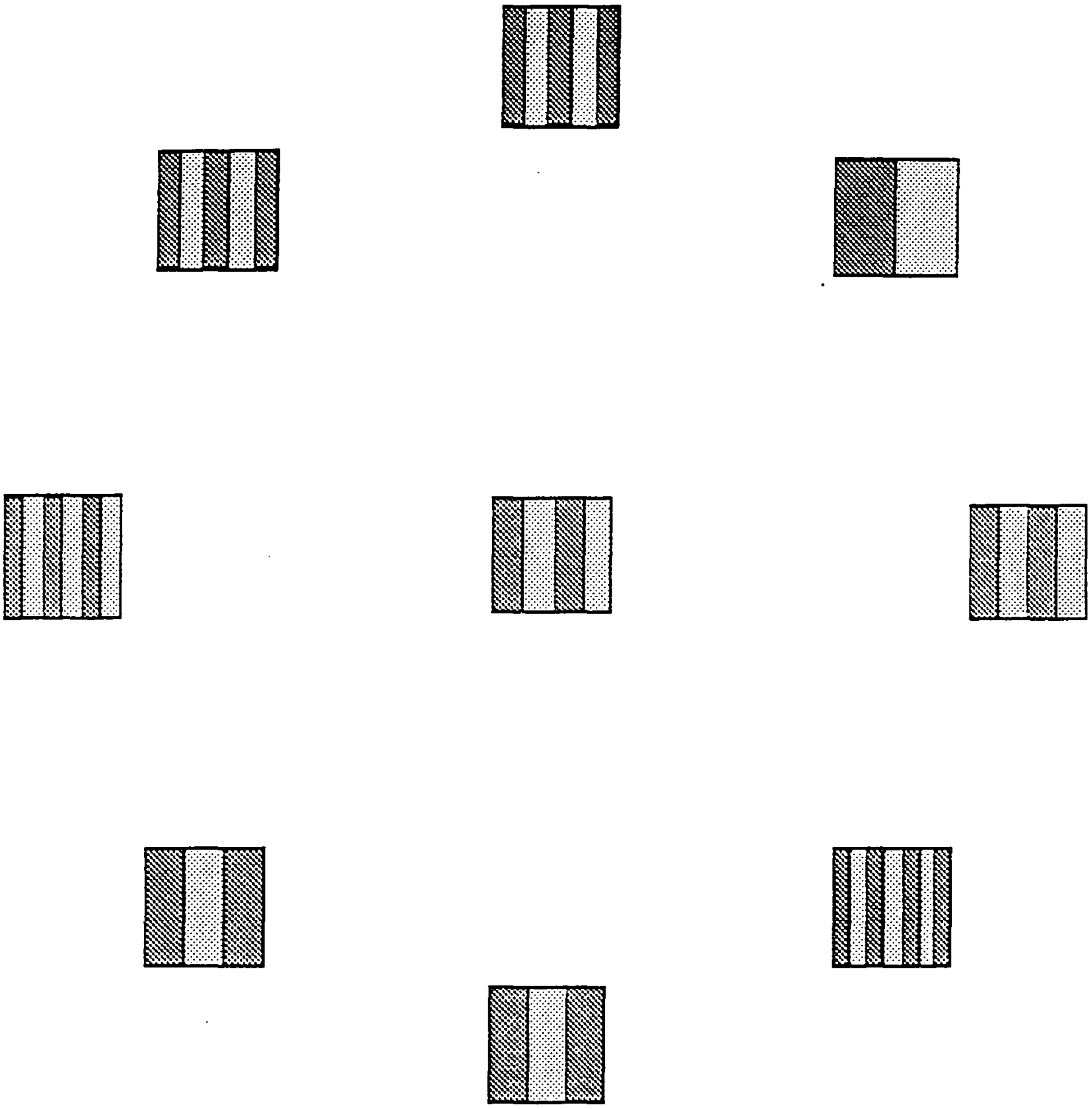


Figure 7.1: Example of a 2 cycle, high similarity, low contrast arrangement.

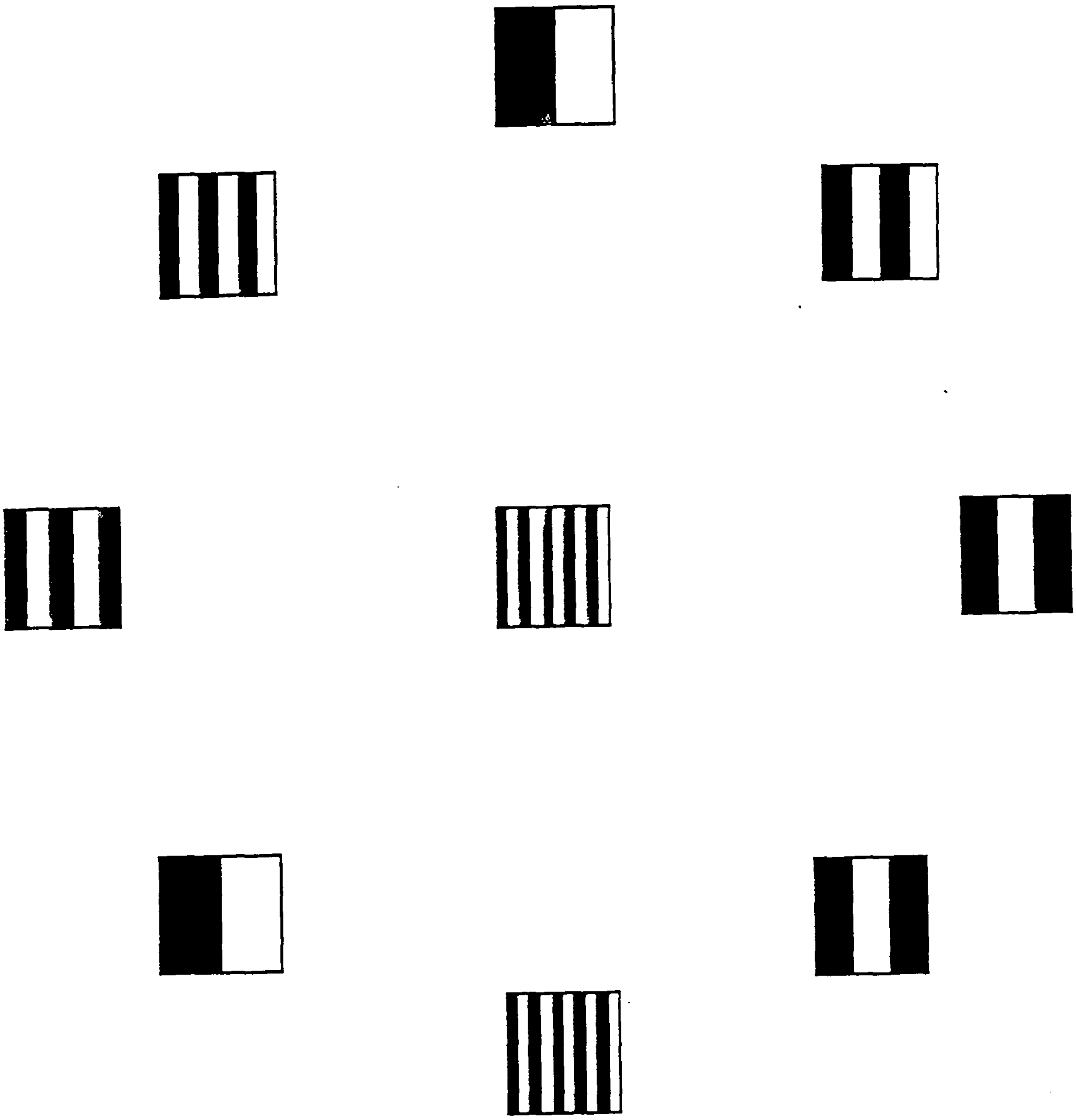


Figure 7.2: Example of a 5 cycle, low similarity, high contrast arrangement.

7.4 RESULTS

A three-way repeated measures ANOVA was conducted on the data with the variables being number of cycles, high versus low contrast, and high versus low similarity. Table 7.1 summarises the results. It can be seen that there were highly significant effects within all variables and two-way interactions. Furthermore, the three-way interaction of variables was also significant.

Factor	SS	MS	F	df	p
Cycles	471515395	58939424	139.21	8	< 0.001
Contrast	8937084	8937084	25.83	1	< 0.001
Similarity	542381387	542381387	469.84	1	< 0.001
Con. x Sim.	4310705	4310705	9.57	1	< 0.01
Con. x Cyc.	11516973	1439622	4.70	8	< 0.001
Sim. x Cyc.	82389888	10298736	35.14	8	< 0.001
Con. x Sim. x Cyc.	5136395	642049	2.07	8	< 0.05

Table 7.1: Results from the three-way ANOVA.

For the cycles variable, 40.3 percent of the variance (ω^2) was accounted for by the cycles as opposed to the between subject and error variances. For the similarity variable, 13.6 percent of the variance was likewise accounted for, and for the contrast variable the ω^2 value was 2.6 percent.

Table 7.2 shows search time results between the nine levels employed within the cycles variable. Figure 7.3 graphically depicts these results. It can be seen that, with the exception of the extreme high frequencies, the pattern followed that hypothesised.

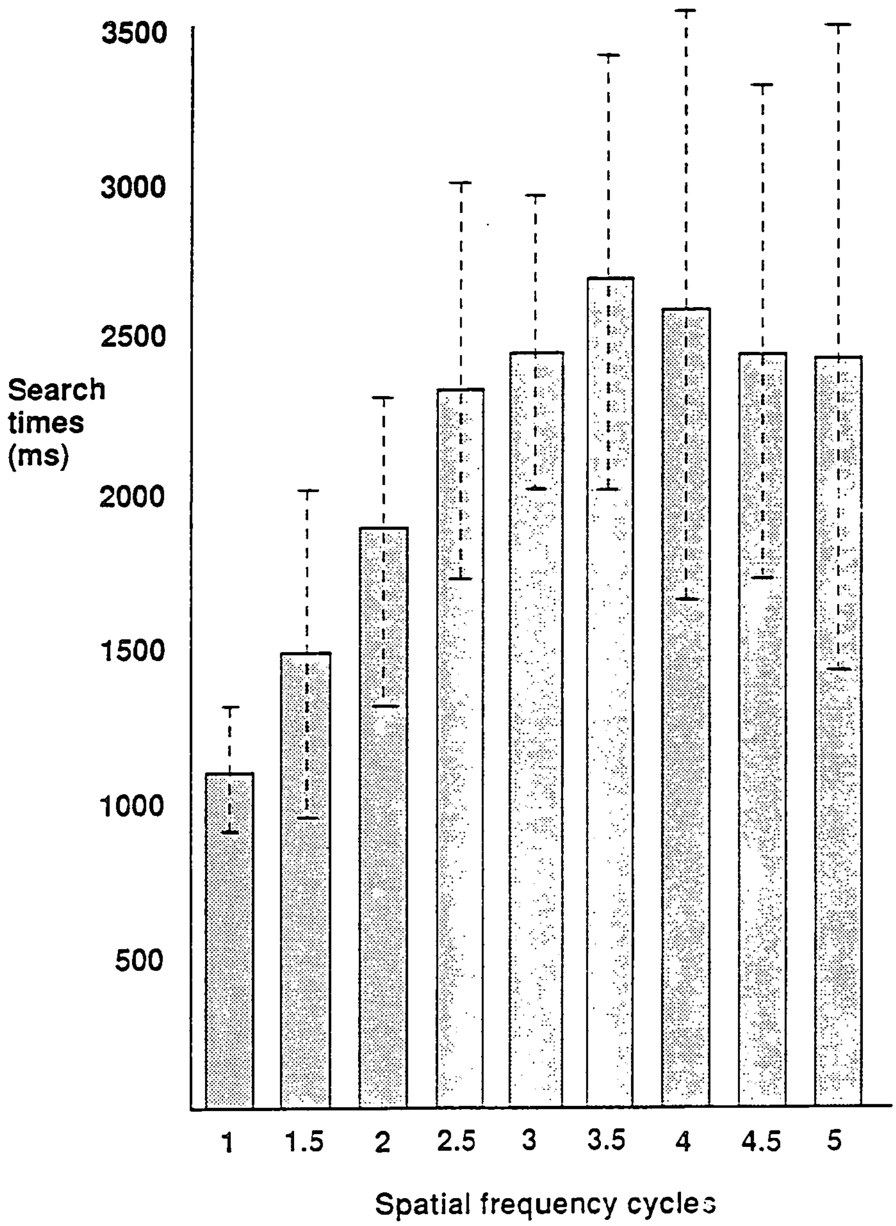


Figure 7.3: Histogram of search times (ms) obtained for the nine different spatial frequency cycles (vertical lines show standard deviations).

Cycles								
1	1.5	2	2.5	3	3.5	4	4.5	5
1095 (198)	1481 (531)	1784 (503)	2339 (651)	2471 (488)	2687 (739)	2582 (953)	2438 (866)	2422 (1069)

Table 7.2: Search times (ms) obtained for the nine different spatial frequency cycles (standard deviations in parentheses).

Table 7.3 summarises search times when the data from Table 7.2 are seperated into that from high similarity and low similarity conditions. Figure 7.4 depicts these results graphically.

CYCLES								
High similarity								
1	1.5	2	2.5	3	3.5	4	4.5	5
1277 (581)	1931 (730)	2216 (804)	2830 (995)	2880 (981)	3271 (1126)	3354 (1047)	3180 (1188)	3347 (1374)
Low similarity								
928 (387)	1033 (399)	1352 (583)	1803 (665)	2056 (755)	2091 (757)	1777 (731)	1707 (709)	1509 (631)

Table 7.3: Search times (ms) obtained for the nine different spatial frequency cycles on high similarity and low similarity arrays (standard deviations in parentheses).

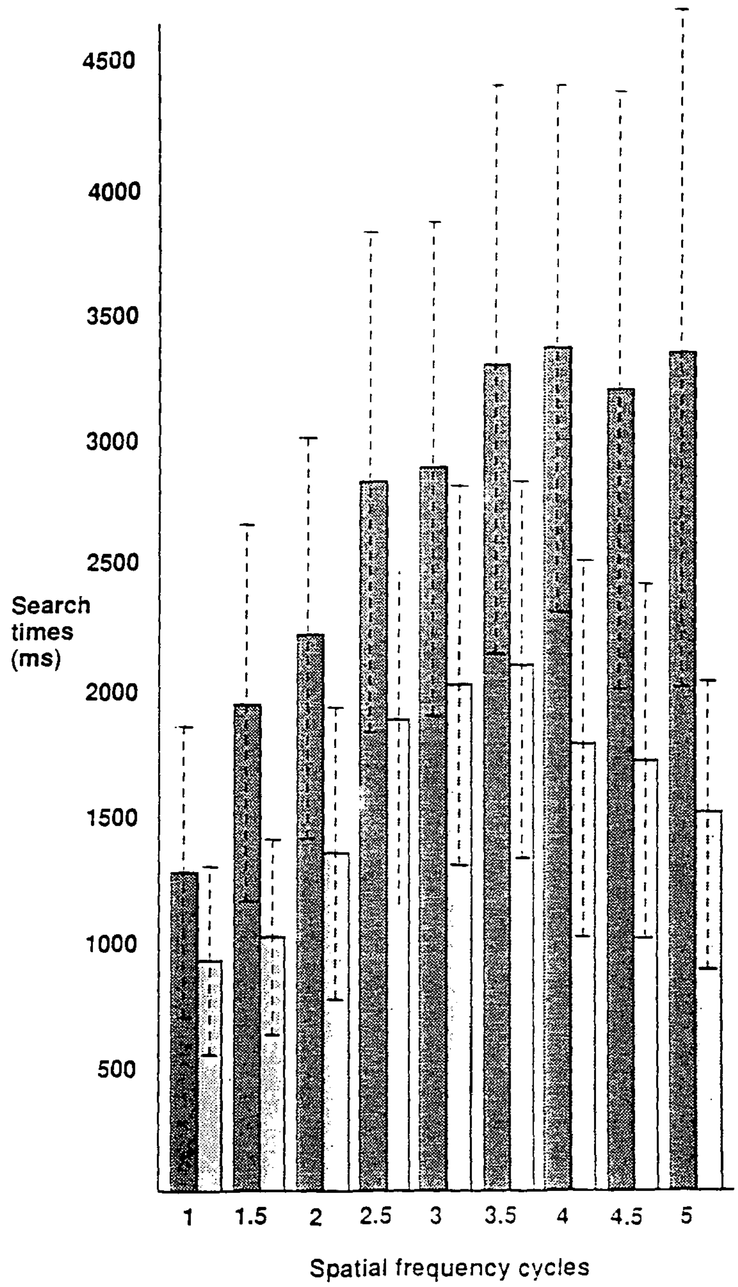


Figure 7.4: Histogram of search times (ms) obtained for the nine different spatial frequency cycles when data are seperated into that from high similarity (dark bars) and low similarity (light bars) arrays (vertical lines show standard deviations).

Post hoc Tukey's WSD analyses were then performed on data from the three-way ANOVA to determine between which cycle frequencies the significant differences lay. Table 7.4 presents a matrix of results from these analyses.

Cycle frequency									
	1	1.5	2	2.5	3	3.5	4	4.5	5
1	—	—	—	—	—	—	—	—	—
1.5	0.001	—	—	—	—	—	—	—	—
2	0.001	0.001	—	—	—	—	—	—	—
2.5	0.001	0.001	0.001	—	—	—	—	—	—
3	0.001	0.001	0.001	NS	—	—	—	—	—
3.5	0.001	0.001	0.001	0.001	0.01	—	—	—	—
4	0.001	0.001	0.001	0.01	NS	NS	—	—	—
4.5	0.001	0.001	0.001	NS	NS	0.01	NS	—	—
5	0.001	0.001	0.001	NS	NS	0.01	NS	NS	—

Table 7.4: Summary of significance levels following the Tukey's test comparisons.

Table 7.5 summarises search times for all subjects on the high/low contrast and high/low similarity variables. It is clear that in each case the differences fell in the hypothesised direction.

	High	Low
Contrast	2070 (835)	2213 (819)
Similarity	2527 (946)	1582 (439)

Table 7.5: Summary (means) of search times (ms) for all subjects on the contrast and similarity dimensions (standard deviations in parentheses).

Table 7.6 summarises the interaction between contrast and similarity.

	Low similarity	High similarity
Low contrast	1704 (517)	2724 (758)
High contrast	1460 (331)	2680 (732)

Table 7.6: Summary of the interaction between the variables contrast and similarity. Figures refer to mean search times (ms) for all subjects (standard deviations in parentheses).

Figure 7.5 depicts the interaction between contrast and similarity.

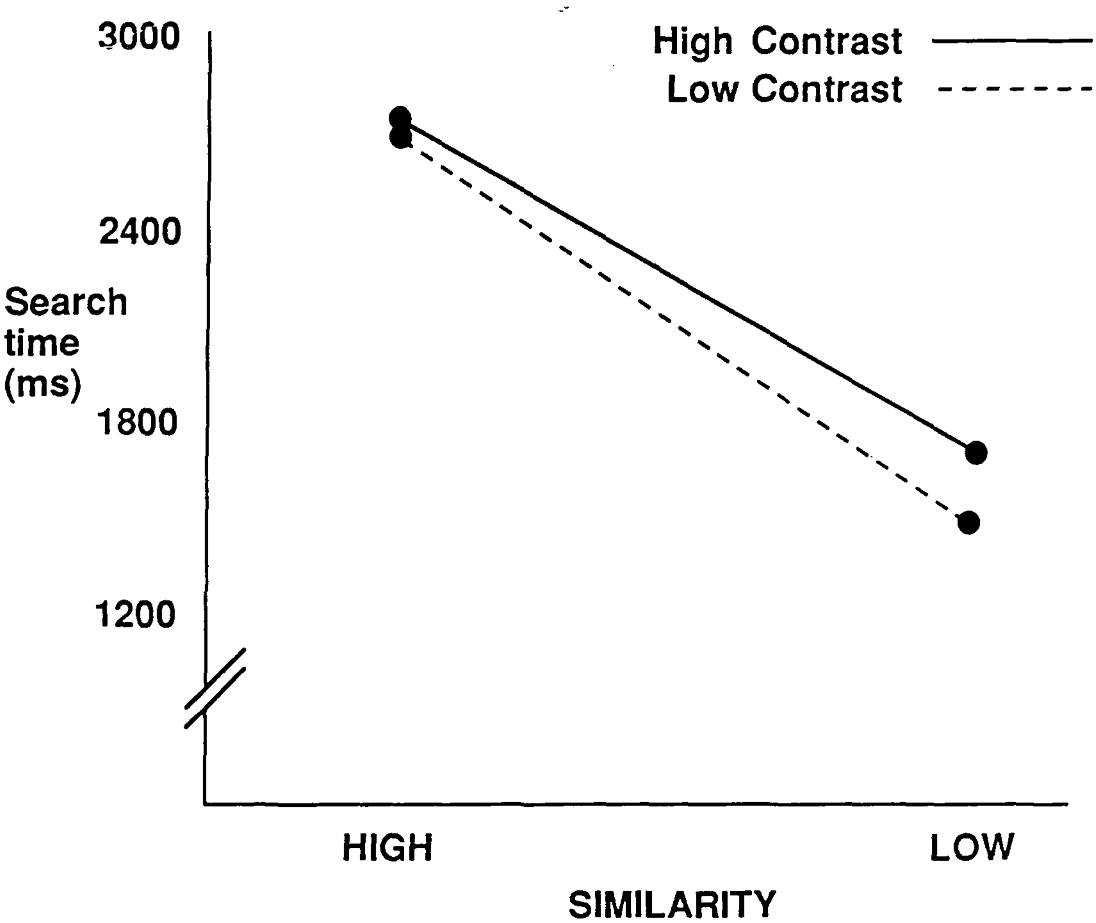


Figure 7.5: The interaction between contrast and similarity.

7.5 DISCUSSION

In line with the experimental findings and background outlined in Chapter 6, it was hypothesised that search times would be faster with grids of lower spatial frequencies. This expected pattern was clearly found with the exception of a slight 'dip' within the 4, 4.5 and 5 cycle grids. As the trend over the six grids of lower frequencies is continuous, further investigation of this phenomenon seems warranted. It is possible that there is some threshold for visual search optimisation around 1.75 cycles/degree and 2 cycles/degree. There does seem to be greater between- and/or within-subject variability at the higher frequency end, as indicated by the higher standard deviations. However, considering the margin of drop between these three higher cycle frequencies, and particularly considering that they nowhere near stray from within the standard deviation of the slowest-searched frequency, (3.5 cycles), it is likely that these are deviations not straying from chance levels within the overall pattern. This possibility is strengthened by the fact that *post hoc* analyses from the ANOVA on cycle frequencies showed very highly significant differences between the majority of cycle frequencies. There was, however, a cluster of non-significant differences at the higher end of the spectrum (see Table 7.4). For future research along these lines it would be advisable to widen the range of cycle frequencies sampled.

A further argument, referring to the icon/verbal label distinction, would take the following lines. Essentially, there is no difference in locatability of patterns between the 3 cycle grids (1.5 cycles/degree) and the 5 cycle grids (2.5 cycles/degree). In 'real' (VDU) terms, although there is much variability between items (particularly icons), the spatial frequencies of most icons and words would fall around these values. Discounting the lower spatial frequency grids used in this study, which may have little relevance to actual VDU items used, differences in spatial frequency has no effect on performance (at least in relation to peripheral processing).

A second hypothesis was that grids of higher contrast (i.e. black/white as opposed to dark/light greys) would promote quicker search times. This was also supported in that the mean search time for high contrast targets ($\bar{X}$ = 2070 ms; sd=835) was significantly ($p < 0.001$) faster than that for low contrast targets ($\bar{X}$ = 2214 ms; sd=819). Although perhaps intuitively 'obvious', there appears to be no other empirical evidence for this phenomenon within the HCI literature. This again offers a new avenue for exploration.

The third hypothesis was that arrays with non-target items of low similarity to the target item would have faster search times than those containing distractors of high similarity to the target. Once more, this hypothesis was supported. Mean search times for arrangements in the low similarity condition were significantly ($p < 0.001$) faster ($\bar{X} = 2527$; sd=946) than those in the high similarity condition ($\bar{X} = 1582$; sd=439). Such differences in the discriminability of items, particularly as perceived in peripheral vision, also warrant further scientific examination.

All three two-way interactions produced highly significant results (see Table 7.1). High contrast positively interacted with low similarity, high contrast positively interacted with the lower frequency of cycles, and low similarity of non-targets positively interacted with lower frequency of cycles. There was even a significant interaction between the three variables.

The next issue to determine was the percentage of variance accounted for by each of the three variables employed. As the seminal work on this issue makes apparent, with a three-way repeated measures analysis of variance, this is not a clear-cut matter (Vaughan & Corballis, 1969). An in-depth look at this article and the rationale of Winer (1971) suggested that this approach (ω^2) was justified. However, for simplicity (and considering the present concern with *main* factors), variance explained for main effects is best calculated individually (as if a one-factor ANOVA had been conducted), and expressed (a) in relation to variances for the between subjects and error variances, and (b) expressed as ratios to each other rather than as a proportion to the total variance for all factors. The obtained variances for the cycles, similarity and contrast variables were 45.7 percent, 13.6 percent and 2.6 percent, respectively. This provides a ratio of effect of approximately 40:14:3, or about 13:5:1. The strength of the cycles variable can thus be seen not only in the fact that differences were significant beyond a 0.001 percent level of expectation, but that the effect of this variable was thirteen times stronger than that of the contrast variable (which again was highly significant). The similarity variable was approximately five times stronger than the contrast variable.

From this analysis of the respective variances accounted for by the three factors and from the three-way ANOVA summarised in Table 7.1, it is clear that contrast exercises a much less powerful influence than the two other variables. The interaction plotted in Figure 7.5 reinforces this fact as it can be seen that contrast has little effect in both high similarity and low similarity cases. The effect is, nevertheless, in the expected direction.

The interaction between similarity and cycles is also interesting. From Figure 7.3, it would seem that, as suggested above, the effect of cycles tails off at a certain point. However, more detailed examination of the data revealed that this explanation does not tell the full story. Seperating the effect of spatial frequency cycles for high and low similarity (see Table 7.3 and Figure 7.4), it is seen that the cycle factor has a consistent effect only for high similarity grids. The cycle effect diminishes with the relatively higher spatial frequencies where targets to be searched for are surrounded by low similarity distractors. It can therefore be said that if discrimination is easy, then cycle frequency does have a consistent effect; perhaps mediated through peripheral vision. In more difficult cases, sequential scanning is more likely required.

One minor point worth mentioning is that within the half-cycle grids (i.e. 1.5, 2.5, 3.5, and 4.5 cycles), there was an unequal proportion of black and white areas; the ratio being 2:1 black to white. It was not felt likely that this would produce any effect as the frequencies of these grids were evenly balanced out across the presentations. The lack of any 'saw-tooth' effect on the alternating half-cycle–full-cycle pattern (see Fig. 7.3) demonstrates that this was so.

Other variables might have been included within this analysis and that of orientation was a contender which was excluded only on grounds of the need to limit the subject's time at task. It is certainly a principal candidate for future research. One spatial frequency analysis using this variable is that of Heeley and Timney (1989).

There may be another implication for spatial frequencies and VDUs. Lunn and Banks (1986) have proposed a basis for a purely visual component of VDU-induced fatique. They suggest that adaptation to the fundamental spatial frequency of lines of text on a VDU provides a single explanation for a wide variety of reports of visual fatigue. Reliable contrast threshold elevations at spatial frequencies of 2, 3, and 5 cycles/degree were found after participants read single-spaced text on a VDU. Lunn and Banks point out that this adaptation also reduces sensitivities to spatial frequencies in the range largely responsible for the reflexive accomodative response (2–6 cycles/degree), and therefore could account for objective optometric measures of disturbed accomodation as well as some subjective effects of viewing VDUs. The authors also discuss implications for VDU design.

7.6 CONCLUSIONS

The variables examined in this study all proved to have strong effects on the peripheral perception of spatial frequency. The cycles variable was particularly strong, having an effect thirteen times that of the contrast variable. On a note which may give impetus to follow-up studies, it might be remarked that although the results do support the notion of the lower spatial frequency characteristics of icons resulting in their search superiority over verbal labels due to their relative ease of peripheral processing, the findings do not demonstrate that the effect is exclusively due to that mechanism. Articulatory distance, global/local features, etc. may also play a role.

EXPERIMENT 11: AN EYE MOVEMENT STUDY OF SPATIAL FREQUENCY AND PERIPHERAL VISION

7.7 INTRODUCTION

To gain further evidence for the peripheral perception/spatial frequency theory of icon search supremacy, it was decided to finely analyse search scan paths in an eye movement monitoring study. The scleral search coil method was again chosen due to the need to monitor both the vertical and horizontal planes. In this experiment, due to the fact that peripheral processing was held very much to be influential, and that targets—if sufficiently conspicuous—might be located almost immediately (e.g. 'a single saccade away'), *no* differences in scan paths or eye movement parameters might be expected. The differences found within Experiment 10 might, very simply, be due to peripheral processing *per se* leading almost immediately to detection. Although the general experimental hypothesis was that there would be significant differences between eye movement variables manifest over the three factors, in this case an overall null hypothesis is worth stressing: That *no* significant differences would emerge within the variables from eye movement recordings.

7.8 METHOD

Four subjects participated although technical problems resulted in data from one subject being discarded. Experimental conditions were as for Experiment 10. Programme SAMFIO was employed with a sampling rate of 10 ms. This programme provided details of fixation positions (and hence scan path), and whilst fixation durations were given, the programme did not provide data on saccade amplitudes.

7.9 RESULTS

7.9.1 Search paths

Figures 7.6 to 7.8 are photographs of search paths obtained from one particular subject (FN). Figure 7.6 shows a phenomenon typically observed. From the reference grid (centre), the subject makes the first saccade immediately to a non-target item (*NE*) which is just half a cycle removed from the target grid. The subject then checks the reference and quickly locates the correct target (*NW*).

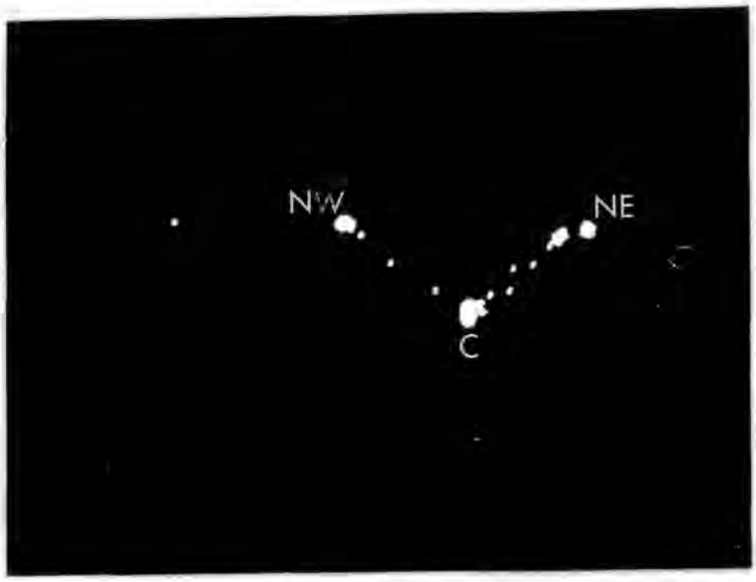


Figure 7.6: A search path showing the effects of peripheral vision.

Figure 7.7 shows a systematic (anti-clockwise) search path to locate the target at E.

Figure 7.7: A systematic, yet relatively long, search path.

Figure 7.8 shows a relatively unsystematic search path by which a target at S is located. This path and that of the other two illustrated here are further discussed within §7.10.1.

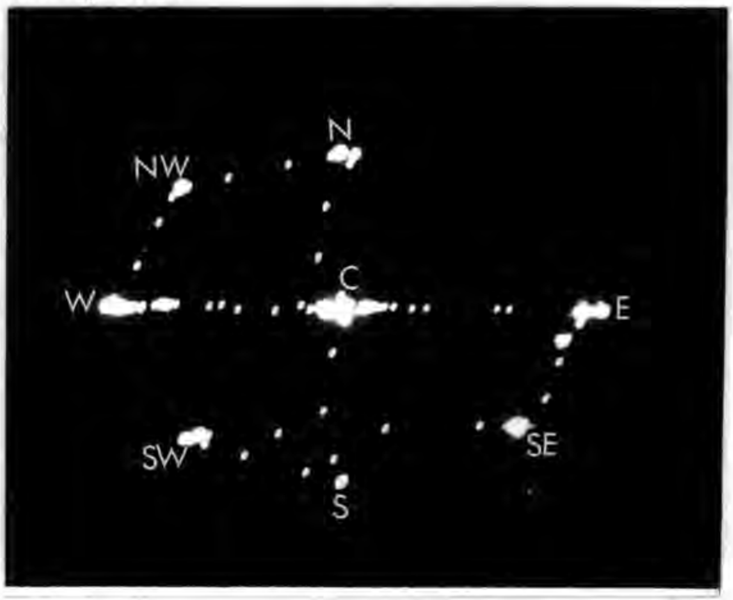


Figure 7.8: An unsystematic search path.

7.9.2 Fixation duration

Table 7.7 shows the results of fixation duration analysis on the cycle frequency variable. Taking the mean scores from each of the three subjects on this variable, a one-way repeated measures ANOVA showed there to be no significant differences between the nine cycles used ($F(8, 16) = 1.663$).

Cycles	1	1.5	2	2.5	3	3.5	4	4.5	5
Mean	255.5	247.0	238.0	252.7	248.0	276.4	282.7	245.6	245.0
sd	156.8	127.4	146.4	137.5	124.9	127.8	130.8	124.3	112.0

Table 7.7: Fixation durations (ms) of mean times obtained between the nine different spatial frequency cycles; data are combined from all three subjects.

Table 7.8 depicts mean fixation duration for high contrast and low contrast grids for the three subjects. Differences were not significant ($t = -2.994$; $df = 2$).

Subject	High contrast	Low contrast
FN	264.1 (111.7)	274.2 (128.2)
JF	267.4 (156.3)	294.1 (175.5)
VP	233.4 (106.9)	259.7 (164.1)
Combined	251.6 (125.5)	272.8 (157.7)

Table 7.8: Fixation durations (ms) of the three subjects for high and low contrast grids (standard deviations in parentheses).

Table 7.9 depicts mean fixation duration for high versus low similarity of non-targets for the three subjects. Differences were not significant ($t = -1.515; df = 2$).

Subject	High similarity	Low similarity
FN	264.7 (117.9)	273.5 (112.4)
JF	271.8 (149.5)	271.9 (171.6)
VP	233.3 (107.0)	260.0 (164.0)
Mean	256.6	268.5

Table 7.9: Fixation durations (ms) of the three subjects for high and low similarity of non-targets (standard deviations in parentheses).

7.9.3 Misses

Number of 'misses' described the number of incidences where eye movement analysis indicated that the subjects had fixated upon the target but had failed to perceive it. Due to the fact that some of the data from the 144 displays was lost, a percentage value was arrived at by dividing the number of targets fixated but not recognised by the total number of relevant items for which data were available, all multiplied by 100. Due to the relatively small number of arrays viewed within the nine different variations of the cycles variable, data were combined from the three subjects. Table 7.10 depicts the misses index for the nine cycles employed.

Cycles	1	1.5	2	2.5	3	3.5	4	4.5	5
Misses index	5	6	0	20	11	7	0	21	14

Table 7.10: Index of misses across the nine different levels of cycles (N=3).

Table 7.11 shows the index of misses in relation to the high versus low contrast condition. Differences again were not significant ($t = -3.355; df = 2$).

	Contrast	
	High	Low
FN	6	13
JF	0	8
VP	27	44

Table 7.11: Index of misses for the high versus low contrast condition.

Table 7.12 shows the index of misses in relation to the high versus low similarity condition. Differences on this variable were not significant ($t = -0.795; df = 2$).

	Similarity	
	High	Low
FN	15	10
JF	0	3
VP	31	44

Table 7.12: Index of misses for the high versus low similarity condition.

7.9.4 First saccades to target

As a further measure of efficiency of search, an index of first fixations to target was derived whereby number of instances where subjects located the target without fixating upon another item was divided by total number of presentations for which data were available.

Table 7.13 depicts the first saccades to target index for the nine cycles. Correlating (Pearson r) the search times (see Table 7.2) for the various spatial frequencies with this index produced a correlation of -0.626 which fell short of statistical significance (critical $\text{value}_{0.05} = 0.666$).

Cycles	1	1.5	2	2.5	3	3.5	4	4.5	5
First sacs.	40	44	44	40	17	36	42	21	43

Table 7.13: Index of first saccades to target across the nine different levels of cycles (N=3).

Analysing this index for high similarity and low similarity displays seperately, however, produced a different pattern (see Table 7.14). Figure 7.9 shows this difference graphically. The correlation between search time and the index for the low similarity arrays was -0.583 (NS), and the corresponding correlation for high similarity arrays was -0.596 (NS).

Cycles	1	1.5	2	2.5	3	3.5	4	4.5	5
High similarity	27	36	20	0	10	12	25	13	25
Low similarity	80	100	75	57	25	67	75	45	67

Table 7.14: Index of first saccades to target across the nine different levels of cycles when data are analysed seperately for high and low similarity.

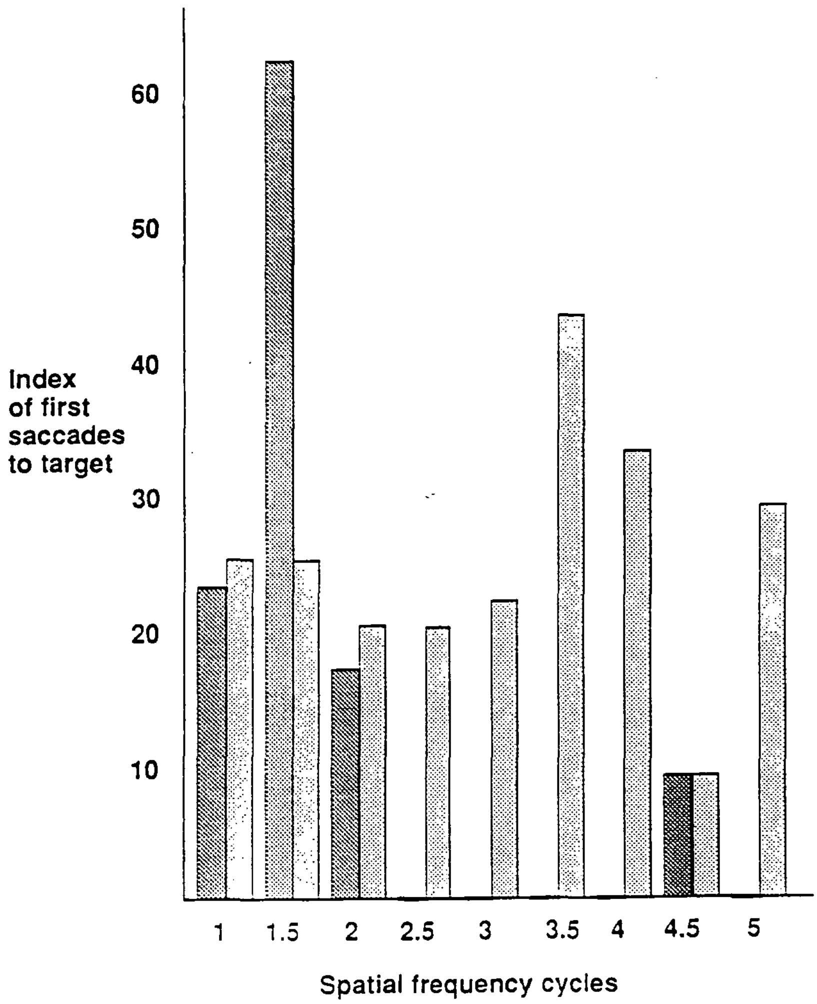


Figure 7.9: Plot of first saccades to target index seperately analysed for high (dark grey) and low (light grey) similarity.

Table 7.15 shows results from the first saccades to target analysis for the high versus low similarity condition. This variable also failed to show significant differences ($t = 1.509; df = 2$).

	Similarity	
	High	Low
FN	15	65
JF	18	56
VP	8	0

Table 7.15: Index of first saccades to target for the high versus low similarity condition.

Table 7.16 shows results from the first saccades to target index for the high versus low contrast condition. Differences between conditions on this variable were also non-significant ($t = -2.294; df = 2$).

	Contrast	
	High	Low
FN	35	36
JF	44	50
VP	19	22

Table 7.16: Index of first saccades to target for the high versus low contrast condition.

7.9.5 Number of items fixated before target

A third measure of visual search was the number of items (including 'missed' targets) fixated before the target was identified. As there were multiple instances within each array, the index is derived as a mean of all incidences of the arrays for which data were available. Table 7.17 shows this index in relation to the nine levels of cycles employed.

Cycles	1	1.5	2	2.5	3	3.5	4	4.5	5
Mean	0.968	1.625	1.454	2.200	3.000	2.429	1.667	2.545	2.282
sd	1.169	2.306	1.916	2.895	3.106	3.567	1.875	2.905	3.561

Table 7.17: Index of items fixated before the target was identified between the nine different levels of cycles (N=3).

Table 7.18 shows results from this analysis for the high versus low similarity condition. The difference between the indices from these similarity conditions was significant ($t = 8.041; df = 2; p < 0.05$).

	High similarity	Low similarity
FN	2.412 (2.076)	0.650 (1.226)
JF	1.778 (2.006)	0.625 (1.008)
VP	4.154 (3.674)	2.778 (3.114)
Mean	2.781	1.350

Table 7.18: Index of number of items fixated before successful location of the target for the high versus low similarity condition (standard deviations in parentheses).

Table 7.19 shows results from this index for the high versus low contrast condition. Again, these differences proved non-significant ($t = 1.418; df = 2$).

	High contrast	Low contrast
FN	1.721 (1.931)	1.409 (1.992)
JF	1.104 (1.627)	1.083 (1.676)
VP	4.154 (3.684)	2.889 (3.060)
Mean	2.326	1.794

Table 7.19: Index of number of items fixated before the successful location of the target for the high versus low contrast condition (standard deviations in parentheses).

7.10 DISCUSSION

7.10.1 Search paths

Figure 7.6 showed an example of a phenomenon commonly seen whereby the subject's first saccade is to a non-target (at NW) just a half-cycle removed from that of the target (at NE). It would appear that peripheral vision incorrectly draws the eye to the non-target. Within such cases, the subject seems to quickly realise their error, presumably as the pattern being observed does not match the pattern currently held in visual memory, and then the subject would re-view the reference item (C) and immediately locate the target. Although no precise analysis was performed, the experimenter is beyond doubt that this phenomenon did not occur with non-targets most like the actual targets in the *low* similarity condition.

A systematic search pattern (of an anti-clockwise direction) is shown in Figure 7.7. With this particular subject (FN), the systematic (circular) searches were of equal proportions in anti-clockwise and clockwise directions. It is of note that systematic does not always imply short search times, as in this case a systematic path could not have been longer.

Figure 7.9 provides an example of an unsystematic search for a target located at *S*. From (a) a re-run of the data analysis by the computer as shown on a display screen and (b) the hard copy print-out of the data analysis showing horizontal and vertical fixation coordinates, the sequence of fixations is clear. From the central reference item, the subject fixates on *N*, then *NW* and *W* before checking the referent. Such returns to the centre were often seen both during long searches where the subject is possibly refreshing their memory image and also by way of making a close comparison when non-targets very similar to the target are encountered. This latter phenomenon possibly accounted for some of the greater overall search times found for the high similarity condition. The subject then saccades to *E*, and then *SE*, having omitted a look at *NE* in the upper half of the display. A saccade is then made to *SW* missing *S* where the target actually is. It appears, however, that the target may have been peripherally perceived at this point as there is an immediate reversal of the eye movement direction back to *S*. The key is then pressed just as the eyes return to the central point.

7.10.2 Fixation durations

There was no discernible pattern of fixation durations across the cycle frequency variable. Despite the fact that a clear pattern was found for search times, this does not appear to have a correlate in this eye movement parameter.

The differences in fixation durations between high contrast and low contrast items and between high versus low similarity of distractors were not significant. Nor were the corresponding differences for the high/low similarity variable. In fact, whereas one might expect the high similarity condition to be characterised by longer fixation durations (on the grounds of lower discriminability), the opposite trend was found for all three subjects (high similarity $\bar{X} = 256.6$; low similarity $\bar{X} = 268.5$).

To summarise results from the analysis of fixation durations, it must be said that no differences in eye movements were found to account for performance superiorities in the locatability of different spatial frequency parameters.

It is, perhaps, unfortunate that the programme did not sample for saccade amplitudes. However, it is felt that similarly neutral findings would have emerged on this measure.

7.10.3 Misses index

This described the incidence of situations where the subject had fixated on the target yet failed to perceive it. The pattern shown between the nine cycle levels is best described as erratic. There was no 'saw-tooth' effect between the half-cycle and full-cycle frequencies.

Tables 7.11 and 7.12 show results from this index over the contrast and similarity dimensions. In neither case did the subjects show a unified pattern between the high and low conditions.

7.10.4 First saccades to target

Again, there was no uniformity across the three subjects as far as the direction of difference between high and low was concerned on both the similarity and contrast variables.

The correlation between the index of first saccades to target and search times on the nine frequencies was not significant.

Considering that by chance alone an index of 12.5 would be expected (one target amongst eight items), then it is clear that only the very lowest frequencies (1 and 1.5 cycles) are substantially greater than this.

Performing seperating analyses on high similarity and low similarity arrays revealed marked differences. As Table 7.14 and Figure 7.9 show (with perhaps the exception of the 4.5 cycle results), whilst the index is fairly consistent across the low similarity arrays, the index pattern across the high similarity arrays is clearly positively skewed with almost all cycle frequencies above 2 producing an index of zero. This is supported by the finding that whereas low similarity items and search times are not correlated, high similarity arrays and search time are negatively correlated; the lower the spatial frequency the faster the search. It would seem that discriminability underlies the interaction between the similarity and cycles variables shown in Figure 7.1.

7.10.5 Number of items fixated before target

Although this index was clearly lower on the 1 cycle condition ($\bar{X} = 0.968$; $sd = 1.169$) compared with the 5 cycle condition ($\bar{X} = 2.282; sd = 3.561$), intermediate indices showed no systematicity.

Although the differences in this index was not significant on the contrast variable, there was consistency shown in that all three subjects showed higher ratings on the high contrast as opposed to the low contrast condition. However, on discriminability grounds, it would be expected that targets with high contrast would be more readily located (i.e. there would be a *lower* number of items fixated before the target was identified).

The single analysis showing significant results ($p < 0.05$) was that of the number of items fixated before target between the high/low conditions of the similarity variable. The index for the high condition ($\bar{X} = 2.781$) was more than twice that of the low condition ($\bar{X} = 1.350$). The most likely explanation for this is in terms of the greater discriminability of the target from the low similarity items resulting in lower search times and attendant lower number of items fixated before the target, although this might be relatively independent of peripheral vision factors.

7.11 CONCLUSIONS

As only two of the twelve analyses of eye movements demonstrated a significant result, it must generally be concluded, in line with the null hypothesis, that the search times found for the three spatial frequency variables associated with the icon search superiority have no correlate in eye movement parameters. Icons are certainly more varied in terms of area covered and shape than are words, but such differences are unlikely to account for their consistent and reliable visual search superiority. To an extent, an increase in spatial frequency is an increase in *complexity;* but this sort of 'explanation' is rather circular. It would appear though that there is an attribute of icons which results in them being located perhaps 'just a saccade away', and that attribute may well be the ease in peripheral perception of their lower spatial frequency composition.

Chapter Eight

Do Antialias Fonts Facilitate Search?

8.1 OVERVIEW

As was emphasised in §1.10, a valid examination of optimising visual search at the VDU must encompass various parameters which include, in addition to those already explored, how recent advances in font design may enhance the locatability of items and, on a higher level, how these new technologies might aid reading *per se.* The impetus behind this relatively 'applied' study came from IBM(UK) Ltd, one of several organisations concerned with the development of 'antialiased' fonts. It was intended, therefore, that this final study would have potential applications for ongoing commercial development. Two experiments are reported. Results of the first were inconclusive; possibly due to artifacts of the fonts used. An index of density was proposed which supports the suggestion that previous reports showing performance enhancements with antialiased fonts were seriously flawed by truly equivalent fonts not having been employed. A second study was carried out using the search coil technique. Both reading and search tasks were employed. Differences in eye movement parameters were found across the fonts used and these are discussed in terms of the proposed density index and fonts generally. In addition, data from a probability vector and transition matrix analysis are described. As far as the search task was concerned, results seemed to suggest that the number of diagonal saccades varies increasingly with decreasing font density.

8.2 GENERAL INTRODUCTION

One hurdle which VDU engineers have struggled to contend with is the perceptual superiority of hardcopy text over screen-displayed material. Whether reading material from a VDU and searching for textual material on a screen is more difficult than with conventional printed material is obviously an important issue. In their 1988 paper tellingly entitled 'Reading from screen versus paper: there is no difference', Oborne and Holton set a benchmark: 'If the advantages offered by new technology are to be accepted, its operation must present no additional obstacles to the operator than the technology it replaces. To ensure that a screen-based reading medium is accepted, then, it is important that reading and comprehending information from a VDU is as easy and as efficient as it presently is using paper' (pp. 1–2). It might be supposed that if the line length, level of detail etc. of the information is presented similarly then reading and comprehension from paper and VDU must be equivalent, but results from several comparative studies do question such a proposal. This issue has been well reviewed by Oborne and Holton (1988) and below a brief overview of their argument is provided.

The fundamental point is that a simplistic comparison between VDUs and paper just does not hold. Operator posture is likely different as VDUs are generally vertical and paper typically laid horizontally; reading distances are often different; contrast between characters and background is probably different; paper generally has positive contrast whilst conventional screens tend to be of negative contrast; and the character set, resolution, justification (left versus full), and familiarity with the medium are also likely to be different.

Numerous studies have addressed these issues yet conclusions still remain equivocal (Oborne & Holton, 1988). At least eleven studies (Kak, 1981; Muter *et al.*, 1982; Gould & Grischkowsky, 1982, 1984; Wright & Lickorish, 1983, 1984; Kruk & Muter, 1984; Mills & Weldon, 1984; Heppner *et al.*, 1985; Creed *et al.*, 1987; Gould *et al.*, 1987*a*) support the screen disadvantage; seven (Switchenko, 1984; Askwall, 1985; Cushman, 1986; Bender *et al.*, 1987; Gould *et al.*, 1987*b*; Oborne & Holton, 1988; Miller *et al.*, 1989) do not; while one (Newsted, 1985) indicates a screen advantage.

As one example, controlling for a number of critical variables that normally exist between the two formats, such as number of characters per line and lines per page (which cumulatively bias towards a paper format), Switchenko (1984) demonstrated no real difference between screen and paper. Additionally, Kruk and Muter (1984) reviewed the contrast difference which exists between screen and paper. For example, Timmers *et al.* (1980) have demonstrated increased word recognition speed with reduced contrast ratios between characters and their backgrounds. However, in the Kruk and Muter study neither the time taken to fill the screen nor the variations in contrast ratios caused significant differences in reading speed on the screen. Varying the distance from the screen (40 cm, 80 cm, 120 cm) was also found to produce no significant differences in reading speed. These results suggest that although there is no single cause for slower reading in the screen condition, at least two factors play important roles: (a) format, which includes numbers of characters per line and lines per page; and (b) inter-line spacing. Nevertheless, this team suggests that the slowing down of reading speed due to these two factors alone was insufficient to account for all differences between the two media.

Of those factors mentioned above, Oborne and Holton emphasise image polarity, i.e. dark characters on light background or vice versa, because in this case the available evidence seems to present fairly clear-cut results. Dark characters on a light background (a positive image) are read quicker (Radl, 1980), more accurately and are more acceptable (Bauer & Cavonius, 1980) than are the light characters on a dark background generally seen on VDU screens. However, they do cause slightly more visual fatigue (Cushman, 1986), perhaps due to increased flicker (Isensee, 1982). It may well be the case that the screen disadvantage obtained in some experiments simply reflects the fact that by using negative appearing characters the screen is at an immediate disadvantage. Cushman has provided support for this contention in that no significant difference for reading speed was demonstrated when the image polarity was the same for paper and VDU presentations.

Oborne and Holton considered this issue again whilst attempting to control as many variables as possible including subject's posture, distance from the screen, image polarity, line length and page layout. No significant differences were found in either reading speed or comprehension between screen and paper, or between dark and light character displays. This is a particularly important finding as it allows future developments of electronic media to progress without fear that reading performance may be degraded. Results were interpreted in terms of viewing position and reading time. Firstly, in studies which had obtained a screen disadvantage the subjects with the paper format presentation had been allowed to change the viewing position of the paper presentation, and could thus vary the distance between eyes and texts, vary the angle of the paper, use fingers to mark their place, and possibly adjust the amount of light falling on the text by turning the paper away from the light source. Secondly, it may be that any screen–paper difference occurs only with longer duration reading periods. This suggestion is supported by the observation that other studies which have demonstrated no screen disadvantage also employed shorter reading periods (e.g. Askwall, 1985). Since the range of tasks which might be performed in the electronic office of the future would likely include both long- and short-duration reading periods, this would be an important question to test empirically.

Finally, Oborne and Holton raise the pertinent question: What sort of text is likely to be presented on VDU screens in the future? As the electronic presentation of text becomes more pervasive, questions need to be raised along the lines of: Do individuals perform better using screen or paper when the material is presented in the most optimum format for that medium?

8.2.1 Previous research

At this point, two terms—'aliasing' and 'antialiasing'—need to be introduced. Technically, the term aliasing refers to the problem arising from a failure to accurately reproduce a signal from digital samples. A signal can be faithfully reproduced from such samples only if the highest frequency in the signal does not exceed one-half the sampling frequency (Oppenheim & Schafer, 1975). Most CRT displays are raster displays that typically feature dot matrix characters. On close-up examination, at the very least, it is possible to notice the jagged edges or lines that appear to contain 'staircasing'; the result of aliasing due to an undersampling of the signal that would be required to produce sharp, continuous characters.

To reduce the effects of aliasing, 'greyscaling' may be used (e.g. Sholtz, 1982). This form of antialiasing employs intermediary shades of grey at appropriate pixel points between the extremes of black and white in order to 'smooth out' the staircasing of characters seen with conventional 'bi-level' or 'binary' (e.g. purely black pixels on a white background) fonts. Antialiasing provides the character set designer with the ability to more finely control the stroke width of a character (and not be constrained by the pixel addressability). Hopefully, this creates characters more similar to high quality print. Antialiasing can, however, reduce the sharpness of the characters (Kennedy *et al.*, 1987; Kennedy & Oakley, 1989).

Three antialiasing techniques are available: (a) averaging several pixels from a high-resolution image into a single pixel in a low resolution image, (b) filtering from a high-resolution image to produce a low-resolution image, and (c) prefiltering at the display resolution. This last form of antialiasing technique is that employed within the system with which this report is concerned. There are implications for both graphics design and text font design. With AA techniques, a typographer can design VDU typefaces that are close visual matches for some of the well known classical printed designs.

8.2.2 Previous research on antialiasing

The literature on antialiased (AA or greyscale fonts has expanded rapidly in the last few years (Lee & Rabideau, 1979; Leler, 1980; Schmandt, 1980, 1983; Warnock, 1980; Crow, 1981; Sholtz, 1982; Gupta, 1986; Turkowski, 1986; Bender *et al.*, 1987; Booth *et al.*, 1987; Gould *et al.*, 1987*b,c*; Cowan, 1988; Cushman & Miller, 1988; Ferwerda & Greenberg, 1988; Naiman & Farrell, 1988; Krantz & Silverstein, 1989; Miller *et al.*, 1989; Weiman & Perrin, 1989; Silverstein *et al.*, 1989). Gould's team has published an on-going series of studies investigating differences in the readability of CRT displays and paper, the titles being self-explanatory: 'Reading is slower from CRT displays than from paper: attempts to isolate a single variable explanation' (Gould *et al.*, 1987*a*), and 'Reading from CRT displays can be as fast as reading from paper' (Gould *et al.*, 1987*b*). In the latter study, CRT speeds equivalent to that of paper were obtained with the use of a high resolution monitor, positive polarity and with an AA font which closely resembled the font used on the paper. It was felt that the image quality of the characters was the key. 'No strong evidence' was found 'that anti-aliasing by itself' accounted for the improved CRT reading speeds; rather, 'it could account for part of the difference' (p. 510).

Initially, Gould *et al.* (1987*a*) conducted a series of experiments and analyses to identify the causes of this reading speed deficit. Typically, each experiment isolated one variable and studied whether it explained the difference. The principle result was that no single variable (e.g. experience in using VDU, display orientation, angular character size, font, polarity, different display units) explained why this reading speed difference occurs. The majority of these studies used proof-reading as the experimental task, yet the reading speed difference was found when participants read for comprehension also. Gould's team tentatively concluded that the difference was due to a combination of variables that could interact in non-linear ways and which affected the image quality of the characters themselves. Gould's results ruled out two other possible explanations: personal variables, such as age, experience, or familiarity with reading from VDUs; and possible inherent defects in CRT technology, such as flicker, alleged subliminal effects, or the fact that the CRT display is self-luminous rather than passively or reflectively illuminated.

The aim of the Gould *et al.* (1987*b*) study was to compare reading from paper and from CRT displays when the display appearance was similar to paper. Various ways were tried to achieve this similarity. With one approach, an electronic scanner was employed to scan pages of print and then display the results on a VDU, but satisfactory results were never achieved. In a further approach, they used a system that printed characters on paper that were similar to the way they appeared on the screen. However, the similarity was only approximate, and there were programming problems with the desired experimental set-up. This second paper described six experiments using the YODA system (an early AA system; Gupta, 1986) to display VDU characters that looked similar to those on paper. They had the same font, polarity (dark on a light background), size, colour, and layout of the two media. When transparencies were made of a page of print (240 dots per inch) and then placed over a screen, the match was 'quite accurate'. The VDU and system used in their series of studies had up to 16 different levels of luminance.

The results identified a set of conditions that, when present on VDUs, led to significantly faster reading; reading equivalent to that achieved with good print on paper. Concerning the fonts used, Gould concluded that the antialiasing by 'itself probably contributes somewhat to the reduction of the CRT/paper reading speed difference. Anti-aliasing eliminated the significant reading speed difference between paper (262 words/min) and aliased CRT characters (240 words/min). The purpose of anti-aliasing is to increase the perceived resolution of a display through the use of gray level. Based upon informal observations, we believe that anti-aliasing is of most help with relatively low-resolution displays and that as display resolution increases, anti-aliasing contributes less and less' (Gould *et al.*, 1987*b*, p. 514).

In addition, Bender *et al.* (1987) showed that the text on a VDU rendered in black, AA characters on a white background (positive polarity) could be read as rapidly from a VDU as text that was printed on paper with a letter-quality printer. Miller *et al.* (1989) then set out to determine if this same typeface showed the same performance characteristics when presented on colour on a VDU. An IBM Personal System/2 Color Display (Model 8514), which is an analog CRT display, presented the AA typeface in high-intensity white characters on a blue background. A typical office courier typeface was used to present letter-quality print on paper. The results showed that when proof-reading for spelling errors, no difference was found to exist between the reading speed or the number of errors detected for the two presentation media. It was concluded that when text is rendered as high-intensity white characters against a blue background on this system in the AA typeface, the text can be read as quickly and as accurately from the screen as it can be read from paper.

Hewitt *et al.* (1989) report a subjective comparison study in which participants compared texts displayed on monitors using AA fonts and using binary fonts with text printed on paper. The following variables showed an effect.

(1) **Monitor:** Two sizes of monitor were used; a 16 inch and a 14 inch, both displaying 1024 × 768 images (approximately 100 pixels per inch). AA fonts were chosen more frequently when text was displayed on the smaller monitor. This is possibly due to the smaller screen size/smaller pixel size/higher resolution of the 14 inch monitor reducing the likelihood of the AA fonts being perceived as fuzzy.

(2) **Font style:** Concerning font style, AA allows the designer of fonts a greater number of addressable points than actually exist within a character box. This advantage was detected by subjects when the Times Roman (serifed) fonts were compared. AA fonts were chosen far more frequently than the binary fonts. Such serifed fonts require larger character boxes, i.e. a larger number of addressable points in order to accurately reflect the true shape of the characters than do sans-serif fonts (e.g. helvetica). This reason may also explain the significant preference for the AA version of the small size (8-point versus 12-point). With smaller character boxes it is difficult to accurately reflect the true shape of even sans-serif fonts.

(3) **Age:** Participants fell within a 'young' (17–24 years) or 'old' ($\geq$ 45 years) *age*

group. The young participants more frequently reported the AA fonts as being fuzzy. Thus, at least in part, possibly due to *visual acuity*, the choice between the AA font and the binary font became a more difficult decision. The AA font was particularly beneficial when the text was displayed in Times Roman font and when displayed using a relatively small (8-point) font.

(4) **Polarity:** The AA fonts were designed and tuned for positive *polarity* presentation and when text was displayed with a small amount highlighted using negative polarity, subjects reported a substantial decrement in 'readability'. This was apparent despite less than 10 percent of the reading material area being reverse highlighted.

Although some participants preferred the lower contrast ratio between the character and background of the monitor-displayed text compared to that of the text printed on paper (9:1 versus 14/15:1), for most subjects the addressability differences of the two different media (monitors approximately 300 pixels per inch, laser printer 300 dots per inch) translated into perceivable differences in clarity. Paper-oriented text was more often perceived as clearer and sharper, and consequently preferred. 'Some participants at different times and to a different extent saw the AA font text as being *fuzzy and out of focus*' (Hewitt *et al.*, 1989, p. 2; italics added). Also, the binary font was mostly felt to be too *thin* compared to the bolder AA font. Such subjective reports are important in that they suggest a clearly visible difference (i.e. 'boldness' or 'density') between binary and AA fonts which may go beyond grey-scaling *per se*.

Use of the IBM Presentation Manager facility, 'Glass' (software which allows any area of the display such as a particular alphanumeral to be greatly magnified), readily confirms this notion. Marked differences in stroke width were apparent between the standard binary font and two alternative forms; the antialiased and its 'bold' version (here termed AA and bold-AA) fonts of one letter (helvetica style). The bold-AA clearly contained a greater degree of antialiasing than the 'standard' AA font. It could be seen that antialiasing consisted of more than pixels of intermediary grey along the staircasing of pure black and white tones. Stroke width and overall density are substantially increased.

Figures which are included in §8.9 clearly demonstrate this (the figures are placed in that section because there were slight differences between the fonts used in Expt. 12 and Expt. 13 and those specific examples relate to the latter study).

EXPERIMENT 12: THE ROLE OF DENSITY IN ANTIALIAS FONTS

8.3 INTRODUCTION

Several AA vs binary studies by IBM human factors engineers have examined variables such as reading speed and spelling mistakes in CRT vs paper presented text, various sizes of font (e.g. 8 point vs 12 point), various font types (e.g. Times Roman vs helvetica), various sizes and differing resolution monitors, positive vs negative presentations, as well as gathering subjective data (IBM, unpublished reports). Overall, results from this diversity of studies on the prototype AA facilities might be regarded as inconsistent and inconclusive. Also, a number of reservations may be expressed about the experimental designs used (e.g. occasions where only some fonts were proportionally spaced and instances where perhaps a non-AA font should have been used as a control but was not).

The present experiment aimed to investigate whether, and to what extent, antialiasing contributes to the legibility of alphabetic characters. To avoid cognitive factors such as comprehension arising, a pure legibility (perceptual) task (visual search) was chosen. It was hypothesised that search times for single target letters presented amidst randomly generated alphabetic arrays would be significantly faster for AA VDU presentations.

8.4 METHOD

8.4.1 Subjects

Time constraints allowed only nine (5M:4F) subjects (mean age = 25.5; sd = 7.7) to participate. All were employees of IBM (UK) Laboratories Ltd., and were accustomed to operating VDUs daily.

8.4.2 Equipment

Fifty-four randomly generated screen displays of upper case alphabetic characters were produced. Spacing between characters was a random 0 to 3 spaces. Each screen contained only one instance of the target letter. Nine different target letters were employed, of three different types: straight or characterised by horizontal and vertical lines (*E, H, I*), angular (*A, V, X*), and rounded (*C, O, S*). These nine characters appeared in both a serifed (Times Roman) and a sans-serif (helvetica) style of type and in three font styles (binary, AA, and Bold AA). Binary font would normally be understood as that typically encountered on standard VDUs, bold font as so but with a thicker stroke width, whilst AA font would usually be regarded as the binary version although with staircasing smoothed out. Hence 54 presentations ($9 \times 2 \times 3$) were used. Some editing was performed to ensure a balanced placement of the targets within the array across the variables involved. Presentation order was randomised both within and between subjects.

8.4.3 Procedure

Presentation of screens was made to subjects via a batch-file system using an IBM PS/2 with a 8514 high resolution (1024×768 addressability) 16 inch monitor. Arrays occupied the entire screen apart from top and bottom status bars. The horizontal desk illumination was about 500 lux. The screen black/white contrast ratio was approximately 9:1. Timing per presentation was automatically recorded and printed out on-line.

When the screen appeared subjects were required simply to search, using whatever strategy they wished, for the relevant target letter. This was revealed to subjects by a pack of cards, appropriately ordered, and lying directly in front of the subject. When the subject had located the target letter they informed the experimenter and verified its location by stating the characters immediately to the left and right of the target. As soon as the target had been found, the experimenter pressed the keyboard space bar which both registered with the timer and also activated the next screen. Subjects were informed that on finding one target they were to turn over the next card to reveal the subsequent target letter and to immediately proceed with that. Precise instructions to subjects are included as Appendix 8.1. From experience gained in previous studies (e.g. Expt. 1), a cut-off point of 180 seconds was made. If subjects had not located the target within three minutes, they were informed that they could move on to the next array.

8.5 RESULTS

A total of 88 instances (18 percent) were recorded of exceeding the time limit. For analysis, figures of 180 s were taken in these cases.

Table 8.1 shows search times for the various fonts.

	Binary	AA	Bold AA
Helvetica	93.2	98.8	95.3
	(67.2)	(61.2)	(66.2)
Times Roman	74.5	77.2	88.2
	(54.5)	(60.9)	(57.0)

Table 8.1: Mean search times (s) for nine subjects across 54 arrays (standard deviations in parentheses).

A three-way repeated measures ANOVA was conducted on the data. The variables examined were 'type' (i.e. helvetica vs Times Roman), 'font' (i.e. binary, AA or bold AA), and 'line' (i.e. straight, angular or curved). Table 8.2 summarises the results.

Factor	SS	MS	F	df	p
Type	1198.79	1198.79	2.93	1	*NS*
Font	1174.02	587.01	1.05	2	*NS*
Line	431.91	215.96	0.26	2	*NS*
Type x Font	929.87	464.94	0.51	2	*NS*
Type x Line	419.53	209.76	0.25	2	*NS*
Font x Line	10807.29	2701.83	2.69	4	*NS*
Type x Font x Line	1155.35	288.84	0.47	4	*NS*

Table 8.2: Results from the three-way ANOVA.

8.6 DISCUSSION

It was hypothesised that significant differences would be found in search times in that one or both antialiased fonts would be more rapidly located than the binary font. This was not shown to be the case. Analysis of results by a three-way repeated measures ANOVA showed no main effects or interactions at any level. In fact, the serifed font showed a pattern opposite to that expected, in that search times for the binary font were slightly quicker than for the AA font, which again had search times somewhat quicker than the 'more' antialiased or denser 'bold AA'. The trend found was not even linear for the sans-serifed font. As such, the two font types did not demonstrate similar patterns of search time differences.

Although the sample size used (N = 9) was only half the minimum hoped for, it is not claimed that this is responsible for these inconclusive findings. Other concerns are greater. For example, the possibility of true differences being 'swamped' by the incidences of exceeding the '180-second rule' is of note. By simply altering the parameter within the generating programme controlling the 0 to 3 blank spaces between characters, it would be possible to reduce the background noise and hence substantially decrease the probability of the time limit being exceeded. Although it is true that future steps could be taken to reduce the fairly high rate of cut-offs, these results are not likely attributable to this influence alone. Certain other factors too may have influenced the results. For example, as previous researchers such as Gould *et al.* (1987*c*) remark, as VDU resolution increases, antialiasing seems to contribute increasingly less to performance measures. The high resolution monitor employed may have partially obscured effects; it may have been 'too good for the job'. It may also be that such a 'pure' perceptual task is not the most appropriate to use in assessing performance of typical VDU users. Real tasks rarely exclude cognitive factors such as reading and comprehension. Lastly, it was unfortunate that the 'system proportional' font was not included as a non-antialiased bold alternative to the standard binary. A system proportional font is that where characters, whether narrow or wide, occupy equal horizontal areas; thus an 'i' would have more space on either side of it than, say, a 'w'. The fact that the others involved were not system proportional poses no great problem as a measure of the effects attributable to the system proportional factor *per se* can be seen by looking at performance on the letters such as *I* (where large gaps may be seen between these and the adjacent letters) and characters such as the *W* (where letters may almost 'touch').

It may also be the case that a ceiling effect contributed to the lack of significant results. The 18 percent of instances exceeding the time limit supports this possibility. It is therefore suggested that a more demanding search task should be employed in similar future research.

Some form of objective quantification of density seemed in order. Obviously, the 12- and 14-point binary fonts consisted of either black or white pixels. In the antialiased forms of this prototype there were some four readily discernible shades of grey, and there were possibly a further one or two shades also, although time constraints associated with equipment availability allowed only a fairly rapid appraisal. The aim was therefore decided to propose a provisional quantified index of density (ID). It was possible to award a black pixel five points, a pure white pixel a zero rating, and the intermediary grades of grey from 1 to 4 points; thus giving a six-point scale. For instance, an upper case standard binary sans-serif *I* of eight pixels high would therefore have an ID of 8×5 or 40 units. A bold *I* of the 'system proportional' style would cover two pixels horizontally and therefore have an ID of $8 \times 2 \times 5$ or 80 units. It was therefore quantitatively twice as dense, and to the naked eye would appear clearly bolder. Taking the standard binary *A* which consisted of 29 black pixels and attributing each of these with 5 points, the ID value is 145. The bold binary *A* consisted of 46 pixels of value 5 and therefore had an ID of 230. However, not all lines are of double width. It is also of interest that this bold character was actually smaller (narrower) than the 'equivalent' standard binary *A*, occupying a space of 12×10 pixels (120 pels^2) as opposed to the standard binary's *A* of 12×11 pixels (132 pels^2). Looking at the AA *A*, it could be seen that this again occupied different space dimensions, being taller than the others, but relatively narrow in width ($13 \times 11 = 143$ pels^2). So, whilst narrower than the bold character, it is clearly much denser. Its ID would, counting the values 1 through to 5 (including four shades of grey) of all pixels occupied, would be 230; identical to that of the binary bold font. Further up the scale is the bold antialiased *A*. Although occupying the same number of pels^2 as its non-bold sibling, its ID is a mighty 334. Although the equivalence in ID between the binary bold and antialiased helvetica upper case *A* is not likely to hold exactly across the alphabet or other font types, and although the possibility of other subtle grey scales in addition to the four considered might alter the ID, something approximating this parity would be expected. This therefore gives more weight to the suggestion that the appropriate font styles to compare would be binary, bold and antialiased. If any difference in performance was to be found between the AA font and the others, this could not be attributable to density or boldness (as measured by the ID) and may well lie in the inherent qualities of antialiasing.

8.7 CONCLUSIONS

Suggestions for new directions can be proposed. In the light of findings from a 'rough and ready' quantitative assessment of density (the ID), it would seem particularly fruitful to compare a binary vs bold vs antialiased fonts in an attempt to ascertain the contribution of antialiasing *per se* to any performance enhancement for VDU users.

It may be advisable therefore, to modify any future work involving this particular type of visual search task so that:

(1) Only one font type is used—probably the less complex sans-serifed helvetica.
(2) Background noise is set such that search task difficulty is reduced to limit likelihood of a ceiling effect obscuring differences.
(3) A mid-range monitor may provide more scope for benefitting from the fine tuning of antialiased fonts.
(4) The task chosen is more akin to those encountered in the 'real-world'.
(5) As it is held that the boldness or density of the characters (e.g. stroke width) may account for more of any performance enhancement than antialiasing *per se,* it would be imperative to include a bold font as one of three fonts to be appraised along with standard binary and standard (non-bold) antialiased. It would then be hypothesised that direction of performance would increase from binary through antialiased to bold (i.e. 'system proportional') or, possibly from binary through bold to antialiased.

Finally, some meaningful quantitative measure is required to compare fonts and to account for differences in readability or legibility. Due to the pattern of results found within the present experiment it is not possible here to formulate an equation within which an index of density may be placed to produce some measure of readability or legibility. However, if the methodology is modified to produce a linear (or even curvilinear) algebraic formula consistent with the direction of difference held by current beliefs, then it may become possible to predict a figure for performance on diverse font types and styles on this equation.

It was expected that any difference in behavioural measures relating to fonts would be explicable in terms of the ID measure.

EXPERIMENT 13: AN EYE MOVEMENT STUDY OF ANTIALIASED FONTS

8.8 INTRODUCTION

Due to ongoing product developments within the IBM Corporation, it was not possible to gain further access to the prototype software employed within Experiment 12. As a somewhat different font system was to be employed in this study, it seemed that the search tasks used might also be diversified. Whereas a timed visual search study was used before, use was made in this study of eye movement parameters. Spatial and temporal performance measures were chosen for study as manifested during: (a) a reading task, and (b) a classic visual search type task where the subject must locate a target (letter X) within a rectangular array of random letters. Specifically, saccade and regression amplitudes and fixation durations were to be recorded. Search time *per se* was not taken as a dependent variable. One of the aims of this next study was to develop a more precise quantitative ID and to use this as a means of comparison between fonts.

The programme Glass was again able to show differences between the various fonts available. Figures 8.1 to 8.4 depict the marked differences in stroke width of (a) the standard 12-point binary, (b) the 12-point AA equivalent, (c) the bold binary 12-point equivalent, and for comparison (d), the 14-point binary font. It can be seen by comparing Figures 8.1 and 8.2 that antialiasing consists of more than pixels of intermediary grey along the staircasing of pure black and white tones. Stroke width and overall density are substantially increased. Further examples may be found in Appendix 8.2.

The relevant device driver on-line guidance (IBM, 1990) states: 'Conventional text on display screens is displayed as an array of dots or picture elements (pixels or pels). Each pixel is displayed in one of the two colors: a foreground color—the color of the text characters—or a background color. With antialiased text, pixels can be displayed using a range of color shades (up to 8 in this implementation) that can have values intermediate between the foreground and background color. Antialiased fonts are designed to make use of this capability to improve the appearance and readability of text.'

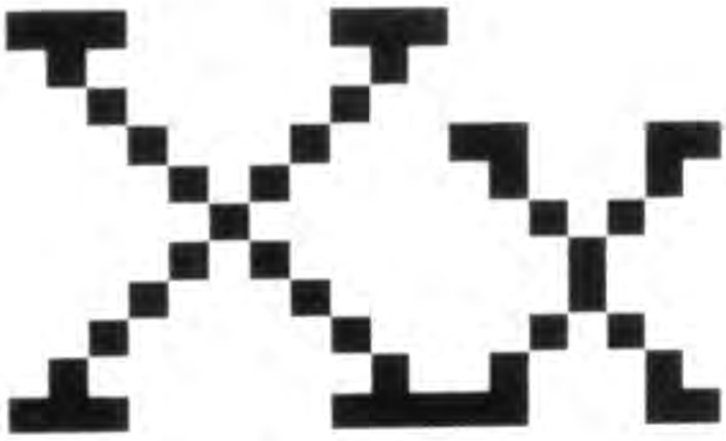


Figure 8.1: Enlarged example of the actual pixel composition of a standard 12-point binary character.

8.9 METHOD

8.9.1 Subjects

Five people participated. All had previous experience wearing a scleral search coil. They were known to have no uncorrected visual impairments.

8.9.2 Equipment and materials

The computer used was an IBM PS/2 model 80 fitted with an XGA Motorway card. This card, along with the necessary software installations, was responsible for generating the particular fonts used. This machine was connected to an IBM 8614 high resolution (1024 × 768 addressability) 16 inch colour monitor. As in the previous experiment, the array size was almost 16^2 inches. The entire system used is commercially available.

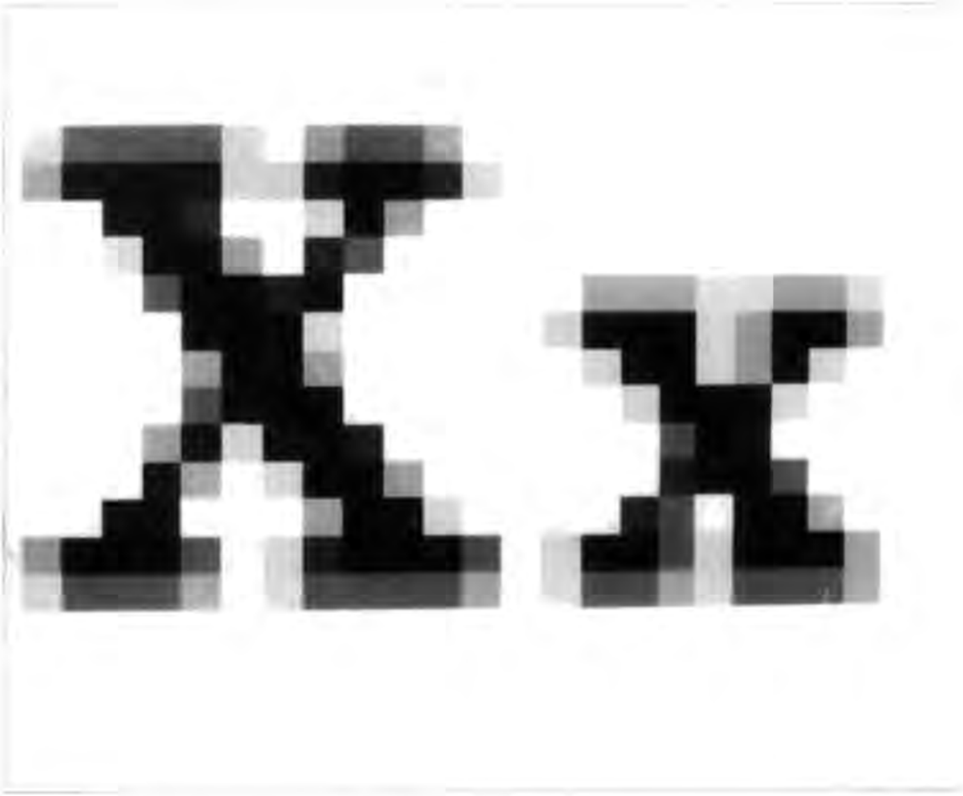


Figure 8.2: Enlarged example of the actual pixel composition of an antialiased 12-point character.

The following fonts were employed:

(a) Times Roman binary 12 point (hereafter termed TR12),
(b) Times Roman antialiased 12 point (TRAA12),
(c) Times Roman bold 12 point (TRB12), and
(d) Times Roman binary 14 point (TR14).

Reading material comprised of single full screens of one of four passages extracted from 'Gulliver's Travels' and may be described as of fairly advanced reading standard. Each subject received one of these screens in one of the four fonts (four presentations). Presentation order of fonts and of the reading passages were randomised between subjects. Figure 8.5 gives an example of such a display in 12-point binary font. Examples of the other three fonts are provided in Appendix 8.3.

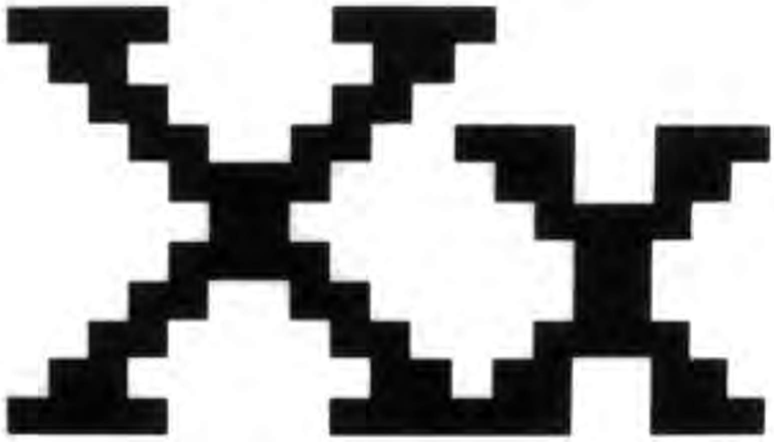


Figure 8.3: Enlarged example of the actual pixel composition of a bold 12-point character.

Visual search material consisted of four seperate screens of randomly generated *upper case* (only) alphabetic characters. Figure 8.6 gives an example of a screen display with 12-point binary font. Examples of the other three fonts may be found in Appendix 8.4. Within each there was only one letter *X;* the target letter. Again, presentation order and font by individual display was balanced between subjects.

The VDU was situated 40 cm from the subject's eyes, with the centre of the screen at about eye level. Although there were slight differences in the VDU illumination, depending upon which of the four fonts was displayed, illumination ranged from 58.0 to 61.3 cd/m^2. Ambient room luminance was approximately 150 lux, measured horizontal to the screen.

Figure 8.4: Enlarged example of the actual pixel composition of a standard 14-point binary character.

8.9.3 Procedure

Reading Presentation order of fonts and of the reading passages were randomised between subjects.

Visual search Again, presentation order and font by individual display was balanced between subjects.

Fine eye movement analyses using the scleral coil method were employed to compare reading and visual search performance manifested during the two tasks (eight screens in total).

Photometer (Minolta luminance meter) readings of the ten shades (black, white and eight grey-tones) were taken by enlarging a character with Glass so that each pixel was represented on the screen at a size of about 1.5 cm^2. At least three readings were taken of each tone and any differences obtained were averaged out.

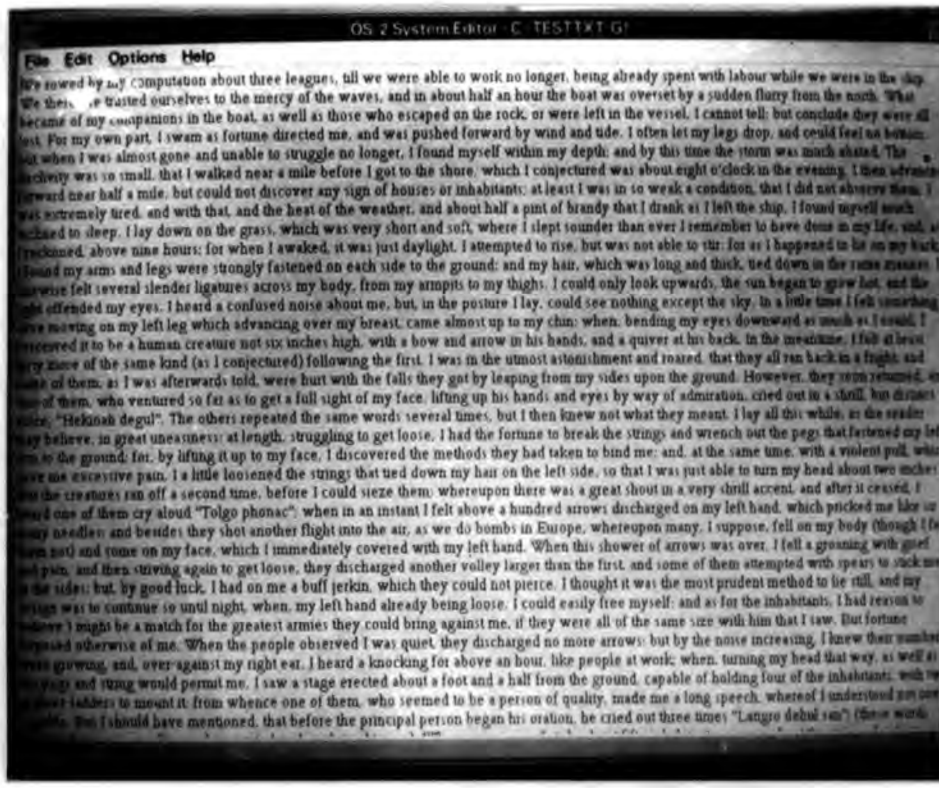


Figure 8.5: Example of reading material (12-point binary font).

8.10 RESULTS

8.10.1 Density index

A sample of nine characters (both upper and lower case) were chosen for analysis. These consisted of three characters typified (in upper case) by horizontal and vertical lines (*E, H, I*), three typified by angular (diagonal) lines (*A, V, X*), and three typified by circular lines (*C, O, S*).

An ID for these eighteen characters was awarded for the four fonts examined (72 scores). This was arrived at by awarding nine units for pure black pixels, 0 units for pure white (e.g. the centre of an 'O'), and from 1 for very light grey through to 8 for extreme dark grey. These scores are summarised in Table 8.3.

Figure 8.6: Example of visual search material (12-point binary font).

Table 8.4 shows the ID ratings for the letters as characterised by horizontal and vertical lines, angular lines, and circular lines. The combined rating

$$\frac{(uc + lc)}{2} \tag{8.1}$$

is produced simply from the mean of both the overall density index for both upper and lower case figures for the four fonts, divided by two.

Lower case	a	c	e	h	i	o	s	v	x	Mean	sd
TR12	153	117	171	243	108	162	90	162	162	152.0	44.6
TRAA12	317	196	220	348	201	244	172	206	244	238.7	58.4
TRB12	270	198	261	423	189	468	162	288	288	283.0	103.2
TR14	360	261	297	459	225	306	261	243	306	302.0	71.4
Upper case	A	C	E	H	I	O	S	V	X	Mean	sd
TR12	261	234	297	342	135	270	225	216	261	249.0	57.6
TRAA12	454	338	442	529	279	473	437	385	447	420.4	75.1
TRB12	432	369	423	522	252	288	378	396	450	390.0	82.1
TR14	405	351	450	594	261	468	405	369	486	421.0	94.1

Table 8.3: Index of density ratings for the nine letters (upper and lower case) used and means (N = 3).

	Horiz/vert	Angular	Circular	(uc+lc)/2
TR12	216 (93.2)	202 (50.5)	181 (69.1)	200.5
TRAA12	336 (129.3)	342 (104.1)	310 (126.4)	329.5
TRB12	345 (129.3)	354 (81.0)	310 (116.6)	336.5
TR14	381 (142.8)	361 (83.1)	342 (82.9)	361.5

Table 8.4: Index of density ratings for those letters characterised (in upper case) by horizontal and vertical lines; E, H, I), angular (diagonal) lines (A, V, X), and by circular lines (C, O, S); and a combined rating ('(uc+lc)/2') across the four fonts.

Figure 8.7 provides a simple graphical depiction of these differences. Here, the values show the combined means of the lower and upper case indices.

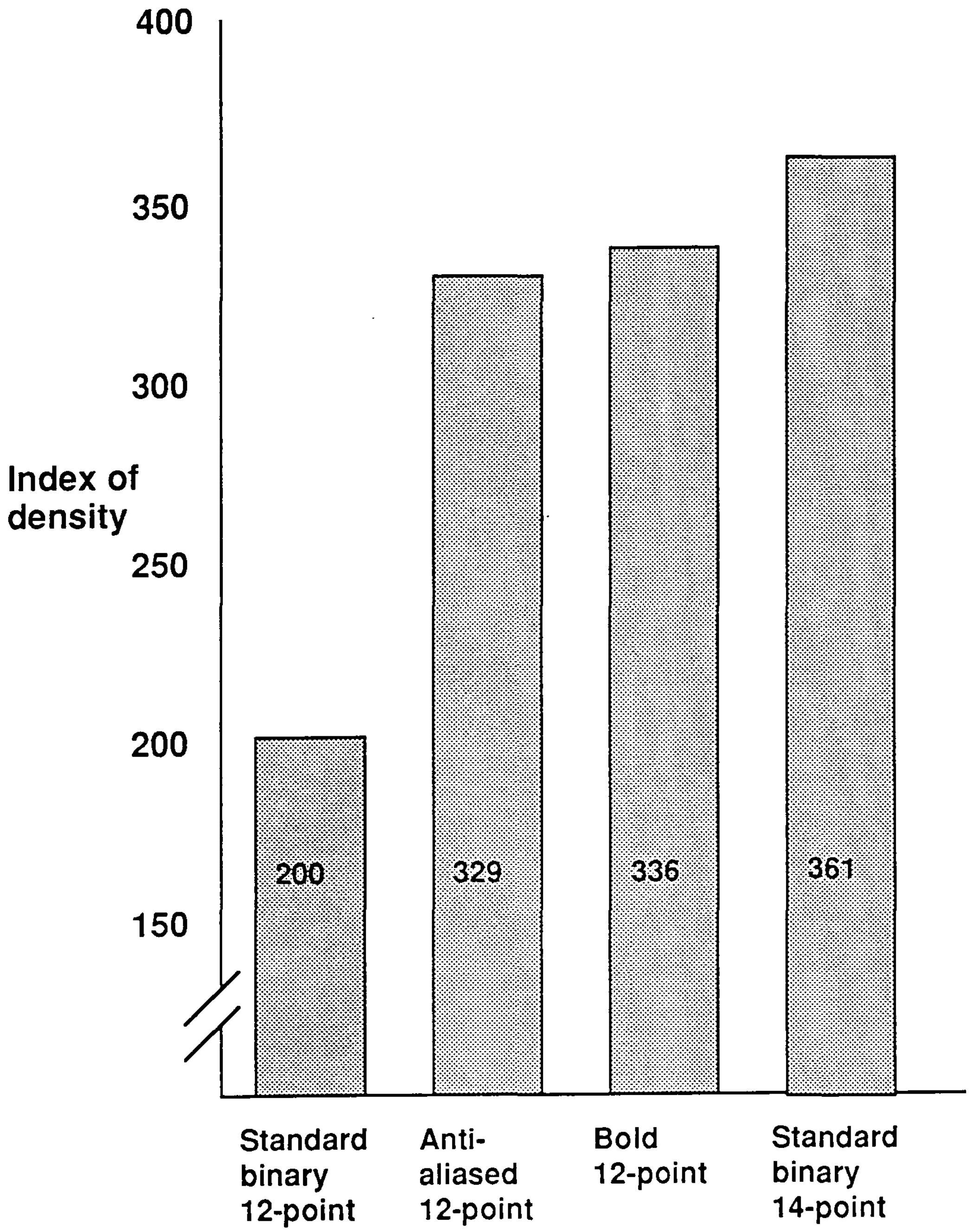


Figure 8.7: Histogram of the combined means of the indices of density for upper and lower case letters.

Figure 8.8 shows luminance ratings (cd/m^2) for the ten tones. It can be seen that with the exception of pure white (which, in any case, did not contribute to the numerical value) and of the extreme light grey, they do follow a linear gradient.

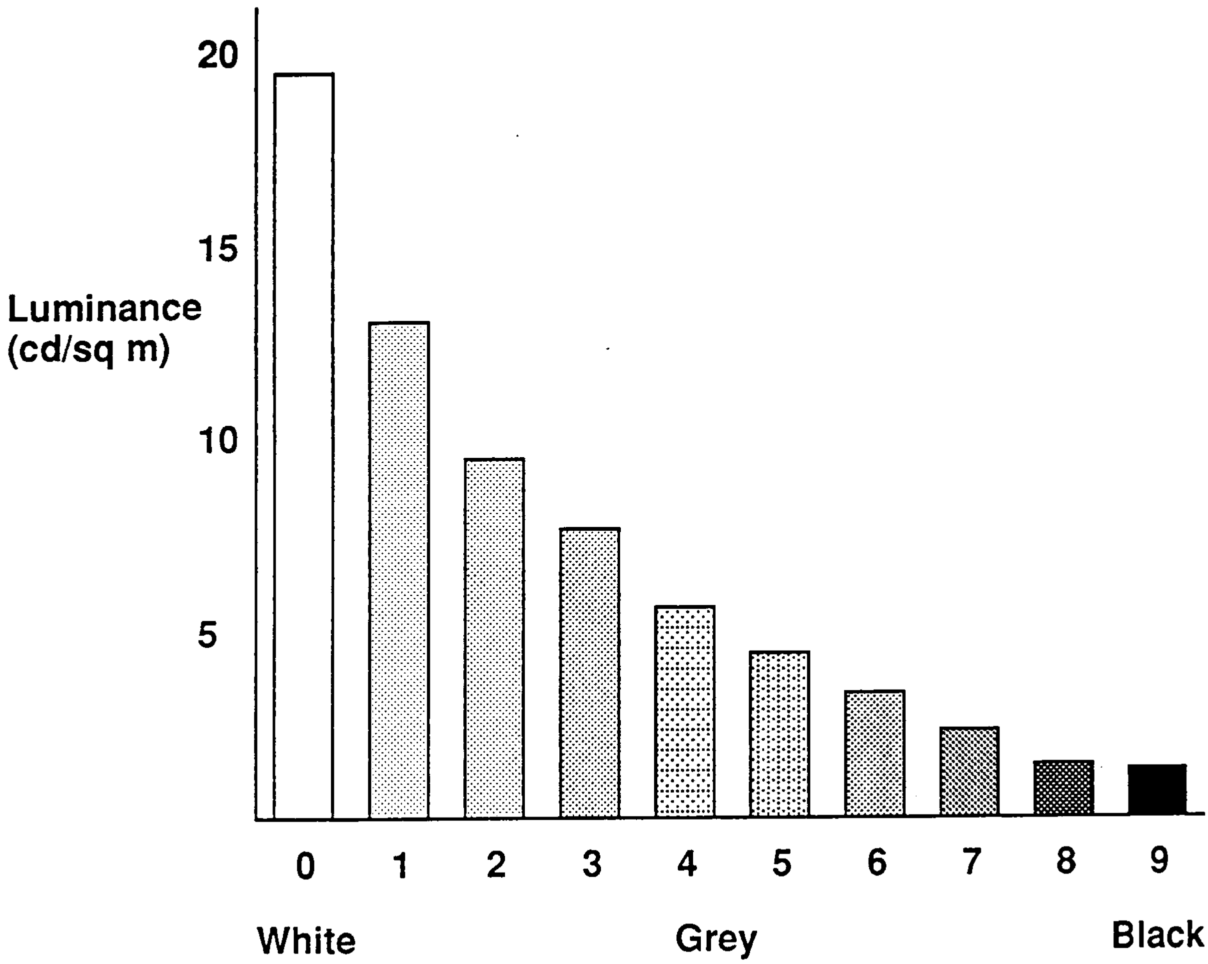


Figure 8.6: Luminance ratings (cd/m^2) for the black, white and grey scales comprising the antialiased font.

8.10.2 Eye movement analyses

Data from the eye movement recordings are summarised in Table 8.5.

	Reading			Search
	Fixations (ms)	Saccades (char. spaces)	Regressions (char. spaces)	Fixations (ms)
TR12	257.3 (21.2)	11.0 (4.5)	7.5 (4.0)	299.7 (20.2)
TRAA12	257.0 (16.7)	8.4 (2.5)	5.5 (1.5)	297.5 (26.1)
TRB12	266.0 (33.3)	8.5 (2.6)	6.5 (3.5)	313.2 (28.2)
TR14	261.4 (25.5)	8.3 (3.1)	5.5 (2.1)	315.6 (35.3)

Table 8.5: Eye movement parameters recorded over reading and search trials for the four fonts (standard deviations in parentheses).

Data were not obtainable (due to a search coil breaking *in situ*) from one subject (AH) on the TR14 reading screen and from all four search screens. Two methods of calculating results were therefore possible. Firstly, it was appropriate to conduct ANOVAs with only three fonts (omitting the incomplete data on TR14) for all five subjects. Alternatively, it was also justifiable to calculate statistical values with only four subjects (excluding AH) on all four fonts.

Excluding the TR14 font, there was a significant difference ($F(2,8) = 5.39; p < 0.05$) between saccade amplitudes on the reading task. Table 8.6 shows the means across these three fonts for this condition. The Newman-Keul's test showed significant differences to lie between TRAA12 and TR12 ($q = 12.82; p < 0.05$) and between TRAA12 and TRB12 ($q = 12.53; p < 0.05$). Saccade amplitudes were larger for the standard binary font.

TR12	TRAA12	TR12B
11.0 (4.5)	8.4 (2.5)	8.5 (2.6)

Table 8.6: Mean saccade amplitudes (as measured by character spaces travelled) and standard deviations (in parentheses) for the reading task from five subjects on three fonts.

Excluding the one subject, there was again a significant difference ($F(3,9) = 4.04; p < 0.05$) between saccade amplitudes on the reading task. These differences lay between TR12 and TRB12 ($q = 10.26; p < 0.05$) and between TR12 and TR14 ($q = 11.18; p < 0.05$). Saccade amplitudes were again larger for the standard binary font (see Table 8.7).

TR12	TRAA12	TR12B	TR14
12.0 (5.2)	8.3 (2.8)	8.5 (3.0)	8.3 (3.1)

Table 8.7: Mean saccade amplitudes (as measured by character spaces travelled) and standard deviations (in parentheses) for the reading task from four subjects on four fonts.

Figure 8.9 shows the mean saccade amplitudes for the four font types.

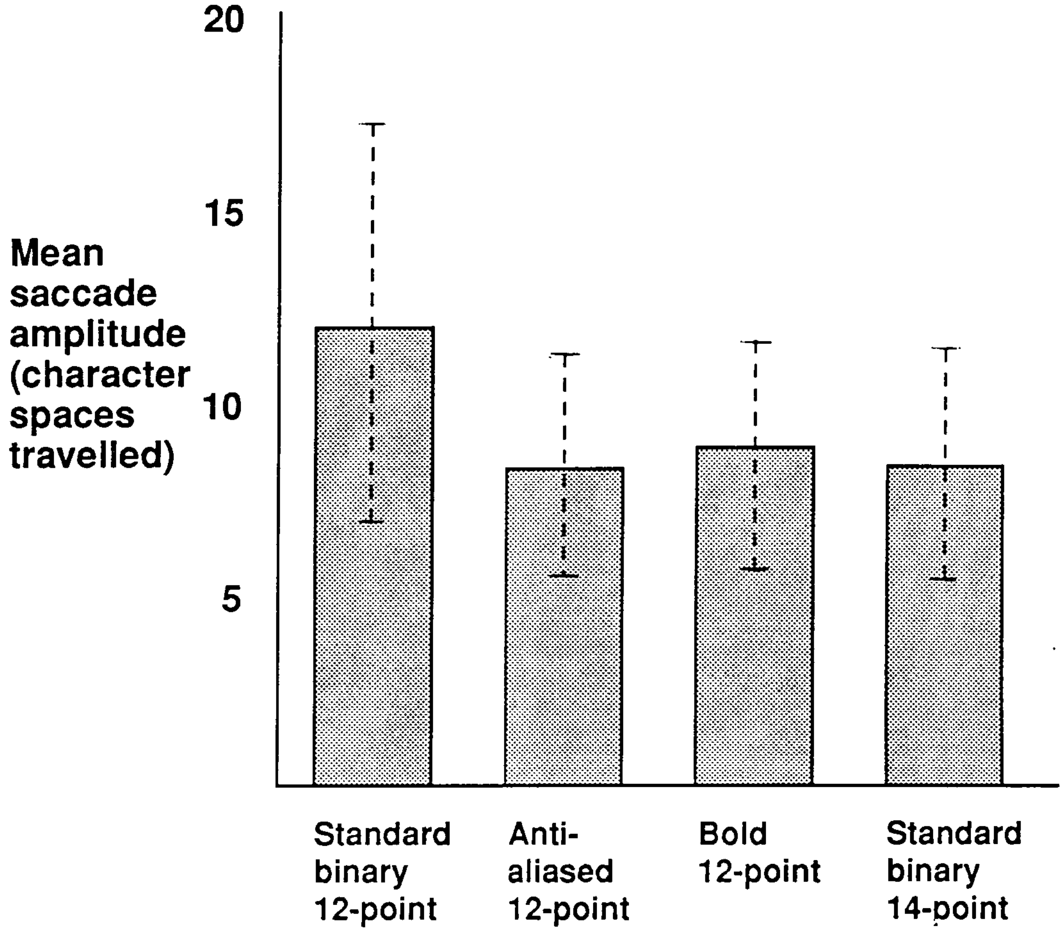


Figure 8.9: Histogram of mean saccade amplitudes (character spaces travelled) for the four fonts (vertical lines show standard deviations).

Analyses of regression amplitudes within the reading task failed to show significant differences whether excluding the font ($F(2,8) = 2.023$) or the subject ($F(3,9) = 2.706$). In the former analysis, however, there was a linear pattern with TR12 showing larger regression amplitudes than TRB12 which, in turn, showed larger amplitudes than TRAA12 (see Table 8.8). In the latter analysis, this trend was again shown (see Table 8.9).

TR12	TRAA12	TRB12
7.5 (4.0)	5.4 (1.5)	6.5 (3.5)

Table 8.8: Mean regression amplitudes (as measured by character spaces travelled) and standard deviations (in parentheses) for the reading task from five subjects on three fonts.

TR12	TRAA12	TRB12	TR14
8.1 (4.4)	5.6 (1.7)	7.1 (3.7)	5.5 (2.1)

Table 8.9: Mean regression amplitudes (as measured by character spaces travelled) and standard deviations (in parentheses) for the reading task from four subjects on four fonts.

As subjects had not been restricted to a systematic (reading strategy) search pattern on the search trials, it was likely that eye movements would partly consist of vertical and diagonal saccades as well as horizontal ones. As this would result in the computerised analysis being quite meaningless, it was not deemed appropriate to analyse saccade amplitudes or regression size from the search recordings.

ANOVAs conducted on fixation durations (ms) recorded during the reading task were not significant for either analyses ($F < 1.0$). Nor was there a significant difference in fixation duration when the reading task and corresponding search task recordings were combined ($F(3,21) = 1.860$).

Analysis of fixations during the search task ($N = 4$) revealed no significant differences ($F(3,9) = 1.351$).

Correlating (Pearson's r) the combined ID ratings with reading saccade amplitudes produced a figure of -0.992 ($df = 3, p < 0.001$). Correlations for the ID and reading regressions and for ID and reading fixations gave Pearson r values of -0.905 ($p < 0.05$) and 0.443 (ns), respectively. The Pearson r for ID and search fixations was 0.546 (ns).

Regression coefficients for the eye movement parameters and the overall ID were as shown in Table 8.10.

Parameter	Regression coefficient
Reading saccades	$14.520 - 0.018 \times ID$
Reading fixations	$252.494 + 0.026 \times ID$
Reading regressions	$9.30 - 0.012 \times ID$
Search fixations	$285.555 + 0.068 \times ID$

Table 8.10: Regression coefficients for the overall index of density and eye movement parameters.

8.10.3 Probability vector and transition matrices analyses

As in Experiments 3 and 8, eye movement data were subjected to probability vector and transition matrices analyses. Table 8.11 shows depicts the proportion of diagonal saccades in the reading and searching tasks. Table 8.12 shows results from transition matrices for the four fonts on reading and searching

	Reading				Searching			
	TR12	TRAA12	TRB12	TR14	TR12	TRAA12	TRB12	TR14
VP	0.008	0.015	0.028	0.014	0.083	0.068	0.081	0.081
DS	0.004	0.004	0.004	0.012	0.042	0.025	0.035	0.023
AC	0.015	0.014	0.013	0.011	0.302	0.104	0.199	0.073
AH	0.002	0.014	0.008	*	*	*	*	*
$\bar{X}$	0.007	0.012	0.013	0.012	0.142	0.066	0.105	0.059

Table 8.11: Proportion of diagonal saccades in the reading and searching tasks (* = data lost).

		Reading								Searching							
		TR12		TRAA12		TRB12		TR14		TR12		TRAA12		TRB12		TR14	
		E	W	E	W	E	W	E	W	E	W	E	W	E	W	E	W
VP	E	86	14	84	16	83	17	85	15	81	19	80	21	82	18	85	15
	W	60	40	60	40	57	43	60	40	21	79	23	77	17	83	17	83
DS	E	78	22	81	19	79	21	79	21	75	25	82	18	79	21	82	18
	W	59	41	52	48	62	38	53	47	34	66	22	78	26	74	19	81
AC	E	80	20	80	20	82	18	77	23	*	*	*	*	*	*	81	19
	W	72	28	74	26	77	23	75	25	*	*	*	*	*	*	22	78
AH	E	81	19	84	16	80	20	*	*	*	*	*	*	*	*	*	*
	W	70	30	61	39	64	36	*	*	*	*	*	*	*	*	*	*

Table 8.12: Transition matrices for the four fonts on reading and searching tasks (* = missing data or insufficient data for analysis).

8.11 DISCUSSION

8.11.1 Eye movement parameters and ID ratings

It was hypothesised, at a very general level, that any differences in behavioural measures relating to fonts would be accountable by the ID measure.

Taking the 'combined rating' as the most comprehensive form of the ID, it can be seen that there is a marked difference between the standard binary 12-point font and the other fonts examined. For example, the binary font has only 61 percent the density or boldness of its supposed antialiased counterpart. In some cases (e.g. upper case H) the bold 'equivalent' had over twice the density of the standard font. Whilst antialiased and bold fonts were comparable (329.5 vs 336.5), there was also little difference between these two and the 14-point standard font (361.5). From this perspective alone, it must be concluded that the TR12 is not comparable with the TRAA12.

There were also slight differences in screen luminance. Readings from one text passage displayed in the four fonts showed (average of several readings) the TRAA12 screen to be of identical luminance to that of the larger font (TR14) screen (59.6 cd/m^2), whilst that for the TRB12 was less (58.0 cd/m^2) and that for the TR12 was a little greater (61.3 cd/m^2). Although not a lot of value can be attached to these marginal differences (presumably a consequence of different IDs), these data do provide some further evidence for the notion that there was more similarity between antialiased, bold and even a 14-point font than between the antialiased font and the binary.

On the *reading task,* there was a significant difference between *saccade amplitudes,* with the standard binary font being significantly greater ($\bar{X} = 11.001, sd = 4.482$) than both the antialiased font and the bold font ($\bar{X} = 8.437; sd = 2.492$ and $\bar{X} = 8.495; sd = 2.576$, respectively). This pattern was again found when analysing across the four fonts (excluding one subject).

Consistent with the view outlined above, saccade amplitudes were very similar between the TRAA12 and TRB12 fonts ($\bar{X} = 8.266; sd = 2.825; \bar{X} = 8.532; sd = 2.973$ respectively)—and even with the TR14 font ($\bar{X} = 8.302$; $sd = 3.062$) but not between the TR12 font ($\bar{X} = 11.098; sd = 5.170$) and the TRAA12.

Although the analysis of the *regressions* did not show any differences beyond what may be a chance expectation, there was again a consistent pattern regardless of how the data were analysed. For regression amplitude, there was considerably more similarity between the TRAA12 font ($\bar{X} = 5.595; sd = 1.695$) and the TR14 font ($\bar{X} = 5.486; sd = 2.068$) than that between the TR12 font ($\bar{X} = 8.057; sd = 4.381$) and its AA version.

There was a highly significant correlation ($r = -0.992$ $p < 0.001$) between the ID ratings and reading saccade amplitudes. There was also a significant correlation between ID and regression amplitudes within the reading task. Although this superficially supports the basic experimental hypothesis, due to the small sample size involved ($N = 4$), little meaning can be attached to this.

Regression coefficients were derived from the eye movement parameters as possible predictors of visual search and/or reading efficiency with various fonts. The X value taken was the overall density index. The Y values were taken as the mean from the average values from each of the five subjects. If the ID of any font can be specified, then these regression coefficients can predict spatial and temporal values of eye movements. For instance, if a prototype font possessed an ID of 300, then average saccade amplitude during reading would be expected to be 9.12 character spaces (as per the calibration within this experiment). The precise figure in terms of *mm* would depend upon the retinal angle subtended by the characters, but could readily be calculated. If the bold or AA equivalent of this font had an ID of 450, then the average saccade would be 6.42 spaces. In other words, a less efficient reading characteristic (shorter saccade amplitude) is shown by the non-standard font. This is the pattern shown within this experiment, and this suggests that AA fonts do not enhance reading. It also suggests, as has been noted over a long period, that use of bold font should be limited to the occasional few words for emphasis, and that it is not beneficial (e.g. can result in feelings of fatigue) if used extensively. Furthermore, as a general truism on reading, bold type typically produces no improvement in reading speed (e.g. Paterson & Tinker, 1940*a*).

Considering the same hypothetical system, regression equations would predict average reading fixations of 260 ms on the standard font, whilst reading fixations would increase only slightly to 264 ms on the AA or bold fonts. Again, this pattern of negligible difference in fixation durations between fonts was seen in this study and it suggests little benefit in displaying bold lettering other than for highlighting for specific attention.

Regressive eye movements in reading would show mean values of 5.7 character spaces in this hypothetical prototype binary font and 3.9 spaces in the AA or bold equivalent. Thus, although regressive movements are shorter with the bold font, and might suggest better reading, results from this experiment show the difference to be within chance expectation.

With the visual search task, the standard binary font would produce fixation durations of an average 306 ms, whilst those of an equivalent bold or AA font would produce mean fixations of 316 ms. Although again not of statistical significance, the direction would suggest that, if anything, the non-standard font produces a less efficient search.

Although a few reservations might remain, it seems most probable that the hypothesised relationship between readability and density is a genuine one. These reservations are considered next.

The results obtained are not likely an artifact of the selection of particular characters for analysis. The character set is almost identical to that used in relevant previous studies suggesting a superiority of antialiased fonts over binary fonts (e.g. Hewitt *et al.*, 1989) and as such is regarded as an appropriate choice.

Nor is it felt that doubt lies in the system of ascribing rating values to the pixels. Although at times difficulty was experienced in determining the appropriate values, particularly within the middle range, the few errors possible made due to perceptual discrimination problems are likely to have been balanced out by a fairly equal number of possible errors of under- or over-estimating brightness. A photometer check made on the subjective judgements, whilst not totally dispelling all doubt, showed that supposedly objective determinations are as prone, if not more so, to fluctuations in luminance re-readings of any particular pixel.

Despite the technical problem of a search coil breaking *in situ,* and consequent loss of potential data, the two alternative forms of data analyses produced essentially the same results.

8.11.2 Probability vector analysis

Firstly, in the *reading task,* the proportion of diagonal saccades can be seen to be very low (the overall mean is 0.011). In other words, looking at the overall mean, one out of one hundred saccades is diagonal. This holds true within every subject and within every font.

Secondly, in the *searching task,* this proportion is also low, with the overall mean being 0.093. This figure is far removed from the value (0.5) expected within a strictly random search strategy. It is of note that the pattern of means show that the proportion of diagonal saccades is highest in binary, less in bold, still less in AA, and is the lowest in the 14-point binary font. The numerical differences arise mainly from only one subject, yet the pattern which the means show is also that of the pattern shown by the other subjects. In particular, with all three subjects, the two non-antialiased 12-point font values are clearly higher than that seen for the antialiased font.

Again, these data alone do not provide an explanation of this result, but they may suggest a relationship between density of the fonts and the proportion of diagonal saccades. Specifically, the greater is the density (boldness) of the font, the more systematic is the stratégy adopted by observers in the searching task.

Thirdly, in each of twelve possible comparisons between reading and searching this proportion is larger in the searching condition than in the reading one. This adds support to the notion of the proportion of diagonal saccades as an 'asystematicity index'.

8.11.3 Transition matrices analysis

Firstly, in the reading task, the 2×2 matrices shown above contain nearly all the information conveyed by the full 8×8 transition matrix. Each 2×2 transition matrix contains, at least, the 94 percent of the pairs of saccades used to obtain the full 8×8 matrix. The matrices given for a particular observer are very similar in the four fonts. In fact, the matrices differ more between observers than between fonts. Considering the means, within the reading task, after a saccade type E, the probability of a saccade of type E is 0.81. After a saccade type W, this probability is 0.63. Therefore, there is a dependency between previous and subsequent saccade directions. However, irrespective of previous saccade directions, the most probable direction is E.

Secondly, in the searching task, two observers give 2×2 transition matrices containing at least a 83 percent of all pairs of saccades. In one font this requirement is also fulfilled by observer AC, but in the other three fonts, this observer's 2×2 matrices contain just 4 percent, 1 percent and 59 percent, respectively, of the total pairs of saccades. After a saccade of type E, the proportion of a saccade of type E

is very similar in all the nine matrices. The mean of this nine proportions is 0.81. More differences are apparent when comparisons are made of each matrix's second rows, although these differences do not seem to be related in any clear way to the fonts involved. Concerning the means, after a saccade of type *W*, the probability of a saccade of type E is 0.22. Therefore, the data seem to imply that, in the searching task, a 'reading' strategy is employed; but in both directions (not only left to right, but also right to left). In both tasks, $p(E|E)$ is 0.81. So, this seems to suggest that in a 'reading' strategy, it is four times more likely that a saccade will be followed by a saccade of the same direction than by one of the opposite direction.

Thirdly, a close examination of the three matrices from subject AC which are omitted shows that, under the binary and bold fonts, this subject used a systematic strategy—but an uncommon one. This individual swept the screen, column by column (rather than row by row, as usually is the case), from left to right. The first column was searched from top to bottom, the second from bottom to top, again top-to-bottom, etc. In the AA font, the subject applied the typical strategy to some part of the screen, yet his idiosyncratic strategy was applied to other parts.

8.12 CONCLUSIONS

8.12.1 Eye movement analyses

It appears to be the case that (a) the antialiased fonts examined are bolder than there binary counterparts, and (b) the bolder is the font the shorter is the saccade amplitude. On a reading task this would suggest (all other factors—e.g. comprehension—being equal) a less efficient reading eye movement pattern being involved. This seems at odds with what is generally claimed and with what has sometimes been found (albeit with short-comings). Subjective comments may hold the explanation. To refer again to the Hewitt *et al.* (1989) study detailed in §8.2.2, several experimental participants (notably younger ones presumably with better visual acuity) have described AA format as appearing fuzzy and out of focus. It is boldly suggested here that there is really *no* perceptual *superiority* with AA displayed text. Rather, the reverse may be the case. The differences emerging within this present experiment indicate that with users possessing untarnished visual acuity, AA presentation may be a hindrance and not a help. It was unfortunate that a correlation study, due to only the four fonts being employed, could not have helped elucidate the issue.

8.12.2 Probability vector and transition matrices analyses

Transition matrix analysis did not reveal any differences between fonts when the task is a text reading one. As expected, when the observer is engaged in a task as stereotyped as reading, there is little opportunity for adopting strategies. The analysis shows only a few redundant saccades; i.e. the proportion of diagonal saccades is almost zero. It may seem incongruous that although differences emerge between fonts in terms of saccade amplitudes, there is no difference within transition matrices analyses. However, although longer saccades result in less saccades to read each line, longer saccades do not imply (a) any more or any less diagonal saccades, and (b) a different relation between the two types of horizontal saccades is taken into account within Table 12.

Some small differences between fonts have been suggested by the searching task. If the four fonts were ordered firstly according to their proportion of diagonal saccades, and secondly according to saccade amplitude in the reading task, there were only two identical rankings. Therefore, the index of density is also linearly related to the proportion of diagonal saccades in the search task. Very tentatively, this result seems to suggest that the number of diagonal saccades varies increasingly with decreasing font density. From the transition matrices, it can be seen that the most common strategy is similar to that used in the reading task, but with one important difference. In the search task, the reading proceeds from left to right, then from left to right, etc. No specific search patterns seem related to the different fonts.

8.12.3 Summary

It must be stressed that the lack of significant results may, in part, be related to the low sample sizes (N=5) employed. However, the concerns outlined above suggest shortcomings more related to the actual fonts examined rather than number of subjects. The development of AA fonts is an ongoing project area within the company concerned and although those AA fonts employed in studies by workers such as Bender *et al.* (1987), Gould *et al.* (1987*b*), Hewitt *et al.* (1989) and Miller *et al.* (1989) varied slightly from those employed within this present study (and probably between themselves), there is little doubt that the issue of density was confounding their results. It is essential that in any further evaluative work, the density factor is fully taken into account.

Chapter Nine

Conclusions:
Implications for Visual Search and VDUs

9.1 GENERAL OVERVIEW

The VDU is the main source of feedback to the user from the computer. Whatever the application currently in use, the need to search the display has always been present, but with the advent of bit-mapped and menu-driven formats, visual search has more than ever been an integral part of operating a computer. The user is more likely to need to search for a file icon, and click this to open rather than to type a command at the cursor line. There is much less emphasis on syntactic structure, and much more emphasis on appearance. Indeed, within a direct manipulation system, user and computer communicate mainly in terms of appearance (Shneiderman, 1983). There are important perceptual and cognitive principles associated with such systems; e.g. 'recognition is easier than recall' or 'seeing and pointing is better than remembering and typing'. An important consideration is how to display information so that the user can quickly locate the items needed.

The aim of this thesis has been to explore aspects of visual search in relation to VDUs. However, the study of any aspect of HCI will involve numerous variables at both the machine and human sides of the interaction. As we anticipated in Chapter 1, in order to do justice to the research area, it has been inappropriate to pursue a unitary path. Many parameters must be considered in a valid examination of visual search and VDUs. These parameters include the overall presentation (e.g. screen shape; Chapter 3), how information may be optimally presented either in central/foveal vision (e.g. cursor presentation; Chapter 4) or in peripheral vision (e.g. status bar information; Chapter 5), when and why icons might be chosen over verbal labels (Chapters 6 and 7), and how new font designs may enhance the locatability of items and perhaps also screen-based reading generally (e.g. antialiasing; Chapter 8). The experimental approach adopted has involved using both relatively pure search tasks (e.g. Expts 1 & 10) and relatively realistic search tasks (e.g. Expts 4 & 6) which have been supplemented with fine analyses of the eye movements involved in order to shed light on the underlying processes which would otherwise remain unknown or at best only inferred.

The first chapter was devoted to describing some relevant ergonomic principles of VDUs, the role of eye movements or absence of them (e.g. peripheral vision), the influence of attention, and the actual monitoring of eye movements. A review of some studies of eye movements and VDUs concluded this chapter.

Chapter 2 provided an in-depth review of visual search. Early work by Gottsdanker (1960), Neisser (e.g. 1963) and Bloomfield (1970) were outlined. After describing feature-integration theory, the importance of attentional focus, etc., some specific applications of visual search were described.

Chapter 3 examined the role of array shape in search tasks. It was concluded that alphanumeric information (although not text passages) presented on a VDU is more easily searched and located when arrayed in horizontally elongated configurations. This has particular implications for multiple-window displays. A model was proposed which stemmed from suggestions by Bloomfield and accommodated more recent data on visual lobe boundaries. Due to the visual lobe geometry, it was suggested that more 'wastage' would occur with material consisting of shorter horizontal lines; i.e. with the vertically elongated arrays.

The studies described within Chapter 3 demonstrated a search superiority with horizontally elongated arrays over vertically elongated ones (at least as far as non-reading materials were concerned). Two possible reasons were suggested for this by eye movement patterns. Firstly, horizontal arrays may be better matched to the visual lobe. Secondly, size of scanning saccades may be affected by the array shape, and hence lead to more efficient scanning behaviour. This allowed a revaluation of Bloomfield's (1970) conclusions. It was suggested that verbal or alphanumeric materials are susceptible to the shape effect, whilst passages of text or reading material appear to relatively immune. This is possibly due to the stereotyped nature of of the visual and cognitive processes involved in reading; i.e. the absence of susceptibility to the shape effect is likely due to higher level factors related to linguistic processing and cognitive load (e.g. Lévy-Schoen *et al.*, 1983; Jacobs, 1986). As a vertically oriented letter set seemed to be equated with reading material in the particular Bloomfield study, it is not surprising that he found his results inconclusive.

The evidence would suggest that a search performance superiority would also be found within horizontally biased configurations for iconic materials, as these are a form of non-reading material. There would seem to be implications for the spatial layout of iconic information within VDUs, particularly within multiple-windowed systems. Similarly, there are implications for menu layouts generally, and particularly iconic menus and those with brief verbal descriptors.

The expected edge effect—that phenomenon whereby the outer area of an array receives fewer fixations than the inner half—did not emerge. A likely explanation for this (stemming from subjective reports) is that many subjects (as a tedium relieving strategy) often commenced the search with a quick scan of the periphery of the arrays. This finding calls into question the previously claimed 'robustness' (Monk, 1981) of the effect, at least as far as free search strategies are concerned.

It is possible that, from an ecological point of view (e.g. Gibson, 1967), the pattern found within this experiment is linked to mankind's horizon-based perceptual system. An alternative framework concerns the fact that acuity falls off peripherally from the fovea almost equally in both the horizontal and vertical planes (Rovamo & Virsu, 1979). The likely explanation here concerns an interaction between eye movement tendencies and text line arrangement.

Given the variability in saccades sizes, although the idea of visual lobes being arranged neatly to fit into the array framework is probably over-simplistic, the different configurations may well induce scanning strategies with different degrees of visual lobe overlap.

A tentative explanation for the performance differences found within the different shapes was offered based on eye movement analyses. It was shown that different shapes were related to different saccade amplitudes, with longer saccades being more associated with horizontally elongated arrays than with vertically biased arrays. It was also demonstrated that the proportion of diagonal saccades was related to the array shape, with the proportion of these redundant saccades being greater in vertical configurations. Data from a transition matrix analysis of the reading task supported this conclusion as there was greater similarity between the free and systematic 2×2 transition matrices when both were horizontal than when both are vertical. Both longer saccade amplitudes and a less systematic search therefore can be the explanation for the differences in performance due to array shape. Two explanations were proposed. Firstly, it was suggested that differences in saccade amplitude account for the differences in search times. Secondly, there is the possibility of a more redundant search within vertically biased configurations due to greater asystematicity. The more redundant is the search, the more likely users are to repeat the scan over the same locations, and the greater is the length of time necessary to find the target.

It was concluded from both (a) an analysis of the frequency of saccades redundant for the task involved (diagonal saccades) and from (b) the analysis of dependency between saccades necessary for the task (E and W), that it seems vertically oriented arrays are characterised by a more asystematic search.

Chapter 4 described a text-editing study showing a performance superiority where information was presented at the cursor rather than on a status line. Issues discussed included task analysis, attentional switches, and internal vs external memory. The interpretation for the results of Experiment 4 was along the lines of Russo (1978) who pointed out that the involvement of eye movements in search was likely to involve a cognitive cost in terms of both the execution time and the preparation time for the activity. It was suggested that designers should more carefully consider these aspects. Whilst there are limits to the amount of information which can be provided with a cursor, several possibilities such as colour, shape and blink rate may be explored.

If information could be displayed peripherally without the need for such attentional breaks, then this should manifest itself in performance enhancements. The combined record from the HIMS/eye monitoring technique supported the notion that there was a distinct pause before a mode change when eye movements were required to verify the mode status through an attentional micro-switch. As was said in §4.11, the major advantages of computer capture of behaviour, rather than video or verbal protocol techniques, is that it is automatic, error free, large amounts of data can be collected efficiently and the data is often readily available for further statistical analysis. In the case of the present experiments, a further behaviour of interest (i.e. eye movements) was easily synchronised with the HIMS recordings.

A second study employing the HIMS and a customised text-editor was then carried out to investigate peripheral salience (conspicuity) and status information. It was concluded that the superior performance observed within a high conspicuity condition was attributable to the relative lack of disruptive saccades. Once again, the principle study was supplemented by an eye movement analysis of the fine visual tasks involved. It was then possible to determine how frequently subjects looked at the specific areas where mode information was displayed. Over the (four) low to high salience conditions, incidences of fixating upon the 'informative' screen areas diminished in a dramatic pattern.

This study also demonstrated a further application of eye movement monitoring in relation to visual search studies in that the scleral coil technique served to substantiate a hypothesised task analysis of HCI. It was suggested within §5.9 that symbol blinking could also be a very effective indicator of system state, especially with display areas positioned in the periphery of the normal visual field where the eye is most sensitive to movement and changes of intensity. As an overall conclusion, it was apparent that the peripheral presentation of status information and other on-screen mode changes is presently grossly under-used, perhaps particularly so for novices.

The role of peripheral visual processing is seen as of particular importance in the ease of locatability of icons. The results of the visual search task described in Experiment 8 showed icons to be more quickly located within an array than textual material. The primary reason for the 'icon locatability supremacy' assumed differences in peripheral visual processing. When the target was relatively simple or when it was an icon, the item was more easily resolved in peripheral vision and thus search performance was enhanced. The principle reason offered for this involved spatial frequencies. As textual material is generally characterised by higher spatial frequencies, it is likely to be more difficult to resolve in peripheral vision (e.g. Wilson *et al.*, 1990). It was suggested that on many occasions within Experiment 8, the iconic information available in peripheral vision was sufficient for recognition, and hence directly accessed without a true search.

To investigate the bases for these differences fixation durations and scan paths were examined. Four methods of analysing the search paths (including transition matrices) suggested greater systematicity of search strategies for complex items. It was concluded within the next experiment (Expt. 9) that search times were greater for words than for icons, and greater when the item was relatively complex than when it was relatively simple. It was suggested that the more useful is peripheral vision, i.e. the more easily resolvable the target is in the periphery, the less essential it is to use the types of saccades assumed necessary for the task. For example, if the target is resolvable in peripheral vision, it is an appropriate strategy to rely on peripheral information and to search only those positions where peripheral vision detected greater similarity with the target. In this sort of case, a systematic search is obviously not produced by the observer. On the other hand, if peripheral vision does not happen to be particularly informative, the observer can readily rely on a systematic search and so avoid redundant searching. This assumption therefore predicts a low level in the proportion of diagonal saccades in conditions where search time is also low. Results from Experiment 9 neatly fitted this prediction.

As indicated by the transition matrix, the structure of the systematic component of the search shows greater differences between words and icons than between simple and complex items. The main difference between words and icons may be explained by supposing longer saccades being produced in the icon condition. The significant effect found on factor *W–I*, but not on *C–S*, suggests that searching for icons may be regarded as involving a more efficient pattern of eye movements.

Data from five of the six participants in Experiment 9 showed greater efficiency in search for icons than for words. In addition, simple items showed more oblique changes than complex items. It was suggested that peripheral vision is used more in locating simple items. Where peripheral vision cannot be used, users are compelled to exercise a more systematic search. In these situations we therefore find lower values within a transition matrix.

Analysis by scan path of the number of items fixated before the target was located suggested inferior strategies being employed on complex words with progressively more efficient strategies with simple words and complex icons to a superior strategy shown for simple icons. Furthermore, it was noted that there was approximately equal numbers of clockwise and anti-clockwise searches as well as those where the scan path was erratic or too brief to interpret.

The fact that average search times were shorter for icons may seem at odds with the suggestion that icons may be characterised by a *less* systematic search. The analogy proposed was that of the difference between searching on a page of a book for a reference date (e.g. '1959') and a Japanese name in its untranslated form. Although people are likely to be less systematic in searching for the Japanese character (they are likely to randomly move their eyes over the page waiting for the name to *pop-out* (see §2.5), search time is likely to be shorter than for the year date.

By way of conclusions, it was proposed that if a target is detectable in peripheral vision, then (a) there are fewer different positions fixated before the target is detected; (b) there are more cases where the first saccade successfully locates the target; and (c) there is a less systematic search (as revealed by a transition matrix). It was therefore a logical progression to carry out a further visual search study to investigate the role of spatial frequencies in icon versus word locatability.

As it is important to know what factors underlie the icon search superiority, a relatively pure search study was then conducted with particular regard for the possible role of peripheral processing. Experiment 10 described a visual search study employing spatial frequency grids to examine the role of variables such as cycle frequency, high/low contrast, and high/low similarity of non-targets.

All variables examined showed strong effects on search times. The frequency of cycles variable was particularly strong, with an effect thirteen times that of the contrast variable. As a comment intended as impetus for further research, it was remarked that although these results did support the notion of the lower spatial frequency characteristics of icons underlying their search superiority, other mechanisms might also be involved in this apparent ease of peripheral processing. Other factors such as articulatory distance, global/local features, as described in §6.2.1, may also contribute.

An eye movement analysis of fixation durations and scan paths was then conducted to determine if eye movement parameters could account for the observed differences. The differences observed within the previous study might simply be due to peripheral processing *per se* leading almost immediately to detection. Indeed, only two of the twelve different eye movement analyses demonstrated a significant result and the general hypothesis was therefore supported. It was concluded that the search times found for the three spatial frequency variables associated with the icon search superiority have no correlates in eye movement parameters. It may be that there is an attribute of icons which results in them being located 'just a saccade away', and that attribute may well be the ease in peripheral perception of their lower frequency composition.

The relatively applied studies detailed in Chapter 8 had potential applications for ongoing commercial development. Experiment 12 aimed to investigate whether, and to what extent, antialiasing contributes to the legibility of alphabetic characters. A perceptual (visual search) task involving straight-forward legibility was chosen in order to avoid cognitive factors such as comprehension from arising. It was hypothesised that search times for single target letters presented amidst randomly generated alphabetic arrays would be significantly faster for AA VDU presentations. This hypothesis was not supported. Suggestions for new directions were proposed. In the light of findings from a tentative numerical assessment of density, it seemed appropriate to compare a binary vs bold vs AA font in an attempt to ascertain the contribution of antialiasing *per se* to any performance enhancement for VDU users.

Experiment 13 made use of eye movement parameters and both a classic visual search task and a reading task. One of the main aims of this final study was to develop a more precise quantitative ID and to use this as a means of comparison between fonts. It was hypothesised that there would be a clear relationship between an ID for the four fonts investigated and the eye movement measures of fixation duration and (where appropriate) saccade and regression amplitudes. It appeared to be the case that (a) these AA fonts are *bolder* than their binary counterparts, and (b) the bolder is the font the shorter is the saccade amplitude. On a reading task this would suggest a less efficient reading eye movement pattern being involved. This seems at odds with what is generally claimed by manufacturers for AA fonts. It was suggested that there really is no perceptual superiority with AA displayed text, and that the reverse may in fact be true. It may be so that for users possessing relatively good visual acuity, AA presentation may be a hindrance rather than a help.

There seems little doubt that the effects of density has been confounding the effects in previous AA studies. It was been recommended that before definitive statements can be made concerning the benefits of AA VDU fonts, an AA font is produced which is *truly equivalent* to its binary counterpart. Perceptual equality between screen-displayed information hard copy remains an important goal to pursue. Chapter 8 highlighted a major shortfall which must be rectified before further research is viable. Unfortunately, indications are that future results from scientific studies may fall contrary to those desired by business companies with vested interests.

A specific form of analysis employed as part of three eye movement studies contained within this thesis is that of analysis by transition matrices. From eye movement recordings it was possible to calculate the proportion of each saccade type (e.g. E), i.e. the probability vector (a first order analysis); and the proportion of a particular saccade type being followed by a saccade type N, saccade type NE, ... type NW, i.e. the transition matrix (a second order analysis). From these two analyses, it is then possible to show the characteristics of the search strategy. For example, if the tasks convey information in rows and/or columns, some information can be expressed as the proportion of diagonal saccades. In a rectangular array, a less systematic search would be characterised by a greater proportion of diagonal saccades.

The transition matrix provided useful information. For example, it was possible to detect whether the search was systematic or asystematic. If it was a systematic search, it was possible to discern whether the strategy is a reading one, a 'snaking' one, and so forth.

In each of the transition matrix analyses detailed within this thesis, the visual material was arranged in such a way as to require only certain types of saccades. In particular, no diagonal saccades were actually required for the tasks involved. Therefore, it was possible to determine if the proportion of diagonal saccades could convey information regarding the possible non-systematicities associated with the experimental conditions. For instance, in Experiment 13, transition matrix analysis did not reveal any differences between fonts when the task is a text reading one. As expected, when the observer is engaged in a task as stereotyped as reading, there is little opportunity for adopting strategies. The analysis shows only a few redundant saccades; i.e. the proportion of diagonal saccades is almost zero. It may seem incongruous that although differences emerge between fonts in terms of saccade amplitudes, there is no difference within transition matrices analyses.

Although earlier notions had suggested that diagonal saccades were equivalent to low search *efficiency*, as in the shapes study (Expt. 3), it became apparent that this is so only under certain conditions. If peripheral vision is important as in the icon study (Expt. 8), then exactly the opposite relation holds.

9.2 A PARADIGM?

This thesis draws to a close with a brief excursion into the philosophy of science. In his treatise on the history of scientific ideas, Thomas Kuhn (1970) defined the term 'paradigms' as 'those accepted instances of actual scientific practice which include law, theory, application, and instrumentation together, that provide models from which spring particular coherent traditions of scientific research' (p. 10). The notion of a paradigm, according to Kuhn, is closely related to that of 'normal research' which refers to the canons and beliefs adopted either implicitly or explicitly by the practitioners of a specific discipline in defining the rules of the game for that discipline (pp. 23–42). These canons and beliefs, by virtue of their status, once established are no longer questioned and become preconceptions in the strict sense of the word.

Viviani (1990) comments that the rather specialised field of eye movement research may afford a particularly good example of a scientific paradigm. Perhaps all the ingredients are there.

> The instrumentation and the experimental procedures have become fairly standardized... Also, the information-processing parlance and its recent cognitive flavoring supply the intellectual scaffolding to support the data. Last but not least, a central dogma is at hand which provides the initial motivation for much of the work in the domain. The dogma takes several forms and disguises, most of which can, however, be summarized by the following quote from a leading scientist in the field: "[Exploratory saccadic] eye movements can at the very least be considered as tags or experimentally accessible quantities that scientists can observe to understand underlying processes of cognition". Once such a view has been adopted, the motivation for studying visual search by humans follows syllogistically from the (uncontroversial) belief that the study of human cognition is a worthy endeavour, both for its own sake, and for the possible practical outgrowths. (p. 354)

Viviani's critique is entitled 'Eye movements in visual search: cognitive, perceptual and motor control aspects', and as this sort of tripartide information processing approach was introduced in §1.8.1 of this thesis on visual search and underlying eye movements in relation to displays, it is worth taking a closer look at this article. It is particularly important to do so as this lengthy chapter within an internationally distributed book may well become much cited.

The main problem which Viviani considers is: What can be learned from eye movement data about the cognitive and perceptual processes involved in the exploration of the visual world? In the first section of his chapter Viviani critically considers the assumptions that might be made in order to lend scientific credibility to the stated problem. He argues that there are several serious (and perhaps fatal) difficulties with many beliefs within this field, and particularly in the assumption that mental processes can be inferred inductively on the sole basis of experimental findings concerning eye movements. He also argues, however, that under certain circumstances, eye movement data may be useful in evaluating specific theories of visual and cognitive processes.

The point of view which emerges from Viviani's critique is pessimistic. He clearly believes that the hope of progressing inductively from experimental data on exploratory eye movements to a theoretical description of the underlying cognitive processes is ill-founded. He states that eye movements will only become useful in the context of a deductive approach; it will only be after a reasonably articulated theory of these processes has been derived that data on overt behavior will genuinely contribute to understanding by verifying or disproving the predictions of the theory.

> Chess-playing affords an interesting example of this essential difficulty in relating eye movements to complex mental procedures. It is known that expert players perceive the very many alternatives that are normally open in the middle game almost as though they were actually displayed on the board. In-depth analysis is then carried out only for very few among these alternatives. Clearly, in such a complex mental task we find aspects that suggest parallel processing and others that suggest serial processing, the two modes being intimately intertwined. Even if we assume that each eye movement between two cases on the board corresponds to a hypothetical move, it would still be impossible to know which mode is actually active. Moreover, several different developments may include the same move. Not surprisingly, only the analysis of on-line verbal reports has, in a few cases, justified a posteriori the corresponding pattern of eye movements (De Groot, 1978). (Viviani, 1990, p. 355)

Some other evidence noted by Viviani such that fixations do not always coincide with the points on the scene that are being attended to comes from work on the allocation of attentional resources (e.g. Posner, 1980). It may well be the case that, under appropriate conditions of pre-cueing, visual attention can be directed almost everywhere in the visual field, regardless of the direction of the actual line of sight.

It is true, as Viviani comments, that search paths that result from different verbal instructions to the observer (perhaps as found in the reading-type vs systematic searches of the present Expt. 2) are radically different, and yet the same set of mental processes (inferences, memory search and retrieval, comparison, etc.) is likely to be involved in every case. It is also true that there have been unfortunate quotes within the literature such as: 'eye movements reflect the human thought processes' (Yarbus, 1967, p. 190).

However, it would seem that much, if not all, of Viviani's criticisms refer to the Posnerian spatial attentional type task or to studies relating to Scanpath Theory (Noton, 1970; Noton & Stark, 1971*a,b*). What Viviani suggests does not undermine the fact that both spatial and temporal eye movement parameters are useful in evaluating the efficiency of search within arrays of varying configurations. His views do not counter the fact that one can monitor the instances of saccades to status line position as opposed to maintaining the eyes at a working position when evaluating the optimum presentation of status information. His views do not weaken the fact that clear differences in search strategies can be demonstrated for high and low conspicuity information presented on screen. Nor do these views weaken findings from the four types of search path methodologies outlined within this thesis for the analysis of the underlying behaviour responsible for the icon locatability supremacy. Viviani's views do not detract from the finding that the eye seems readily drawn to items of similar spatial frequency to the target, thereby substantiating the notion of important peripheral processing occurring during visual search. His views do not invalidate the fact that eye movement recordings have been invaluable in making comparisons of reading materials since Tinker's pioneering studies up to present-day studies of AA and other VDU-displayed fonts. Despite perhaps being regarded as a article of worth, it is suggested that many applied psychologists who employ eye movement recordings are yet to encounter a critique which truly threatens their work.

The differences of opinion between Viviani and the present author serve to underline the fundamental differences in approach between them. It reflects the fact that this thesis has been primarily concerned with peripheral processes rather than with the central, cognitive processes which control them. To end this debate, a most fitting quote would be that of Graf *et al.* (1983) who concluded that: 'For ergonomical research, eye movements are as useful as search times for evaluating display formats. Furthermore, by the analysis of eye movements we get information about the scanning behaviour of the eye. Eye movements give a clue about the exact position of the problems in the displayed format' (p. 353).

9.3 SUGGESTIONS FROM THESE STUDIES

From Experiments 1, 2 and 3 it is suggested that designers make more use of horizontally elongated array shapes, particularly with multiple-windowed systems. It has also been suggested from Experiments 4 and 5 that designers should give greater acknowledgement to cursor-presented information. Colour and shape change and blink rate may well be harnessed to provide rapidly assimilated VDU information. Similarly, the use of peripherally presented information should also be put to greater advantage. The use of peripherally presented information would seem to be enhanced when information is displayed in iconic form; and, again, there are implications and potential recommendations for screen design. Finally, it is suggested that it is probably more useful for designers to concentrate their efforts on *increasing* the letter/background contrast, rather than employing devices such as antialiased fonts which effectively reduce this contrast and may serve no further purpose than to act as 'sales-gimmicks'.

9.4 RECOMMENDATIONS FOR FUTURE WORK

This thesis concludes with some recommendations for future research. Firstly, there is much scope for exploring the optimisation of status information. Cursor-presented information is one potential avenue and use of colour for enhancing conspicuity is another. The HIMS protocol analyser has proved a useful device in this field.

Secondly, it would be worth exploring whether icons are susceptible to a horizontal bias effect. It would be hypothesised that as is so of alphanumeric characters (though not of reading material), this effect would be observed. It would be a simple matter to devise an array study using blocks of random iconic characters along the lines of the array studies detailed within.

Thirdly, although support has been given for the spatial frequency/peripheral processing theory in accounting for the icon search supremecy, the possible role of articulatory distance differences, global versus local features etc., have not been ruled out. Further work is also in order here.

Fourthly, spatial frequency analysis seems likely to be a major research field over the next decade and there are substantial applications of this methodology in exploring the optimisation of displays.

It is worth reiterating once more that the studies contained within this thesis demonstrate the value of eye movement monitoring in relation to visual search and VDUs. As a further research tool which has proved valuable in this area, the benefits of transition matrices analysis must be underlined. Again, there appears to be considerable potential for their continued use.

References

Anderson, J.R. (1976) *Language, Memory, and Thought.* Erlbaum, Hillsdale, NJ.

Anderson, J.R. (1980) Acquisition of cognitive skill. *Psychological Review,* **89,** 369–406.

Anderson, J.R. (1982) Speech imagery is not always faster than visual imagery. *Memory and Cognition,* **10,** 371–380.

Anderson, T.W. & Goodman, L.A. (1963) Statistical inference about Markov chains. In: R.D. Luce, R.R. Bush & E. Galanter (eds.), *Handbook of Mathematical Psychology, Vol. 2.* Wiley, London, pp. 241–262.

Anstis, S.M. (1974) A chart demonstrating variations in acuity with retinal position. *Vision Research,* **14,** 589–592.

Antes, J.R. (1974) The time course of picture viewing. *Journal of Experimental Psychology,* **103,** 62–70.

Arend, V., Muthig, K.-P. & Wandmacher, J. (1987) Evidence for global feature superiority in menu selection by icons. *Behaviour and Information Technology,* **6,** 411–426.

Askwall, S. (1985) Computer supported reading vs reading text on paper: a comparison of two reading situations. *International Journal of Man–Machine Studies,* **22,** 425–439.

Astley, S.M., Taylor, C.J., Boggis, C.R.M. & Asbury, D.L. (1991) Cue generation and combination for mammographic screening. In: D. Brogan (ed.), *Visual Search II.* Taylor & Francis, London, (forthcoming).

Atkinson, R.C. & Crothers, E.J. (1964) A comparison of paired associate learning models having different acquisition and retention axioms. *Journal of Mathematical Psychology,* **1,** 285–315.

Baecker, R.M. & Buxton, W.A.S. (1987) *Readings in Human–Computer Interaction.* Morgan Kaufman, California.

Baker, C.A., Morris, D.F. & Steedman, W.C. (1960) Target recognition on complex displays. *Human Factors, 2,* 51–61.

Bannon, L., Cypher, A., Greenspan, S. & Monty, M. L. (1983) Evaluation and analysis of users' activity organisation. In: *Proceedings of the ACM Human Factors in Software Conference, CHI'83,* pp. 54–57.

Barbur, J.L., Forsyth. P.M. & Wooding, D.S. (1991) Eye movements and search performance—measurement and modelling work. In: D. Brogan (ed.), *Visual Search II.* Taylor & Francis, London, (forthcoming).

Barnard, P., Grudin, J. & Maclean, A. (1989) Developing a science base for the naming of computer commands. In: J. Long & A. Whitefield (eds.), *Cognitive Ergonomics and Human-Computer Interaction.* CUP, pp. 95–133.

Bartz, A.G. (1976) Peripheral detection and central task complexity. *Human Factors,* **18,** 63–70.

Bauer, D. & Cavonius, C. (1980) Improving the legibility of visual display units through contrast reversal. In: E. Grandjean & E. Vigliani (eds.), *Ergonomic Aspects of Visual Display Terminals.* Taylor & Francis, London, pp. 137–142.

Bauer, D., Bonacker, M. & Cavonius, C.R. (1981) Influence of VDU screen brightness on the visibility of reflected images. *Displays,* **2,** 242–244.

Baylor, D.A., Nunn, B.J. & Schnapf, J.L. (1987) Spectral sensitivity to cones of the monkey, *Macaca fascicularis. Journal of Physiology,* **390,** 145-160.

Beck, J. (1967) Perceptual grouping produced by line figures. *Perception and Psychophysics,* **2,** 491–495.

Bellamy, L.J. & Courtney, A.J. (1981) Development of a search task for the measurement of peripheral visual acuity. *Ergonomics,* **24,** 497–509.

Bender, W., Crespo, R.A., Kennedy, P.J. & Oakley, R. (1987) CRT typeface design and evaluation. In: *Rising to New Heights with Technology: Proceedings of the Human Factors Society 31st Annual Meeting.* HFS, Santa Monica, vol. 2, pp. 1311–1314.

Bergen, J.R. & Julesz, B. (1983) Parallel versus serial processing in rapid pattern discrimination. *Nature,* **303,** 696–698.

Bewley, W. L., Roberts, T. L., Schroit, D. & Verplank, W. L. (1983) Human factors testing in the design of Xerox's 8010 'Star' Office Workstation. In: *Proceedings of CHI'83,* ACM, pp. 72–77.

Bharucha-Reid, A.T. (1960) *Markov Processes and their Applications.* McGraw-Hill, New York.

Bhatnager (1987) Cited in Drury (1991).

Biederman, I. (1987) Recognition-by-components: A theory of human image understanding. *Psychological Reviews,* **94,** 115–147.

Biederman, I., Rabinowitz, J.C., Glass, A.L. & Stacy, E.W. (1974) On the information extracted from a glance at a scene. *Journal of Experimental Psychology,* **103,** 597–600.

Bigelow, C. & Day, D. (1983) Digital typography. *Scientific American,* **249,** 106–119.

Blakemore, C. & Campbell, F.W. (1969) On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology,* **203,** 237–260.

Bloomfield, J.R. (1970) *Visual Search.* Ph.D. Thesis, University of Nottingham.

Bloomfield, J.R. (1975) Studies on visual search. In: C.G. Drury & J.G. Fox (eds.), *Human Reliability in Quality Control.* Taylor & Francis, London, pp. 31–43.

Boer, L.C. & Keuss, P.J.G. (1982) Global precedence as a post-perceptual effect: An analysis of speed-accuracy tradeoff functions. *Perception and Psychophysics,* **31,** 358–366.

Boersema, T. & Zwaga, H.J.G. (1990) Searching for routing signs in public buildings: the distracting effects of advertisements. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 151–158.

Booth, K.S., Bryden, M.P., Cowan, W.B., Morgan, M.F. & Plante, B.L. (1987) On the parameters of human visual performance: An investigation of the benefits of antialiasing. *IEEE Computer Graphics and Applications,* **8,** 210–211.

Borah, J. (1987) *How to Evaluate an Eye Tracking System.* Applied Science Laboratories, Waltham, MA.

Bouma, H. (1978) Interaction effects in parafoveal letter recognition. *Nature,* **226,** 177-178.

Boynton, R.M. (1979) *Human Color Vision.* Holt, Rinehart & Winston, New York.

Bovzkov, V., Bohdanecký & Radil-Weiss, T. (1977) Target point selection during scanning eye movements. *Biological Cybernetics,* **27,** 215–220.

Brems, D.J. & Whitten, W.B. (1987) Learning and preference for icon-based interfaces. In: *Proceedings of the 31st Annual Meeting of the Human Factors Society.* HFS, Santa Monica.

Briand, K.A. & Klein, R.M. (1987) Is Posner's 'Beam' the same as Treisman's 'Glue'?: On the relation between visual orienting and feature integration theory. *Journal of Experimental Psychology: Human Perception and Performance,* **13(2),** 228–241.

Brigham, E.O. (1974) *The Fast Fourier Transform.* Prentice-Hall, Englewood Cliffs, NJ.

Broadbent, D.E. (1958) *Perception and Communication.* Pergamon, London.

Broadbent, D.E. (1971) *Decision and Stress.* Academic Press, London.

Broadbent, D.E. (1982) Task combination and selective intake of information. *Acta Psychologica,* **50,** 253–290.

Broadbent, D.E. (1987) Simple models for experimentable situations. In: P.E. Morris (ed.), *Modelling Cognition.* Wiley, London.

Brown, B. & Monk, T. (1975) The effect of local target and whole background constraint on visual search times. *Human Factors,* **17,** 81–88.

Bruce, M. & Foster, J.J. (1984) *The Visibility of Coloured Characters on Coloured Backgrounds in Viewdata Displays* (draft). Manchester Polytechnic. (Cited in Long, 1984).

Bury, K.F., Davies, S.E. & Darnell, M.J. (1985) Window management: a review of issues and some results from user testing. *IBM HFC-53.* IBM Human Factors Center, San Jose.

Butler, B.E. & Currie, A. (1986) On the nature of perceptual limits in vision. *Psychological Research,* **48,** 201–209.

Buswell, G.T. (1937) How adults read. *Supplementary Educational Monographs* (45). University of Chicago, IL.

Cakir, A., Hart, D.J. & Stewart, T.F.M. (1980) *Visual Display Terminals.* Wiley, Chichester.

Callaghan, T.C. (1991) Early perception: searching for textures and single elements. In: D. Brogan (ed.), *Visual Search II.* Taylor & Francis, London, (forthcoming).

Callaghan, T.C. (1990) Interference and dominance in texture segregation. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 80–87.

Campbell, F.W. & Robson, J.G. (1968) Application of Fourier analysis to the visibility of gratings. *Journal of Physiology,* **197,** 551–566.

Card, S.K., Moran, T.P. & Newell, A. (1983) *The Psychology of Human-Computer Interaction.* Erlbaum, Hillsdale, NJ.

Card, S.K., Pavel, M. & Farrell, J.E. (1985) Window-based computer dialogues. *Human-Computer Interaction—INTERACT'84.* Elsevier, Amsterdam, pp. 239–243.

Carr, K. (1988). Survey of eye movement measuring equipment. *AVA Newsletter,* (August). Applied Vision Association, London, pp. 10–24.

Carroll, J.M. (1982) Learning, using and designing command paradigms. *Human Learning,* **1,** 31–62.

Carter, E.C. & Carter, R.C. (1981) Color and conspicuity. *Journal of the Optical Society of America,* **71,** 723–729.

Chaikin, J.D., Corbin, H.H. & Volkmann, J. (1962) Mapping a field of short time visual search. *Science,* **137,** 1327–1328.

Chase, W.G. & Ericcson, K.A. (1981) Skilled memory. In: J.R. Anderson (ed.), *Cognitive Skills and their Acquisition.* Erlbaum, Hillsdale, NJ.

Chi (1989) Cited in Drury (1991).

Christ, R.E. (1975) Review and analysis of color coding research for visual displays. *Human Factors,* **17,** 542–570.

Cohen, B. (1981) Vestibular and oculomotor physiology. *Annals of the New York Academy of Science,* **374.**

Cole, B.L. & Hughes, P.K. (1984) A field trial of attention and search conspicuity *Human Factors,* **26,** 299–313.

Cole, B.L. & Jenkins, S.E. (1982) The effect of the density of background elements on the conspicuity of objects. *Vision Research,* **22,** 1241–1252.

Collewijn, H., van der Mark, F. & Jansen, T.C. (1975). Precise recording of human eye movement. *Vision Research,* **15,** 447–450.

Cowan, W. (1988) Antialiasing and image quality. In: *Society for Information Display 1988 Seminar Lecture Notes* (Vol. 2). SID, Playa del Rey, CA, p. 44.

Crane, H.D. & Steele, C.M (1978) Accurate three-dimensional eyetracker. *Applied Optics,* **17,** 691–705.

Crane, H.D. & Steele, C.M. (1984) *Generation V dual-Purkinje-image eyetracker.* SRI International, Menlo Park, CA.

Crassini, B. (1986) On the spatial distribution of visual attention. *Journal of Experimental Psychology: Human Perception and Performance,* **12,** 380–382.

Creed, A., Dennis, I., & Newstead, S. (1987) Proof-reading on VDUs. *Behaviour and Information Technology,* **6,** 3–13.

Crow, F.C. (1981) A comparison of antialiasing techniques. *IEEE Computer Graphics Applications,* **1,** 40–48.

Curry & Paxton (1980) *The Curpax Test Type.* Curry & Paxton, London.

Cushman, W.H. (1986) Reading from microfiche, from a VDT, and from the printed page; subjective fatigue and performance: a preliminary report. In: E. Grandjean (ed.), *Ergonomics and Health in Modern Offices.* Taylor & Francis, London, pp. 322–328.

Cushman, W.H. & Miller, R.L. (1988) Resolution and gray-scale requirements for the display of legible alphanumeric characters. In: J. Morreale (ed.), *1988 Society for Information Display International Symposium Digest of Technical Papers.* SID, Playa del Rey, CA, pp. 432–434.

Cypher, A. (1986) The structure of users' activities. In: D.A. Norman & S.W. Draper (eds.), *User Centred System Design: New Perspectives on Human–Computer Interaction.* Erlbaum, Hillsdale, NJ, pp. 243–264.

Davies, S.P. (1990) The nature and development of programming plans. *International Journal of Man–Machine Studies,* **32,** 461–481.

Davies, S.P., Lambert, A.J. & Findlay, J.M. (1989) The effects of availability of menu information during command learning in a word processing application. *Behaviour and Information Technology,* **2,** 135–144.

De Groot, A.D. (1978) *Thought and Choice in Chess.* Mouton, The Hague.

Dehaene, S. (1989) Discriminability and dimensionality effects in visual search for featural conjunctions: A functional pop-out. *Perception and Psychophysics,* **46,** 72–80.

Dennett, D. (1978) *Brainstorms.* MIT Press, Cambridge, MA.

Derrington, A.M., Krauskopf, J. & Lennie, P. (1984) Chromatic mechanisms in laterate geniculate nucleus of macaques. *Journal of Physiology,* **357,** 241–265.

Diaper, D. (1989) *Task Analysis for Human-Computer Interaction.* Ellis Horwood, Chichester.

Ditchburn, R.W. (1980) The function of small saccades. *Vision Research,* **20,** 271–272.

Draper, S.W. (1984) The nature of expertise in UNIX. In: B. Shackel (ed.), *INTERACT'84: First IFIP Conference on Human-Computer Interaction.* IFIP, Amsterdam.

Drury, C.G. (1991) Exploring search strategy. In: D. Brogan (ed.), *Visual Search II.* Taylor & Francis, London, (forthcoming).

Drury, C.G. (1990) Visual search in industrial inspection. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 263–275.

Duncan, J. (1980) The locus of interference in the perception of simultaneous stimuli. *Psychological Review,* **87,** 272–300.

Easterby, R.S. (1970) The perception of symbols for machine displays *Ergonomics,* **13,** 149–158.

Egeth, H. & Pachella, R. (1969) Multidimensional stimulus identification. *Perception and Psychophysics,* **5,** 314–365.

Egeth, H., Jonides, J. & Wall, S. (1972) Parallel processing of multi-element displays. *Cognitive Psychology,* **3,** 674–698.

Ellis, S.R. & Smith, J.D. (1985) Patterns of statistical dependency in visual scanning. In: R. Groner, G.W. McConkie & C. Menz (eds.), *Eye Movements and Human Information Processing.* Elsevier, Amsterdam, pp. 221–238.

Engel, F.L. (1971) Visual conspicuity: directed attention and retinal locus. *Vision Research,* **11,** 563–576.

Engel, F.L. (1977) Visual conspicuity: visual search and fixation tendencies of the eye. *Vision Research,* **17,** 95–108.

Enns, J. (1986) Seeing textons in context. *Perception and Psychophysics,* **39,** 32–38.

Enns, J.T. (1990) Three-dimensional features that pop out in visual search. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 37–45.

Enns, J.T. & Gilani, A.B. (1988) Three-dimensionality and discriminability in the object-superiority effect. *Perception and Psychophysics,* **44,** 243–256.

Enns, J.T. & Prinzmetal, W. (1984) The role of redundancy in the object-line effect. *Perception and Psychophysics,* **35,** 22-32.

Enns, J.T. & Rensink, R.A. (1991) Rapid visual search. In: D. Brogan (ed.), *Visual Search II.* Taylor & Francis, London, (forthcoming).

Enoch, J.M. (1959) Effect of the size of a complex display upon visual search. *Journal of the Optical Society of America,* **49,** 280-286.

Eriksen, C.W. & Eriksen, B.A. (1971) Visual perceptual processing rates and backward and forward masking. *Journal of Experimental Psychology,* **89**, 306–313.

Eriksen, C.W. & Hoffman, J.E. (1972) Temporal and spatial characteristics of selective encoding from visual displays. *Perception and Psychophysics,* **12**, 201–204.

Eriksen, C.W. & Murphy, T. (1987) Movement of the attentional focus across the visual field: A critical look at the evidence. *Perception and Psychophysics,* **14**, 155–160.

Eriksen, C.W. & St. James, J.D. (1986) Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics,* **40**, 225–240.

Eriksen, C.W. & Webb, J.M. (1989) Shifting of attentional focus within and about a visual display. *Perception and Psychophysics,* **45**, 175–183.

Eriksen, C.W. & Yeh, Y. (1985) Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance.* **11**, 583–597.

Fairclough, S.H. & Maternaghan, M.K. (1991) Changes in driver's visual behaviour due to the introduction of complex vs simple route navigation information. In: D. Brogan (ed.), *Visual Search II.* Taylor & Francis, London, (forthcoming).

Farrell, J.E., Casson, E.J., Haynie, C.R. & Benson, B.L. (1988) Designing flicker-free video display terminals. *Displays,* **9**, 115–122.

Fechner, G.T. (1860) Elemente der Psychophysik. (H.E. Adler, trans.), *Elements of Psychophysics.* Holt, Rinehart & Winston, New York.

Fechner, G.T. (1871) Zur experimentalen Aesthetic. *(Toward an Experimental Aesthetic.)* Berlin.

Fernandez, D.P. & Alberola, C.R. (1990) Do visual search variables affect memory search. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 63–71.

Ferwerda, J.A. & Greenberg, D.P. (1988) A psychophysical approach to assessing the quality of antialiased images. *IEEE Computer Graphics and Applications,* **8**, 85–95.

Findlay, J.M. & Scott, D. (1990) Visual science and the user interface: what should go where? In: *Proceedings of the Annual Meeting of the Applied Vision Association.* AVA, Manchester, p. 8

Findlay, J.M., Davies, S.P., Kentridge, R., Lambert, A.J. & Kelly, J. (1988) Optimum display arrangements for presenting visual reminders. In: D.M. Jones & R. Winder (eds.), *People and Computers IV.* CUP, pp. 453–464.

Findlay, D. & Wilkinson, J. (1984) The effects of glare on the contrast sensitivity function. *Human Factors,* **26**, 283–287.

Fisher, D.L., Coury, B.G., Tengs, T.O. & Duffy, S.A. (1989) Minimising the time to search visual displays: the role of highlighting. *Human Factors,* **31**, 167–182.

Foley, J.D., Wallace, V.L. & Chan, P. (1984) The human factors of computer graphics interaction techniques. *IEEE Computer Graphics and Applications,* **4,** 13–48.

Ford, A., White, C.T. & Lichtenstein, M. (1959) Analysis of eye-movements during free search. *Journal of the Optical Society of America,* **49,** 287–292.

Francolini, C.M. & Egeth, H.E. (1980) On the nonautomaticity of 'automatic' activation: evidence of selective seeing. *Perception and Psychophysics,* **27,** 331–342.

Friedman, A. & Liebelt, L.S. (1981) On the time course of viewing pictures with a view toward remembering. In: D.F. Fisher, R.A. Monty & J.W. Senders (eds.), *Eye Movements: Cognition and Visual Perception.* Erlbaum, Hillsdale, NJ, pp. 137–155.

Gait, J. (1986) Pretty pane tiling of pretty windows. *IEEE Software,* September, 9–14.

Gardiner, M.M. & Christie, B. (1987) *Applying Cognitive Psychology to User-interface Design.* Wiley, Chichester.

Garner, W.R. (1983) Asymmetric interactions of stimulus dimensions in perceptual information processing. In: T.J. Tighe & B.E. Shepp (eds.), *Perception, Cognition and Development: Interactional Analyses.* Erlbaum, Hillsdale, NJ.

Gerrissen, J.F. (1990) Computer emulation of preattentive mechanisms in vision: emergent WHAT and WHERE. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 277–291.

Gibson, J.J. (1966) *The Senses Considered as Perceptual Systems.* Houghton Mifflin, Boston.

Gibson, J.J. (1979) *The Ecological Approach to Visual Perception.* Houghton Mifflin, Boston.

Gibson, J.J. & Gibson, E.J. (1955) Perceptual learning: Differentiation or enrichment? *Psychological Reviews,* **62,** 1–32.

Ginsburg, A.P. (1979) Spatial filtering and vision: implications for normal and abnormal vision. In: J. Enoch, Proenza & Jampolski (eds.), *Clinical Applications of Visual Psychophysics.*

Ginsburg, A.P., Evans, D.W., Sekuler, R. & Harp, S.A. (1982) Contrast sensitivity and pilots' performance in aircraft simulators. *American Journal of Optometry and Physiological Optics,* **59,** 105–108.

Giroux, L. & Belleau, R. (1986) What's on the menu? The influence of menu content on the selection process. *Behaviour and Information Technology,* **5,** 169–172 .

Gittens, D. (1986) Icon–based human–computer interaction. *International Journal of Man–Machine Studies,* **24,** 519–543.

Gleitman, H. & Jonides, J. (1976) The cost of categorization in visual search: Incomplete processing of targets and field items. *Perception and Psychophysics,*

20, 281–288.

Gottsdanker, R. (1960) The relation between the nature of the search situation and the effectiveness of alternative strategies of search. In: *Visual Search Techniques.* NAS-NRC Pub. 712, pp. 181–186.

Gould, J.D. & Dill, A.B. (1969) Eye-movement parameters and pattern discrimination. *Perception and Psychophysics,* **6**, 311-320.

Gould, J.D. & Grischkowsky, N. (1982) Doing the same work with a CRT terminal and with hardcopy. In: *Proceedings of the Human Factors Society 26th Annual Meeting.* HFS, Santa Monica, pp. 165–166.

Gould, J.D. & Grischkowsky, N. (1984) Doing the same work with hardcopy and with cathode ray tube (CRT) computer terminals. *Human Factors,* **26**, 323–337.

Gould, J.D., Alfaro, L., Barnes, V., Finn, R., Grischkowsky, N. and Minuto, A. (1987*a*) Reading is slower from CRT displays than from paper: Attempts to isolate a single variable explanation. *Human Factors,* **29**, 269–299.

Gould, J.D., Alfaro, L. Finn, R., Haupt, B. & Minuto, A. (1987*b*) Reading from CRT displays can be as fast as reading from paper. *Human Factors,* **29**, 497–517.

Gould, J.D., Alfaro, L., Finn, R., Haupt, B. Minuto, A. & Salaun, J. (1987*c*) Why reading was slower from CRT displays than from paper. In: J.M. Carroll & P.P. Tanner (eds.), *Human Factors in Computing Systems—IV and Graphics Interface.* Elsevier, Amsterdam, pp. 7–11.

Goulet, C., Bard, C. & Fleury, M. (1990) Visual search strategies and information processing in a racquet sport situation. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 185–192.

Graf, W. & Krueger, H. (1989) Ergonomic evaluation of user-interfaces by means of eye-movement data. In: M.J. Smith & G. Salvendy (eds.), *Work with Computers: Organizational, Management, Stress and Health Aspects.* Elsevier, Amsterdam, pp. 659–665.

Graf, W., Elsinger, P. & Krueger, H. (1987) Methods for the ergonomical evaluation of alphanumeric computer-generated displays. In: H.-J. Bullinger & B. Shackel (eds.), *INTERACT'87: Human-Computer Interaction.* Elsevier, Amsterdam, pp. 349–353.

Graham, N. (1985) Detection and identification of near-threshold visual patterns. *Journal of the Optical Society of America,* **2**, 1468–1482.

Grandjean, E. & Vigliani, E. (1980) *Ergonomic Aspects of Visual Display Terminals.* Taylor & Francis, London.

Grice, G.R., Canham, L. & Schafer, C. (1982) Development of associative and perceptual interference. *Perception and Psychophysics,* **32**, 375–387.

Grindley, G.C. & Townsend, V. (1968) Voluntary attention in peripheral vision and its effects on acuity and differential thresholds. *Quarterly Journal of Experimental Psychology,* **20,** 11–19.

Grindley, G.C. & Townsend, V. (1970) Visual search without eye movements. *Quarterly Journal of Experimental Psychology,* **22,** 62–67.

Gupta, S. (1986) *YODA: An advanced display for personal computers.* IBM Research Report RC-11618. IBM, Yorktown Heights, NY.

Haber, R.N. (1970) How we remember what we see. *Scientific American,* **222,** 104–115.

Haber, R.N. & Hershenson, M. (1980) *Psychology of Visual Perception* (2nd edn.). Holt, Rinehart & Winston, New York.

Halbert, D. (1984) *Programming by Example.* Unpublished doctoral dissertation, Stanford University, CA.

Hallett, P.E. (1986) Eye movements. In K. Boff, L. Kaufman & J. Thomas (eds.), *Handbook of Perception and Human Performance I.* pp. 10.25–10.28.

Heeley, D.W. & Timney, B. (1989) Spatial frequency discrimination at different orientations. *Vision Research,* **29,** 1221–1228.

Helander, M.G., Billingsley, P.A. & Schurick, J.M. (1984) An evaluation of human factors research on visual display terminals in the workplace. In: F.A. Muckler (ed.), *Human Factors Review, 1984.* HFS, Santa Monica, CA.

Helmholtz, H. von (1925) *Physiological Optics,* 1850. (J.P.C. Southall, ed.) Optical Society of America.

Helsen, W. & Pauwels, J.M. (1991) A cognitive approach to visual search in sport. In: D. Brogan (ed.), *Visual Search II.* Taylor & Francis, London, (forthcoming).

Hemenway, K. (1982) Psychological issues in the use of icons in command menus. In: *Proceedings of the Conference on Human Factors in Computing Systems,* Gaithersburg, Maryland, pp. 20–23.

Heppner, F.H., Anderson, J.G.T., Farstrup, A.E. & Weiderman, N.H. (1985) Reading performance on a standardized test is better from print than from computer display. *Journal of Reading,* **20,** 321–325.

Hewitt, D.R., Kennedy, P.J. & Scrivener, R. (1989) Which do people prefer? Text displayed on a monitor using anti-aliased fonts, text displayed using (normal) bi-level fonts, or text printed on paper? *Technical Report,* IBM (UK), Hursley.

Hochberg, J. (1970) Attention, organization and consciousness. In: *D.I Mostofsky, (ed.), Attention: Contemporary Theory and Analysis.* Appleton-Century-Crofts, New York, pp. 99–124.

Hoffman, J.E. (1975) Hierarchical stages in the processing of visual information. *Perception and Psychophysics,* **18,** 348–354.

Hoffman, J.E. (1979) A two-stage model of visual search. *Perception and Psy-*

chophysics, **25,** 319–327.

Hoffman, J.E. (1980) Interaction between global and local levels of a form. *Journal of Experimental Psychology: Human Perception and Performance,* **6,** 222-234.

Hubel, D.H. & Weisel, T.N. (1968) Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology,* **195,** 215–243.

Hutchins, E.L., Hollan, J.D. & Norman, D.A. (1986) Direct manipulation interfaces. In: D.A. Norman & S.W. Draper (eds.), *User Centred System Design: New Perspectives on Human–Computer Interaction.* Erlbaum, Hillsdale, NJ, pp. 87–124.

Hwang, S.-L., Wang, M.-Y. & Ho, C.-C. (1987) The study of Chinese character information on a video display terminal. In: S.S. Asfour (ed.), *Trends in Ergonomics/Human Factors IV.* Elsevier, Amsterdam, pp. 145–151.

IBM (1990) *XGA Device Driver On-line Guidance.* IBM Corporation.

Inhoff, A.W. & Rayner, K. (1986) Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics,* **40,** 431–439.

Intons-Peterson, M. & Fournier, J. (1986) External and internal memory aids: when and how often do we use them? *Journal of Experimental Psychology: General,* **115,** 267–280.

Isensee, S. (1982) The perception of flicker and glare on computer CRT displays. In: *Proceedings of the 26th Annual Meeting of the Human Factors Society.* HFS, Baltimore.

Jacobs, A.M. (1986) Eye movement control in visual search: how direct is visual span control? *Perception and Psychophysics,* **39,** 47–58.

Jenkins, S.E. & Cole, B.L. (1982) The effect of density of background elements on the conspicuity of objects. *Vision Research,* **22,** 1241–1252.

Jervell, H.R. & Olsen, K.A. (1985) *Behaviour and Information Technology,* **4,** 249–254.

Johnston, W.A. & Dark, V.J. (1986) Selective attention. *Annual Review of Psychology,* **37,** 43–75.

Jonides, J. (1981) Voluntary versus automatic control over the mind's eye's movement. In: J.B. Long & A.D. Baddeley (eds.), *Attention and Performance, IX.* Erlbaum, Hillsdale, NJ.

Jonides, J. (1983) Further toward a model of the mind's eye movement. *Bulletin of the Psychonomic Society,* **21,** 247–250.

Jonides, J. & Gleitman, H. (1972) A conceptual category effect in visual search: O as letter or a digit. *Perception and Psychophysics,* **12,** 457–460.

Julesz, B. (1984) Towards an axiomatic theory of preattentive vision. In: G.M. Edelman, W. Einar-Gall & W. Maxwell Cohen (eds.), *Dynamic Aspects of Neocortical*

Function. Neurosciences Research Foundation, New York.

Julesz, B. & Bergen, J.R. (1983) Textons, the fundamental element in preattentive vision and perception of textures. *Bell System Technical Journal,* **62,** 1619–1645.

Kahneman, D.C. (1973) *Attention and Effort.* Prentice-Hall, New Jersey.

Kak, A.V. (1981) Relationships between readability of printed and CRT-displayed text. In: *Proceedings of the Human Factors Society 25th Annual Meeting.* Human Factors Society, Santa Monica, pp. 137–140.

Kanaya, S. (1990) Vision and visual environment for VDT work. *Ergonomics,* **33,** 775–785.

Kaufman, L. & Richard, W. (1969) Spontaneous fixation tendencies for visual forms. *Perception and Psychophysics,* **5,** 85–88.

Keeley, S.M. (1969) Visual detection as a function of attentional demands and perceptual system error. *Perception and Psychophysics,* **6,** 73–77.

Kelly, D.H. (1961) Visual responses to time-dependent stimuli. I. Amplitude sensitivity measurements. *Journal of the Optical Society of America,* **52,** 940.

Kinchla, R.A. & Wolfe, J.M. (1979) The order of visual processing: 'Top-down', 'bottom-up', or 'middle-out'. *Perception and Psychophysics,* **25,** 225–231.

Kinchla, R.A., Solis-Macias, V. & Hoffman, J. (1983) Attending to different levels of structure in a visual image. *Perception and Psychophysics,* **33,** 1–10.

Kleigl, R. & Olson, R.K. (1981) Reduction and calibration of eye movement data. *Behavior Research Methods, Instrumentation and Computers,* **13,** 107–111.

Klein, R.M. (1980) Chronometric analysis of saccadic eye movements: Reflexive and cognitive control. In: D. Landers & R. Christina (eds.), *The Psychology of Motor Behaviour and Sport: 1977.*

Klein, R. & Farrell, M. (1989) Search performance without eye movements. *Perception and Psychophysics,* **46,** 476–482.

Kleiner (1989) Cited in Drury (1991).

Kohler, W. (1929) *Gestalt Psychology.* Liveright, New York.

Kolers, P.A. (1969) Some formal characteristics of pictograms. *American Scientist,* **57,** 348–363.

Kolers, P.A., Duchnicky, R.L. & Ferguson, D.C. (1981) Eye movement measurement of readability of CRT displays. *Human Factors,* **23,** 517–527.

Korfhage, R.R. & Korfhage, M.A. (1984) The nature of visual languages. *Proceedings of the 1984 IEEE Workshop on Visual Languages.* IEEE, Hiroshima, pp. 177–183.

Krantz, J.H. & Silverstein, L.D. (1989) The gray-scale asymptote for anti-aliasing graphic images on color matrix displays. In: *Society for Information Display International Symposium Digest of Technical Papers,* (Vol. XX.) Baltimore, MD:

SID, pp. 139–141.

Kruk, R.S. & Muter, P. (1984) Reading from continuous text on video screens. *Human Factors,* **26**, 339–346.

Kuhn, L.S. (1970) *The Structure of Scientific Revolutions.* University of Chicago Press.

La Berge, D. (1983) Spatial extent of attention to letters in words. *Journal of Experimental Psychology: Human Perception and Performance,* **9**, 371–379.

Lamb, M.R. & Robertson, L.C. (1988) Do response time advantage and interference reflect the order of processing of global- and local-level information? *Perception and Psychophysics,* **46**, 254–258.

Lamb, M.R., Robertson, L.C. & Knight, R.T. (1989) Attention and interference in the processing of hierarchical patterns: Information from patients with right and left temporal-parietal lesions. *Neuropsychologia,* **27**, 471–483.

Lambert, A.J. & Hockey, G.R. (1986) Selective attention and performance with a multidimensional visual display. *Journal of Experimental Psychology: Human Perception and Performance,* **12**, 484–495.

Lansdale, M.W. (1988) On the memorability of icons in an information retrieval task. *Behaviour and Information Technology,* **7**, 131–151.

Lansdale, M.W., Simpson, M. & Stroud, T.R.M. (1987) Comparing words and icons as cue enrichers in an information retrieval task. In: H.-J. Bullinger & B. Shackel, (eds.), *Human-Computer Interaction—INTERACT'87.* Elsevier, Amsterdam, pp. 911–916.

Lansdale, M.W., Jones, M.R. & Jones, M.A. (1989) Visual search in iconic and verbal interfaces. In: E.D. Megaw (ed.), *Contemporary Ergonomics 1989.* Taylor & Francis, London, pp. 422–429.

Lasaga, J.M. (1989) Gestalts and their components: Nature of information precedence. In: B. Shepp & S. Ballesteros (eds.), *Object Perception: Structure and Process.* Erlbaum, Hillsdale, NJ.

Lasaga, J.M. & Cecala, A.J. (1990) Similarity between targets and stimulus patterns at different levels of structural complexity. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 89–97.

Laurel, B.K. (1986) Interface as mimesis. In: D.A. Norman & S. Draper (eds.), User Centred Design: New Perspectives on Human-Computer Interaction. Erlbaum, Hillsdale, NJ, pp. 67–86.

Lee, E.S. & MacGregor, J.N. (1985) Minimizing user search time in menu retrieval systems. *Human Factors,* **27**, 157–162.

Lee, R.P.-J. & Rabideau, G.F. (1979) Digital image quality as a function of sampling density and gray levels. In: *Occupational Health and Safety: Proceedings of the*

12th Annual Meeting of the Human Factors Association of Canada. Human Factors Association of Canada, pp. 9.1–9.5.

Leler, W.J. (1980) Human vision, anti-aliasing, and the cheap 4000 line display. *Proceedings of SIGGRAPH 1980.* ACM, pp. 308–313.

Lennie, P. & D'Zmuru, M. (1988) Mechanisms of colour vision. *CRC Critical Reviews in Neurobiology,* **3,** 333–400.

Lévy-Schoen, A. & O'Regan, K. (1979) The control of eye movements in reading. In: P.A. Kolers, M.E. Wrolstad & H. Bouma (eds.), *Processing of Visible Language.* Plenum, New York.

Lévy-Schoen, A., O'Regan, J.K., Jacobs, A.M. & Coeffe, C. (1984) The relation between visibility span and eye movements in various scanning tasks. In: A.G. Gale & F. Johnson (eds.), *Theoretical and Applied Aspects of Eye Movement Research.* Elsevier, Amsterdam, pp. 133–142.

Lie, I. (1969) The momentary sensitivity contour of the retina: an approach to the study of attention. *Stud. Psychol.,* **11,** 157–160.

Lodding, K.N. (1983) Iconic interfacing *IEEE Computer Graphics and Applications,* **3,** 11–20.

Loftus, G.R. (1972) Eye fixations and recognition memory for pictures. *Cognitive Psychology,* **3,** 525–551.

Long, J. (1987) Cognitive ergonomics and human-computer interaction. In: P. Warr (ed.), *Psychology at Work,* (3rd edn). Penguin, Harmondsworth.

Long, J. (1989) Cognitive ergonomics and human-computer interaction: An introduction. In: J. Long & A. Whitefield (eds.), *Cognitive Ergonomics and Human-Computer Interaction.* CUP, pp. 4–34.

Long, T. (1984) Human factors principles for the design of computer colour graphics displays. *British Telecommunications Technology Journal,* **2,** 1–11.

Lunn, R. & Banks, W.P. (1986) Visual fatigue and spatial frequency adaptation to video displays of text. *Human Factors,* **28,** 457–464.

McConkie, G.W. & Rayner, K. (1975) The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics,* **17,** 578–586.

McCormick, E.J. (1964) *Human Factors Engineering.* McGraw-Hill, New York.

McCormick, E.J. & Sanders, M.S. (1982) *Human Factors in Engineering and Design.* McGraw-Hill, New York.

MacGregor, J.N. & Lee, E.S. (1987) Menu search: random or systematic? *International Journal of Man-Machine Studies,* **26,** 627–631.

MacGregor, J.N., Lee, E.S. & Lam, N. (1986) Optimizing the structure of menu indexes: a decision model of menu search. *Human Factors,* **28,** 387–399.

Mackworth, N.H. (1976) Stimulus density limits the useful field of view. In: R.A.

Monty & J.W. Senders (eds.), *Eye Movements and Psychological Processes.* Erlbaum, Hilldsdale, NJ.

McLeod, P., Driver, J. & Crisp, J. (1988) Visual search for a conjunction of movement and form is parallel. *Nature,* **332,** 154–155.

Maffei, L. & Fiorentini, A. (1977) The unresponsive regions of visual cortical receptive fields. *Vision Research,* **16,** 1255-1267.

Manes, S. (1985) Pushing picture-perfect programmes: smash that icon! *PC Magazine,* June, p. 64.

Marr, D. (1982) *Vision.* Freeman, San Francisco.

Marshak, W.P. & Osarczuk, J.C. (1987) Signal detection performance as a function of Fourier description of symbols. In: *Proceedings of the Human Factors Society 31st Annual Meeting.* HFS.

Martin, M. (1979) Local and global processing: The role of sparsity. *Memory and Cognition,* **7,** 476–484.

Matin, E. (1974) Saccadic suppression: A review and an analysis. *Psychological Bulletin,* **81,** 899–917.

May, J.G., Kennedy, R.S., Williams, M.C., Dunlap, W.P. & Brannan, J.R. (1990) Eye movement indices of mental workload. *Acta Psychologica,* **75,** 75–89.

Maylor, E.A. & Hockey, G.R.J. (1985) Inhibitory component of externally controlled covert orienting in visual space. *Journal of Experimental Psychology: Human Perception and Performance,* **11,** 777-787.

Maylor, E.A. & Hockey, G.R.J. (1987) Effects of repetition on the facilitatory and inhibitory components of orienting in visual space. *Neuropsychologia,* **25,** 41–54.

Megaw, E.D. & Richardson, J. (1979) Target uncertainty and visual scanning strategies. *Human Factors,* **21,** 303–315.

Merchant, J., Morrisette, R. & Porterfield, J.L. (1974). Remote measurement of eye direction allowing subject motion over one cubic foot. *IEEE Transactions of Biomedical Engineering,* **BME-21,** 309–317.

Miller, J. (1981) Global precedence in attention and decision. *Journal of Experimental Psychology: Perception and Performance,* **7,** 1161–1174.

Miller, M.E., Crespo, R.T. & Kennedy, P.J. (1989) An investigation of a color anti-aliased CRT typeface. *Technical Report TR54.518,* IBM, Boca Raton.

Millodot, M. (1982) Accomodation and refraction of the eye. In: H.B. Barlow & J.D. Mollon (eds.), *The Senses.* CUP, pp. 62–81.

Mills, C.B. & Weldon, L.J. (1984) Reading from Computer Screens. University of Maryland Human/Computer Interaction Laboratory, *CAR-TR-94.*

Miyao, M., Hacisalihzade, S.S., Allen, J.S. & Stark, L.W. (1989) Effects of VDT resolution on visual fatigue and readability: an eye movement approach. *Er-*

gonomics, **32**, 603–614.

Miyata, Y. & Norman, D.A. (1986) Psychological issues in the support of multiple activities. In: D.A. Norman & S.W. Draper (eds.), *User Centred System Design: New Perspectives on Human–Computer Interaction.* Erlbaum, Hillsdale, NJ, pp. 265–284.

Mohr, W. (1984) *Visuelle Wahrnehmung und Zeichenfunktion.* Roderer, Regensburg.

Monk, T.H. (1974) Sequential effects in visual search. *Acta Psychologica,* **38**, 315–312.

Monk. T.H. (1976) Target uncertainty in applied visual search. *Human Factors,* **18**, 607–612.

Monk, T.H. (1977) Sequential expectancy in visual search. *Human Factors,* **19**, 601–606.

Monk, T.H. (1981*a*) Search. In: J.S. Warm (ed.), *Sustained Attention in Human Performance.* Wiley, London.

Monk, T.H. (1981*b*) The interaction between the edge effect and target conspicuity in visual search. *Human Factors,* **31**, 167–182.

Moray, N. (1959) Attention in dichotic listening: affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology,* **11**, 56–60.

Moray, N. (1975) A data base for theories of selective listening. In P.M.A. Rabbitt & S. Dornic (eds.), *Attention and Performance V.* Academic Press, New York, pp. 119–135.

Morris, D., Theaker, C.J., Phillips, R. & Love, W. (1988) Human-computer interface recording. *The Computer Journal,* **5**, 437-444.

Morrison, R.E. (1984) Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance,* **10**, 667–682.

Mourand, R.R. & Rockwell, T.H. (1972) Strategies of visual search by novice and experienced drivers. *Human Factors,* **14**, 325–335.

Müller, H.J., Humphreys, G.W., Quinlan, P.T. & Riddoch, M.J. (1990) Combined-feature coding in the form domain. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 47–55.

Murphy, T.D. & Eriksen, C.W. (1987) Temporal changes in the distribution of attention in the visual field in response to precues. *Perception and Psychophysics,* **15**, 149–158.

Muter, P. & Mayson, C. (1986) The role of graphics in item selection from menus. *Behaviour and Information Technology,* **5**, 89–95.

Muter, P., Latrémouille, S.A., Treurniet, W.C. & Beam, P. (1982) Extended reading of continuous text on television screens. *Human Factors,* **24,** 501–508.

Muthig, K.P. & Reinhardt, G. (1987) Comparison of verbal and iconic text editing commands: some results from scanning studies. *Zeitschrift für Psychologie.* Cited in Arend *et al.*, 1987.

Naiman, A. & Farrell, J. (1988) Modelling the display and perception of gray-scale characters. In: J. Morreale (ed.), *1988 Society for Information Display International Symposium Digest of Technical Papers.* SID, Playa del Rey, CA, pp. 424–427.

Naka, K.-I. & Rushton, W.A.H. (1966) S-potentials from colour units in the retina of fish (*Cyprinidae*). *Journal of Physiology,* **185,** 536–555.

Nakayama, K. & Silverman, G.H. (1986) Serial and parallel encoding of visual feature conjunctions. *Investigative Ophthalmology and Visual Science.*

National Research Council, Committee on Computerized Speech Recognition Technologies (1984) *Automatic speech recognition in severe environments.* NRC, Commission on Engineering and Technical Systems, Washington DC.

Nattkemper, D. (1991) Target detection in continuous visual search. In: D. Brogan (ed.), *Visual Search II.* Taylor & Francis, London, (forthcoming).

Nattkemper, D. & Prinz, W. (1990) Eye movements in continuous search: effects of category relationship and stimulus degradation. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 125–132.

Navon, D. (1977) Forest before trees: The precedence of global features in visual perception. *Cogntive Psychology,* **9,** 353–383.

Navon, D. (1981) Do attention and decision follow perception? Comment onMiller. *Journal of Experimental Psychology: Human Perception and Performance,* **9,** 1175–1182.

Navon, D. & Norman, J. (1983) Does global precedence really depend on visual angle? *Journal of Experimental Psychology: Human Perception and Performance,* **9,** 955–965.

Neale, M.D. (1958) *Neale Analysis of Reading Ability.* MacMillan Educational, London.

Neely. J. (1977) Semantic priming and retrieval from lexical memory: roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General,* **106,** 226–254.

Neil, A.S. & Simons, R.M. (1984) Playback: A method for evaluating the usability of software and its documentation. *IBM Systems Journal,* **23,** 78-82.

Neisser, U. (1963) Decision-time without reaction-time: Experiments in visual scanning. *American Journal of Psychology,* **76,** 376–385.

Neisser, U. (1964) Visual search. *Scientific American,* **210,** 94–101.

Neisser, U. (1967) *Cognitive Psychology.* Appleton-Century-Crofts, New York.

Neisser, U. & Beller, H.K. (1965) Searching through word lists. *British Journal of Psychology,* **56,** 349–358.

Neisser, U. & Stoper, A. (1965) Redirecting the search process. *British Journal of Psychology,* **56,** 359–368.

Neisser, U. Novick, R. & Lazar, R. (1963) Searching for ten targets simultaneously. *Perceptual and Motor Skills,* **17,** 955–961.

Newsted, P.R. (1985) Paper versus online presentations of subjective questionnaires. *International Journal of Man-Machine Studies,* **23,** 231–247.

Nodine, C.F. (1982) Compositional design as a perceptual determinant of aesthetic judgement. *Rev. Res. Vis. Arts Educ.,* **15,** 43–54.

Nodine, C.F. & Kundel, H. (1980) Hidden figures, eye movements, and the perception of art. In: R. Groner & P. Fraisse (eds.), *Cognition and Eye Movements.* VEB, Berlin.

Norman, D.A. (1986) Cognitive engineering. In: D.A. Norman & S.W. Draper (eds.), *User Centred System Design: New Perspectives on Human-Computer Interaction.* Erlbaum, Hillsdale, NJ.

Norman, D.A. & Bobrow, D.G. (1975) On data-limited and resource-limited processes. *Cognitive Psychology,* **7,** 44–64.

Noton, D. (1970) A theory of visual pattern perception. *IEEE Trans. Syst. Sci. Cybernet.,* **6,** 349–357.

Noton, D. & Stark, L. (1971*a*) Scanpaths in eye movements during pattern perception. *Science,* **171,** 308-311.

Noton, D. & Stark, L. (1971*b*) Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research,* **11,** 929–942.

Oborne, D.J. & Holton, D. (1988) Reading from screen versus paper: there is no difference. *International Journal of Man-Machine Studies,* **28,** 1–9.

Oldfield, R.C. (1960) Statement in *Visual Search Techniques.* NAS-NRC pub. 712, pp. 240–246.

Oléron, G. & Tardieu, H. (1978) Influence of scrolling upon the recall of texts. In: M.M. Gruneberg, P.E. Morris & R.N. Sykes (eds.), *Practical Aspects of Memory.* Academic Press, London.

Oppenheim, A.V. & Schafer, R.W. (1975) *Digital Signal Processing.* Prentice-Hall, New Jersey.

O'Regan, J.K. (1979) Saccade size control in reading: evidence for the linguistic control hypothesis. *Perception and Psychophysics,* **25,** 501–509.

O'Regan, J.K. (1983) Elementary perceptual and eye movement control processes in reading. In: K. Rayner (ed.), *Eye Movements in Reading: Perceptual and*

Language Processes. Academic Press, New York, pp. 121–138.

O'Regan, J.K. & Lévy-Schoen, A. (1987) Eye movement strategy and tactics in word recognition and reading. In: M. Coltheart (ed.), *Attention and Performance XII, The Psychology of Reading.* Erlbaum, Hillsdale, NJ, pp. 363–383.

O'Regan, J.K., Lévy-Schoen, A. & Jacobs, A.M. (1983) The effect of visibility on eye movement parameters in reading. *Perception and Psychophysics,* **34,** 457–464.

Pachella, R.G. & Fisher, D.F. (1969) Effects of stimulus degradation and stimulus similarity on the tradeoff between speed and accuracy in absolute judgements. *Journal of Experimental Psychology,* **81,** 7–9.

Parasuraman, R. & Davies, D.R. (1984) (eds.) *Varieties of Attention.* Academic Press, New York.

Papin, J.P., Asset, A.M., Bategat, D. Dejours, C. & Viard, D. (1976) Technique d'étude de l'utilisation visuelle des instruments de bord par l'équipage. *Médicine Aéronotique et Spatiale, Medicine Subaquatique et Hyperbare,* **60,** 251–254.

Parkinson, S.R., Hill, M., Sisson, N. & Viera, C. (1988) Effects of breadth, depth and number of responses on computer menu search performance. *International Journal of Man-Machine Studies,* **28,** 683–692.

Parton, D., Huffman, K., Pridgem, P., Norman, K. & Shneiderman, B. (1985) Learning a menu selection tree: training methods compared. *Behaviour and Information Technology,* **4,** 81–91.

Pashler, H. (1987) Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception and Psychophysics,* **41,** 191–201.

Paterson, D.G. & Tinker, M.A. (1940*a*) *How to Make Type Readable.* Harpers, London.

Paterson, D.G. & Tinker, M.A. (1940*b*) Influence of line width on eye movements. *Journal of Experimental Psychology,* **27,** 572–577.

Paterson, D.G. & Tinker, M.A. (1947) The effect of typography upon the perceptual span in reading. *American Journal of Psychology,* **60,** 388–396.

Pelli, D. (1988) Reading, contrast polarity, and contrast adaptation. *Investigative Ophthalmology and Visual Science* (Suppl), **29,** 142.

Pentland, A.P. (1986) Perceptual organization and the representation of natural form. *Artificial Intelligence,* **28,** 293–331.

Perlman, G. (1984) Making the right choices with menus. In: B. Shackel (ed.), *Proceedings of the IFIP Conference on Human-Computer Interaction — INTERACT' 84,* Elsevier, Amsterdam.

Perry, V.H. & Cowey, A. (1985) The ganglion cell and cone distributions in the monkey's retina: implications for central magnification factors. *Visual Research,* **25,** 1795–1810.

Podgorny, P. & Shepard, R.N. (1983) The distribution of visual attention over space. *Journal of Experimental Psychology: Human Perception and Performance,* **9,** 380–393.

Pomerantz, J.R. (1981) Perceptual organization in information processing. In: M. Kubovy & J.R. Pomerantz (eds.), *Perceptual Organisation.* Erlbaum, Hillsdale, NJ.

Pomerantz, J.R. (1983) Global and local precedence: Selective attention in form and motion perception. *Journal of Experimental Psychology: General,* **122,** 516–540.

Posner M.I. (1980) Orienting of attention. *Quarterly Journal of Experimental Psychology,* **32,** 3–25.

Posner, M.I. & Cohen, Y. (1984) Components of visual orienting. In: H. Bouma & D.G. Bouwhuis (eds.), *Attention and Performance X.* Erlbaum, Hillsdale, NJ.

Posner, M.I. & Snyder, C.R.R. (1975) Facilitation and inhibition in the processing of signals. In: P.M.A. Rabbitt & S. Dornic (eds.), *Attention and Performance V.* Academic Press, New York, pp. 669–682.

Posner, M.I., Nissen, M.J. & Klein, R. (1976) Visual dominance: an information-processing account of its origin and significance. *Psychological Review,* **83,** 157–171.

Potter, M.C. (1983) Representational buffers: The eye-mind hypothesis in picture perception, reading, and visual search. In: K. Rayner (ed.), *Eye Movements in Reading: Perceptual and Language Processes.* Academic Press, New York.

Prinz, W. (1983) Asymmetrical control araes in continuous visual search. In: R. Groner, C. Menz, D.F. Fisher & R.A. Monty (eds.), *Eye Movements and Psychological Functions: International Views.* Erlbaum, Hillsdale, NJ, pp. 85–100.

Prinz, W., Meinecke, C. & Hielschel, M. (1987) Effects of stimulus degradation on category search. *Acta Psychologica,* **64,** 187–206.

Pyke, R.K. (1926) *Report on the Legibility of Print.* H.M.S.O., London.

Radl, G.W. (1980) Experimental investigations for optimal presentation mode and colours of symbols on the CRT screen. In: E. Grandjean & E. Vigliani (eds.), *Ergonomic Aspects of Visual Display Terminals.* Taylor & Francis, London.

Rapoport, A. & Chammah, A.M. (1965) Sex differences in factors contributing to the level of cooperation in the prisoner's dilemma game. *Journal of Personality and Social Psychology,* **2,** 831–838.

Ratcliff, R. (1978) A theory of memory retrieval. *Psychological Review,* **85,** 59–108.

Rayner, K. (1979) Eye movements in reading: Eye guidance and integration. In: P.A. Kolers, M.E. Wrolstad & H. Bouma (eds.), *Processing of Visible Language* (Vol. 1). Plenum Press, New York.

Rayner, K. (1983) *Eye Movements in Reading: Perceptual and Language Processes.* Academic Press, London.

Rayner, K. & Fisher, D.C. (1987) Eye movements and the perceptual span during visual search. In: K. O'Regan & A. Levy-Schoen (eds.), *Eye Movements: From Physiology to Cognition.* Elsevier, Amsterdam, pp. 293–302.

Rayner, K., Well, A.D., Pollatsek, A. & Bertera, J.H. (1982) The availability of useful information to the right of fixation in reading. *Perception and Psychophysics,* **31,** 537–550.

Remington, R.W. (1980) Attention and saccadic eye movements. *Journal of Experimental Psychology,* **6,** 726–744.

Riggs, L.A. (1965) Visual acuity. In: C.H. Graham (ed.), *Vision and Visual Perception.* Wiley, New York, pp. 321–349.

Rizzolatti, G., Riggio, L., Dascola, I. & Umiltá, C. (1987) Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia,* **25,** 31–40.

Roberts, T. (1984) Perspectives of a modern user–interface designer. In: G. Salvendy (ed.), *Human–Computer Interaction.* Elsevier, Amsterdam, pp. 61–66.

Robinson, D.A. (1963) A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Transactions in Biomedics and Electronics,* **BME-10,** 137–145.

Rogers, Y. (1989*a*) Icon design for the user interface. *International Review of Ergonomics,* **2,** 129–154.

Rogers, Y. (1989*b*) Icons at the interface: their usefulness. *Interacting with Computers,* **1,** 12–24.

Rogers, Y. & Oborne, D.J. (1987) Pictorial communication of abstract verbs in relation to human computer interaction. *British Journal of Psychology,* **78,** 99–112.

Rosch, E., Simpson, C., & Miller, R.S. (1976) Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance,* **2,** 491–502.

Rosner, M. & Belkin, M. (1989) Video display units and visual function. *Survey of Ophthalmology,* **33,** 515–522.

Rovamo, J. & Virsu, V. (1979) An estimation and application of the human cortical magnification factor. *Experimental Brain Research,* **37,** 495–510.

Russo, J.E. (1978) Adaptation of cognitive processes to the eye movement systems. In: J.W. Senders, D.F. Fisher & R.A Monty (eds.), *Eye Movements and the Higher Psychological Functions.* Erlbaum, Hillsdale, NJ, pp. 89–109.

Saariluoma, P. (1990) Chess players' search for task relevant cues: are chunks relevant? In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 115–121.

Sagi, D. & Julesz, B. (1985) 'Where' and 'what' in vision. *Science,* **228,** 1217–1219.

Sanders, A.F. (1963) *The Selective Process in the Functional Visual Field.* From Inst. for Perception RVO-TNO. Nat. Def. Res. Organization TNO. Soesterberg, Netherlands.

Schmandt, C. (1980) Soft typography. *Information Processing, 1980: Proceedings of the IFIPS.* pp. 1027–1032.

Schmandt, C. (1983) Greyscale fonts designed from video signal analysis. *Proceedings NCGA, 1983.* pp. 549–558.

Schneider, W. & Shiffrin, R.M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review,* **84,** 1–66.

Schoonard, J.W., Gould, J.D. & Miller, L.A. (1973) Studies of visual inspection. *Ergonomics,* **16,** 365–379.

Schwartz, J.P. & Norman, K.L. (1986) The importance of item distinctiveness on performance using a menu selection system. *Behaviour and Information Technology,* **5,** 173–182.

Scott, D. & Findlay, J.M. (1990) The shape of VDUs to come: a visual search study. In: E.J. Lovesey (ed.), *Contemporary Ergonomics 1990.* Taylor & Francis, London, pp. 353–358.

Scott, D. & Findlay, J.M. (1991*a*) Visual search and VDUs. In: D. Brogan (ed.), *Visual Search II.* Taylor & Francis, London.

Scott, D. & Findlay, J.M. (1991*b*) *Visual Search, Eye Movements and Display Units.* Technical Rep. HF147. IBM (UK) Laboratories, Hursley.

Scott, D. & Findlay, J.M. (1991*c*) Optimum display arrangements for presenting status information. *International Journal of Man–Machine Studies* (in press).

Scott, D. & Findlay, J.M. (1991*d*) Future attractions: a visual search study of computer icons and words. In: E.J. Lovesey (ed.), *Contemporary Ergonomics.* Taylor & Francis, London.

Shepherd, M., Findlay, J.M. & Hockey, G.R.J. (1986) The relationship between eye movements and spatial attention. *Quarterly Journal of Experimental Psychology,* **38A,** 475–491.

Sekuler, R. & Blake, R. (1985) *Perception.* Knopf, New York.

Shackel, B. (1967) Eye movement recording by electro-oculography. In: P.H. Venables & I. Martin (eds.), *A Manual of Psychophysiological Methods.* Elsevier, Amsterdam.

Shannon, C.E. (1949) The mathematical theory of communication. *Bell System Technical Journal,* **27,** 379–423.

Shapiro, K.L. (1990) Attention in visual space affects oculomotor performance. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 169–176.

Shapiro, K.L. & Raymond, J.E. (1969) Training for efficient oculomotor strategies enhances skill acquisition. *Acta Psychologica.* Cited in Shapiro (1990).

Shapiro, K.L., Egerman, B. & Klein, R.M. (1984) Effects of arousal on human visual dominance. *Perception and Psychophysics,* **35,** 547–552.

Shaw, M.L. (1978) A capacity allocation model for reaction time. *Journal of Experimental Psychology: Human Perception and Performance,* **4,** 586-598.

Shaw, M.L. & Shaw, P. (1977) Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception and Performance,* **3,** 201-211.

Shiffrin, R.M. & Schneider, W. (1977) Controlled and automatic human information processing II: Perceptual learning, automatic attending, and a general theory. *Psychological Review,* **84,** 127–190.

Shneiderman, B. (1983) Direct manipulation: a step beyond programming languages. *IEEE Computer,* **16,** 57–69.

Shneiderman, B. (1987) *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* Addison-Wesley, Reading, MA.

Sholtz, P.N. (1982) *Making high-quality colored images on raster displays.* IBM Research Report RC-9632. Yorktown Heights, NY.

Silverstein, L.D., Monty, R.W., Gomer, F.E. & Yeh, Y.Y. (1989) A psychophysical evaluation of pixel mosaics and gray-scale requirements for color matrix displays. In: *Society for Information Display International Symposium Digest of Technical Papers, Vol. XX.* SID, Baltimore, MD, pp. 128-131.

Slater, A. M. and Findlay, J. M. (1975). Binocular fixation of the newborn baby. *Journal of Experimental Child Psychology,* **20,** 248–273.

Smith, S.L. & Thomas, D.W. (1964) Colour versus shape coding in information displays. *Journal of Applied Psychology,* **48,** 137–146.

Snowberry, K., Parkinson, S.R. & Smith, S. (1983) Computer display menus. *Ergonomics,* **26,** 699–712.

Snyder, H.L. & Taylor, D.F. (1976) *Compterized Analysis of Eye Movements During Static Display Visual Search.* Aerospace Medical Research Laboratory Report No. AMRL-TR-79-91, Ohio.

Smith, V.C. & Pokorny, J. (1975) Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research,* **15,** 161–171.

Stammers, R.B., George, D.A. & Carey, M.S. (1989) An evaluation of abstract and concrete icons for a CAD package. In: E.D. Megaw (ed.), *Contemporary Ergonomics 1989.* Taylor & Francis, London, pp. 416–421.

Starch, D. (1923) *Principles of Advertising.* Chapter XXV. Layout and typography. A.W. Shaw, Chicago.

Stark, L. & Ellis, S.R. (1981) Scan paths revisited: Cognitive models direct active looking. In: D.F. Fisher, R.A. Monty & J.W. Senders (eds.), *Eye Movements: Cognition and Visual Perception.* Erlbaum, Hillsdale, NJ.

Starkes, J.L. (1987) Skill in field hockey: The nature of the cognitive advantage. *Journal of Sport Psychology,* **9,** 146–160.

Steinman, S.B. (1987) Serial and parallel search in pattern vision. *Perception,* **16,** 389–398.

Sternberg, S. (1967) Two operations in character recognition: Some evidence from reaction-time measurement. *Perception and Psychophysics,* **2,** 45–53.

Stone, L.J. (1990) *Windows of the future: A visual search comparison of computer icons and words.* Undergraduate dissertation. Dept. of Psychology, University of Durham.

Switchenko, D.M. (1984) Reading from CRT versus paper: the CRT disadvantage re-examined. In: *Proceedings of the 25th Annual Meeting of the Human Factors Society.* HFS, Baltimore.

Tassinari, G., Aglioti, S., Chelazzi, L., Marzi, C.A. & Berlucchi, G. (1987) Distribution in the visual field of the costs of voluntarily allocated attention and of the inhibitory after-effects of covert orienting. *Neuropsychologia,* **25,** 55–71.

Timmers, H., van Nes, F.L. & Blommaert, F.J.J. (1980) Visual word recognition as a function of contrast. In: E. Grandjean & E. Vigliani (eds.), *Ergonomic Aspects of Visual Display Terminals.* Taylor & Francis, London.

Tinker, M.A. (1963) *Legibility of Print.* Ames, Iowa State University.

Townsend, J.T. & Ashby, F.G. (1983) *Stochastic Modelling of Elementary Psychological Processes.* CUP.

Travis, D.S. (1990) Applying visual psychophysics to user interface design. *Behaviour & Information Technology,* **9,** 425–438.

Treisman, A.M. (1960) Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology,* **12,** 242–248.

Treisman, A.M. (1964) Verbal cues, language, and meaning in selective attention. *American Journal of Psychology,* **77,** 206–209.

Treisman, A. (1985) Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing,* **31,** 156–177.

Treisman, A.M. (1986) Features and objects in visual processing. *Scientific American,* **255,** 114–126.

Treisman, A.M. (1988) Features and objects: The fourteenth Bartlett Memorial lecture. *Quarterly Journal of Experimental Psychology,* **40A,** 201–237.

Treisman, A.M. & Gelade, G. (1980) A feature-integration theory of attention. *Cognitive Psychology,* **12,** 97–136.

Treisman, A.M. & Gormican, S. (1988) Feature analysis in early vision: Evidence from visual search asymmetries. *Psychological Review,* **95,** 15–48.

Treisman, A.M. & Paterson, R. (1984) Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception and Performance,* **10,** 12-31.

Treisman, A. & Southern, J. (1985) Search asymmetry: a diagnostic for preattentive processing of seperable features. *Journal of Experimental Psychology: General,* **114,** 285–310.

Tsal, Y. (1983) Movements of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance,* **9,** 523–530.

Tullis, T.S. (1983) The formatting of alphanumeric displays: a review and analysis. *Human Factors,* **25,** 657–682.

Turkowski, K. (1986) Anti-aliasing in topological colour spaces. *Computer Graphics,* **20,** 307–314.

Visick, P., Johnson, P. & Long, J. (1984) A comparative analysis of keyboards and voice recognition in a parcel sorting task. *Proceedings of the Ergonomics Society Annual Conference.* Taylor & Francis, London.

Unema, P. & Rötting M. (1990) Differences in eye movements and mental workload between experienced and inexperienced motor-vehicle drivers. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 193–202.

Umiltà, C. (1988) Orienting of attention. In: F. Boller & J. Grafman (eds.), *Handbook of Neuropsychology.* Elsevier, Amsterdam, pp. 175–193.

van Nes, F.L. & Bouman, M.A. (1967) Spatial modulation transfer in the human eye. *Journal of the Optical Society of America,* **57,** 401–407.

Vaughan, G.M. & Corballis, M.C. (1969) Beyond tests of significance: Estimating strength of effects in selected ANOVA designs. *Psychological Bulletin,* **72,** 204–213.

Vitu, F. & O'Regan, J.K. (1990) Eye movements in reading: optimal landing position and usefulness of peripheral vision. In: D. Brogan (ed.), *Visual Search.* Taylor & Francis, London, pp. 133–140.

Viviani, P. (1990) Eye movements in visual search: cognitive, perceptual and motor control aspects. In: E. Kowler (ed.), *Eye Movements and their Role in Visual and Cognitive Processes.* Elsevier, Amsterdam, pp. 353–393.

Voss, M. (1981) Narrowing of the visual field as indicator of mental work load? In: J. Moraal & K.F. Kraiss (eds.), *Manned System Design: Methods, Equipment, and Applications.* Plenum Press, New York.

Wandmacher, J. & Müller, U. (1987) On the usability of verbal and iconic command representations. *Zeitschrift für Psychologie.* (cited in Arend *et al.*, 1987).

Ward, L.M. (1983) On processing dominance: Comment on Pomerantz. *Journal of*

Experimental Psychology: General, **112,** 541–546.

Ward, R. & McClelland, J.L. (1989) Conjunctive search for one and two identical targets. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 664–672.

Warnock, J. (1980) The display characters using gray level sample arrays. *Proceedings SIGGRAPH, 1980.* pp. 302–307.

Weiman, N. & Perrin, N. (1989) Effects of flat-panel pixel structures and anti-aliasing techniques across three image types upon image quality judgements. In: *Society for Information Display International Symposium Digest of Technical Papers, Vol. XX.* SID, Baltimore, MD, pp. 132–134.

Weisstein, N. & Harris, C.S. (1974) Visual detection of line segments: An object-superiority effect. *Science,* **186,** 752–755.

Welford, A.T. (1968) *Fundamentals of skill.* Methuen, London.

Wertheimer, M. (1923) Untersuchungen zur Lehre von der Gestalt. II. *Psychologischen Forschung,* **4,** 301–350.

Westheimer, G. (1979) The spatial senses of the eye. *Investigative Ophthalmology and Visual Science,* **20,** 893–912.

Whitefield, A. (1989) Constructing appropriate models of computer users: the case of engineering designers. In: J. Long & A. Whitefield (eds.), *Cognitive Ergonomics and Human-Computer Interaction.* CUP, pp. 66–94.

Whiteside, J., Jones, S., Levy, P.S. & Wixon, D. (1985) User performance with command, menu and iconic interfaces. In: *Proceedings of Human Factors in Computing.* ACM, San Francisco, pp. 185–191.

Widdel, H. (1983) A method for measuring the visual lobe area. In: R. Groner, C. Menz, D.F. Fisher & R.A. Monty (eds.), *Eye Movements and Psychological Functions: International Views.* Erlbaum, Hillsdale, NJ, pp. 73–84.

Williams, L.G. (1966) Target conspicuity and visual search. *Human Factors,* **8,** 80–92.

Wilson, H.R., Levi, D., Maffei, L., Rovamo, J. & Devalois, R. (1990) The perception of form: Retina to striate cortex. In: L. Spillmann & J.S. Werner (eds.), *Visual Perception: The Neurophysiological Foundations.* Academic Press, San Diego, pp 231–272.

Wright, P. & Lickorish, A. (1983) Proof-reading text on screen and paper. *Behaviour and Information Technology,* **2,** 227–235.

Wright, P. & Lickorish, A. (1984) Ease of annotation in proof-reading tasks. *Behaviour and Information Technology,* **3,** 185–194.

Wyszecki, G. & Stiles, W.S. (1982) *Colour Science* (2nd edn). Wiley, New York.

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

Young, L.R. & Sheena, D. (1975). Survey of eye movement recording methods. *Behaviour Research Methods, Instrumentation and Computers* **7**, 397–429.

APPENDIX 3.1:
INSTRUCTIONS TO SUBJECTS INVOLVED IN EXPT. 1

You will be shown a total of 100 slides. Each contains a random display of letters of the alphabet. You will find that the length and width of these arrays vary.

Your task is simply to locate the target letter X. The letter to search for will always be X. There is only one letter X in each slide, and in each slide there is only one letter X.

As soon as each slide appears, you are to search, using whatever manner you wish, for the letter X. Immediately you have spotted it, press the key. Do not press the key until you are sure you have located the X. Then tell me what letter is to left and right of the target (for instance B, X, blank). Another slide will then appear.

If you press mistakenly (before really finding the X), just say so and continue. Each search is timed.

You will be told more about the study when you have finished.

Do you have any questions?

APPENDIX 3.2a: INDIVIDUAL SUBJECTS' DATA FROM EXPT. 2
READING STRATEGY

Height by width	Fixations (ms)	Saccade amp. (char. spaces)	Regression amp. (char. spaces)
25x100			
MS			
Mean (sd)	340 (153)	5.303 (2.218)	3.714 (1.704)
Med. Mode	300 220	5.000 5.000	3.000 3.000
AH			
Mean (sd)	274 (77)	13.345 (7.578)	4.800 (1.751)
Med. Mode	260 220	12.000 9.000	5.000 5.000
JF			
Mean (sd)	295 (100)	12.026 (6.827)	4.778 (2.669)
Med. Mode	280 240	11.000 9.000	4.000 3.000
30x83			
MS			
Mean (sd)	302 (125)	6.528 (2.799)	3.364 (1.329)
Med. Mode	280 260	6.000 5.000	3.000 3.000
AH			
Mean (sd)	262 (67)	12.398 (6.579)	4.900 (2.767)
Med. Mode	260 240	11.000 7.000	4.000 9.000
JF			
Mean (sd)	280 (112)	10.546 (7.160)	4.000 (2.681)
Med. Mode	280 140	10.000 5.000	3.500 2.000
50x50			
MS			
Mean (sd)	291 (85)	8.475 (4.625)	3.250 (2.062)
Med. Mode	280 220	7.000 6.000	3.500 5.000
AH			
Mean (sd)			
Med. Mode			
JF			
Mean (sd)	271 (85)	8.700 (5.739)	5.500 (2.276)
Med. Mode	240 140	7.000 (6.000)	6.000 8.000
83x30			
MS			
Mean (sd)	281 (84)	7.311 (2.767)	3.750 (2.217)
Med. Mode	260 220	7.000 7.000	3.000 3.000
AH			
Mean (sd)	274 (85)	7.927 (4.175)	6.333 (2.887)
Med. Mode	260 260	7.000 7.000	8.000 8.000
JF			
Mean (sd)	275 (118)	9.636 (6.489)	5.071 (2.165)
Med. Mode	240 240	8.000 6.000	5.500 6.000
100x25			
MS			
Mean (sd)	284 (94)	5.771 (2.200)	3.750 (0.957)
Med. Mode	260 240	6.000 5.000	3.500 3.000
AH			
Mean (sd)	292 (97)	6.843 (3.984)	6.750 (2.017)
Med. Mode	280 280	6.000 4.000	6.500 6.000
JF			
Mean (sd)	313 (146)	6.711 (4.015)	5.156 (2.567)
Med. Mode	300 140	6.000 6.000	5.000 6.000
COMBINED			
25x100			
Mean	303	10.285	4.431
Med. Mode	280 227	9.333 7.667	4.000 3.667
30x83			
Mean	281	9.824	4.088
Med. Mode	273 213	9.000 5.667	3.500 4.667
50x50			
Mean	281	8.587	4.375
Med. Mode	260 180	7.000 6.000	4.750 7.500
83x30			
Mean	277	8.291	5.051
Med. Mode	253 240	7.333 6.667	5.500 5.667
100x25			
Mean	296	6.442	5.219
Med. Mode	280 220	6.000 4.000	5.000 5.000

APPENDIX 3.2b: INDIVIDUAL SUBJECTS' DATA FROM EXPT. 2
FREE SEARCH

Height by width	Fixations (ms)	Saccade amp. (char. spaces)	Regression amp. (char. spaces)
25x100			
MS			
Mean (sd)	381 (174)	6.611 (4.191)	6.250 (2.765)
Med. Mode	240 280	6.000 3.000	5.500 10.000
AH			
Mean (sd)	278 (79)	14.632 (11.857)	4.333 (1.155)
Med. Mode	280 280	12.000 6.000	5.000 5.000
4 JF			
Mean (sd)	244 (80)	16.750(11.127)	2.500 0.707
Med. Mode	240 140	18.000 3.000	2.500 3.000
30x83			
MS			
Mean (sd)	262 (86)	5.000 (3.391)	3.000 (0.000)
Med. Mode	240 220	5.000 1.000	3.000 3.000
AH			
Mean (sd)	302 (84)	10.667 (6.952)	4.600 (2.881)
Med. Mode	280 280	9.000 6.000	4.000 8.000
JF			
Mean (sd)	276 (110)	11.409 (7.918)	5.158 (2.455)
Med. Mode	260 140	10.000 6.000	5.000 4.000
50x50			
MS			
Mean (sd)	314 (126)	7.893 (4.581)	5.189 (4.149)
Med. Mode	280 240	7.000 5.000	6.000 7.000
AH			
Mean (sd)	292 (90)	7.982 (6.838)	6.214 (2.517)
Med. Mode	300 320	6.000 3.000	5.500 5.000
JF			
Mean (sd)	248 (98)	11.574 (7.084)	4.667 (2.646)
Med. Mode	220 140	10.000 8.000	5.000 8.000
83x30			
MS			
Mean (sd)	300 (108)	6.815 (4.059)	5.500 (2.719)
Med. Mode	280 200	6.000 4.000	5.000 3.000
AH			
Mean (sd)	294 (90)	4.368 (2.910)	6.400 (0.548)
Med. Mode	280 260	4.000 2.000	6.000 6.000
JF			
Mean (sd)	238 (94)	8.684 (6.263)	5.500 (2.777)
Med. Mode	220 140	7.000 2.000	7.000 2.000
100x25			
MS			
Mean (sd)	328 (144)	5.597 (3.151)	5.188 (2.509)
Med. Mode	320 320	5.000 5.000	5.500 2.000
AH			
Mean (sd)	304 (134)	4.000 (2.132)	3.857 (2.410)
Med. Mode	300 140	4.000 1.000	3.000 3.000
JF			
Mean (sd)	322 (114)	7.808 (4.691)	8.500 (0.707)
Med. Mode	320 260	6.500 5.000	8.500 9.000
COMBINED			
25x100			
Mean	317	12.664	4.361
Med. Mode	287 233	12.000 4.000	4.333 6.000
30x83			
Mean	280	9.025	4.253
Med. Mode	260 213	8.000 4.333	4.000 5.000
50x50			
Mean	285	9.150	5.357
Med. Mode	267 233	7.667 5.333	5.500 6.667
83x30			
Mean	277	5.802	5.800
Med. Mode	260 200	5.667 2.667	6.000 3.667
100x25			
Mean	318	5.802	5.848
Med. Mode	313 240	5.167 3.667	5.833 4.667

APPENDIX 4.1a: EXAMPLE (ONE PAGE FROM SEVEN) OF ONE SERIES OF LETTERS USED IN THE TEXT-EDITING TASK (UNDEGRADED, (HIGHLIGHTED HARDCOPY VERSION) FROM EXPERIMENT 4.

139 Elm Park Mansions,
Park Walk,
London SW10

Ms Esther Rantzen,
BBC Television,
London W12

Dear Esther,

Congratulations on another great show last night! The main 'how to get your knickers off' item was hilarious! Both Mrs Root and I chuckled all the way through it.

You've got the formula just right. If only more people in TV (television) realised that it's possible to be healthily vulgar without descending to schoolboy smut.

I enclose some humourous definitions and a comic poem for Cyril. I hope you can use them. On the assumption you can, I'm invoicing the BBC separately.

Just one slight criticism of the show last night. I thought your dress was rather revealing for what is essentially family viewing. One could see your legs quite clearly. I hope you won't mind me saying this. One doesn't want to see women's legs in one's lounge-room at a time when youngsters are still up and about.

Could you possibly oblige with a photo?

Yours for A Comical Definition!

Henry Root.

APPENDIX 4.1b: EXAMPLE (ONE PAGE FROM SEVEN) OF ONE SERIES OF LETTERS USED IN THE TEXT-EDITING TASK (DEGRADED OR SCREEN-DISPLAYED VERSION) FROM EXPERIMENT 4.

139 Elm Park Mansions,
Park Walk,
London SW10

Ms Esther Rantzen,
BB Television,
London W12

Dear Esther,

Congratulations on another show last night! The main
'how to get your knickers off' item was hilarious. Mrs
Root and I chuckled all the way through it.

I enclose some humourous definitions and a comic poem for
Cyril.

Aust one slight criticism of the show last night. I thought
your dress was rather revealing for what is essentially
family viewing. One could see your legs quite clearly. I hope
you won't mind me. One doesn't want to see
women's legs in one's lounge-room at a time whe younsters
are still up and about.

Could you possibly oblige with a photo?

Youbs for A Comical Definition!

Henry Root

APPENDIX 4.2: INSTRUCTIONS TO SUBJECTS INVOLVED IN EXPERIMENT 4

This is a word processing exercise. One aspect which we are interested in is how people use insert and type-over modes.

If you are familiar with the Durham University text editing system you will notice some similarities and some differences with the present system. For instance, some of the standard commands are not available here.

I would like to point out that this is not a test of your personal ability. Rather, it is a way of looking at how people in general learn new tasks.

Your task is simply to insert within these letters those characters or words which are marked in yellow using the following commands:-

CTRL-A for type-over
CTRL-S for character insertion
CTRL-D for forward character deletion
(i.e. deletes the character in front of the cursor)
CTRL-F for line deletion

Where yellow sections are involved you will need to be in insert mode.

Where letters are marked in orange the wrong letter has been typed. You are advised, when correcting single letter errors, to type over the mistake, although it is possible to do this by deletion and insertion.

As well as the mode being displayed on the top of the screen, to help you know which mode you are in the cursor can give the same information:-

■ cursor indicates over-write
⌶ cursor indicates character insertion

Please try to use this information as much as possible.

To scroll the page on use:-

F1 forward 10 lines
F2 forward 20 lines
F3 back 10 lines
F4 back 20 lines

APPENDIX 5.1: INSTRUCTIONS TO SUBJECTS INVOLVED IN EXPERIMENT 6

The study involves using a text editor to correct spelling mistakes and omissions. As someone who has used the Curlew editor before, you will find that there both similarities and differences with the present system. Here is a practice screen.

There are four seperate pieces of text to edit. You will find slight differences between screens across the four sessions. You will find that the commands for the most important functions are printed at the top of the keyboard. For instance control-A (two keystrokes) puts you into type-over mode. Control-S puts you into insert mode. These [indicate] are the important keys.

Your task is simply to correct the text on the right hand side of this split screen so that it is identical to the 'correct' version on the left hand side of the screen. You will find it best to move the cursor with the keyboard arrow keys. Please try not to use the return key as this, unfortunately, puts the cursor into the left hand side of the screen. If this happens do not worry about it—just move the cursor over to the right and carry on. You can practice now if you wish.

Where there is only one character to be corrected (e.g. an 'a' printed for an 'e', then you will need to be in 'type-over' mode. Where there are many characters or words to insert, make sure that you are in insert mode. You will find that you will have to make frequent changes between insert and type-over modes. It is important that you are in the appropriate mode before typing.

This other equipment [indicate] enables us to analyse your performance at a later time. Is that clear so far?

(a) On the next session the words 'insertion' or 'type-over' are present at the top of the screen to let you know what mode you are in. Start as soon as the screen appears and please let me know when you have finished.

(b) On the next session the words 'insertion' or 'type-over' are present at the top of the screen to let you know what mode you are in. Start as soon as the screen appears and please let me know when you have finished.

(c) On the next session no information will be presented when you are in insert mode. When you are in type-over mode, however, you will see a short row of yellow asterisks at the top of the screen. Start as soon as the screen appears and please let me know when you have finished.

(d) On the next session, no information will be presented when you are in insert mode. When you are in type-over mode, however, you will see a complete row of yellow asterisks at the top and bottom of the screen. Start as soon as the screen appears and please let me know when you have finished.

APPENDIX 6.1: INDIVIDUAL SUBJECTS' DATA FROM EXPT. 9

Height by width	Fixations (ms)	Saccade amp. (char. spaces)	Regression amp. (char. spaces)
25x100			
MS			
Mean (sd)	340 (153)	5.303 (2.218)	3.714 (1.704)
Med. Mode	300 220	5.000 5.000	3.000 3.000
AH			
Mean (sd)	274 (77)	13.345 (7.578)	4.800 (1.751)
Med. Mode	260 220	12.000 9.000	5.000 5.000
JF			
Mean (sd)	295 (100)	12.026 (6.827)	4.778 (2.669)
Med. Mode	280 240	11.000 9.000	4.000 3.000
30x83			
MS			
Mean (sd)	302 (125)	6.528 (2.799)	3.364 (1.329)
Med. Mode	280 260	6.000 5.000	3.000 3.000
AH			
Mean (sd)	262 (67)	12.398 (6.579)	4.900 (2.767)
Med. Mode	260 240	11.000 7.000	4.000 9.000
JF			
Mean (sd)	280 (112)	10.546 (7.160)	4.000 (2.681)
Med. Mode	280 140	10.000 5.000	3.500 2.000
50x50			
MS			
Mean (sd)	291 (85)	8.475 (4.625)	3.250 (2.062)
Med. Mode	280 220	7.000 6.000	3.500 5.000
AH			
Mean (sd)			
Med. Mode			
JF			
Mean (sd)	271 (85)	8.700 (5.739)	5.500 (2.276)
Med. Mode	240 140	7.000 (6.000)	6.000 8.000
83x30			
MS			
Mean (sd)	281 (84)	7.311 (2.767)	3.750 (2.217)
Med. Mode	260 220	7.000 7.000	3.000 3.000
AH			
Mean (sd)	274 (85)	7.927 (4.175)	6.333 (2.887)
Med. Mode	260 260	7.000 7.000	8.000 8.000
JF			
Mean (sd)	275 (118)	9.636 (6.489)	5.071 (2.165)
Med. Mode	240 240	8.000 6.000	5.500 6.000
100x25			
MS			
Mean (sd)	284 (94)	5.771 (2.200)	3.750 (0.957)
Med. Mode	260 240	6.000 5.000	3.500 3.000
AH			
Mean (sd)	292 (97)	6.843 (3.984)	6.750 (2.017)
Med. Mode	280 280	6.000 4.000	6.500 6.000
JF			
Mean (sd)	313 (146)	6.711 (4.015)	5.156 (2.567)
Med. Mode	300 140	6.000 6.000	5.000 6.000
COMBINED			
25x100			
Mean	303	10.285	4.431
Med. Mode	280 227	9.333 7.667	4.000 3.667
30x83			
Mean	281	9.824	4.088
Med. Mode	273 213	9.000 5.667	3.500 4.667
50x50			
Mean	281	8.587	4.375
Med. Mode	260 180	7.000 6.000	4.750 7.500
83x30			
Mean	277	8.291	5.051
Med. Mode	253 240	7.333 6.667	5.500 5.667
100x25			
Mean	296	6.442	5.219
Med. Mode	280 220	6.000 4.000	5.000 5.000

APPENDIX 7.1: DETAILS OF THE SYSTEMATIC ORGANISATION OF NON-TARGETS IN HIGH AND LOW SIMILARITY CONDITIONS (EXPT. 10)

HIGH

As non-targets:
a target of 1 cycle would have 1.5 (2), 2 (2), 2.5, 3, 3.5 cycles;
a target of 1.5 cycles would have 1 (2), 2 (2), 2.5, 3, 3.5 cycles;
a target of 2 cycles would have 1, 1.5 (2), 2.5 (2), 3, 3.5 cycles;
a target of 2.5 cycles would have 1.5, 2 (2), 3 (2), 3.5, 4 cycles;
a target of 3 cycles would have 2, 2.5 (2), 3.5 (2), 4, 4.5 cycles;
a target of 3.5 cycles would have 2.5, 3 (2), 4 (2), 4.5, 5 cycles;
a target of 4 cycles would have 2.5, 3, 3.5 (2), 4.5 (2), 5 cycles;
a target of 4.5 cycles would have 2.5, 3, 3.5, 4 (2), 5 (2) cycles;
a target of 5 cycles would have 2.5, 3, 3.5, 4 (2), 4.5 (2) cycles.

LOW

As non-targets:
a target of 1 cycle would have 3, 3.5, 4, 4.5 (2), 5 (2) cycles;
a target of 1.5 cycles would have 3, 3.5, 4, 4.5 (2), 5 (2) cycles;
a target of 2 cycles would have 3, 3.5, 4, 4.5 (2), 5 (2) cycles;
a target of 2.5 cycles would have 1, 3.5, 4, 4.5 (2), 5 (2) cycles;
a target of 3 cycles would have 1 (2), 1.5, 2, 4.5, 5 (2) cycles;
a target of 3.5 cycles would have 1 (2), 1.5 (2), 2, 2.5, 5 cycles;
a target of 4 cycles would have 1 (2), 1.5 (2), 2, 2.5, 3 cycles;
a target of 4.5 cycles would have 1 (2), 1.5 (2), 2, 2.5, 3 cycles;
a target of 5 cycles would have 1 (2), 1.5 (2), 2, 2.5, 3 cycles.

APPENDIX 7.2: INSTRUCTIONS TO SUBJECTS INVOLVED IN EXPERIMENT 10

This study concerns search times and various patterns.
You will be shown a total of 144 different slides.
Within each one you will see (demo) one central pattern and eight patterns arranged around it. Your task is simply to locate the pattern in the surrounding ring which is identical to that in the centre. That is the target pattern. There is only one correct target within those eight surrounding patterns, although some are similar.
Each time, the projector will display slides like this one and immediately they appear you are to find the target. As soon as you have done so just press the key in front of you. This will stop the automatic timer which has been started when the slide comes up. To let me know that you have found the right target, I just want you to say each time roughly where the target is. For instance, 'north-west', or '10/11 o'clock'; just as you prefer. I will then put up the next slide. Please try to keep your head the same distance from the screen throughout. There will be a break about half way through.
Any questions?

APPENDIX 8.1:
INSTRUCTIONS TO SUBJECTS INVOLVED IN EXPERIMENT 12

This study aims to investigate legibility of various fonts.

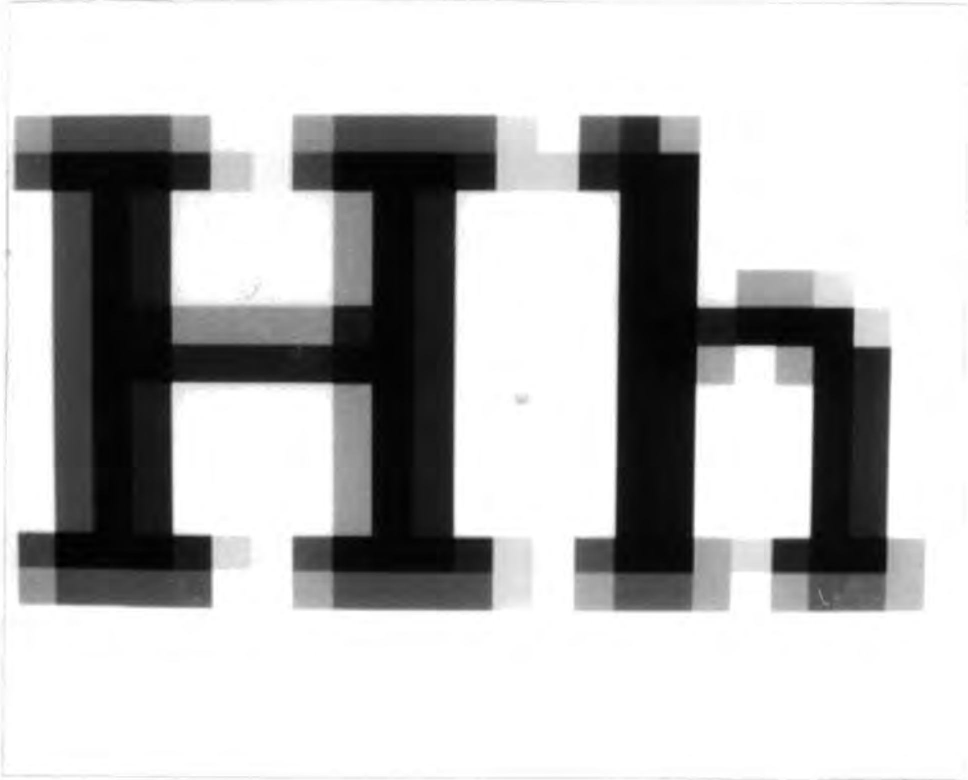
You will be shown a total of 54 screens of randomly arrayed alphabetic characters. Your task is simply to locate the target letter. The target letter will be revealed each time to you from this pack immediately before the next screen appears. It will vary. When you have found the target just tell me what letters are immediately to the left and right of it. For instance, you might say 'B X C' where X was the target letter, or 'P O edge' where O was a target. I will then press the key to bring up the next screen and you may turn over the next target card. Look at that and immediately commence searching. As such, there is little delay between finding one target and the next screen appearing, but there will be a break between each dozen or so screens, whilst the computer reloads.

Please try to maintain your normal VDU viewing distance.

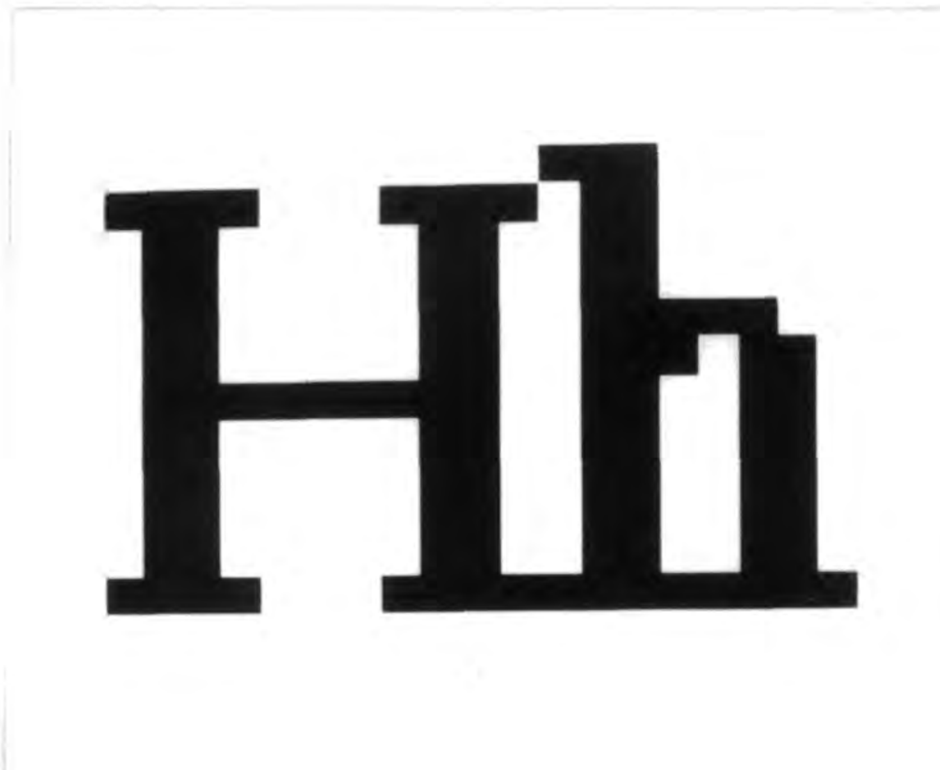
Is that clear?

APPENDIX 8.2a:
AN EXAMPLE OF THE FONTS USED IN EXPERIMENT 13
(letter characterised by horizontal/vertical strokes; 12 point binary)

APPENDIX 8.2b:
AN EXAMPLE OF THE FONTS USED IN EXPERIMENT 13
(letter characterised by horizontal/vertical strokes; 12 point AA)



APPENDIX 8.2c:
AN EXAMPLE OF THE FONTS USED IN EXPERIMENT 13
(letter characterised by horizontal/vertical strokes; 12 point bold)

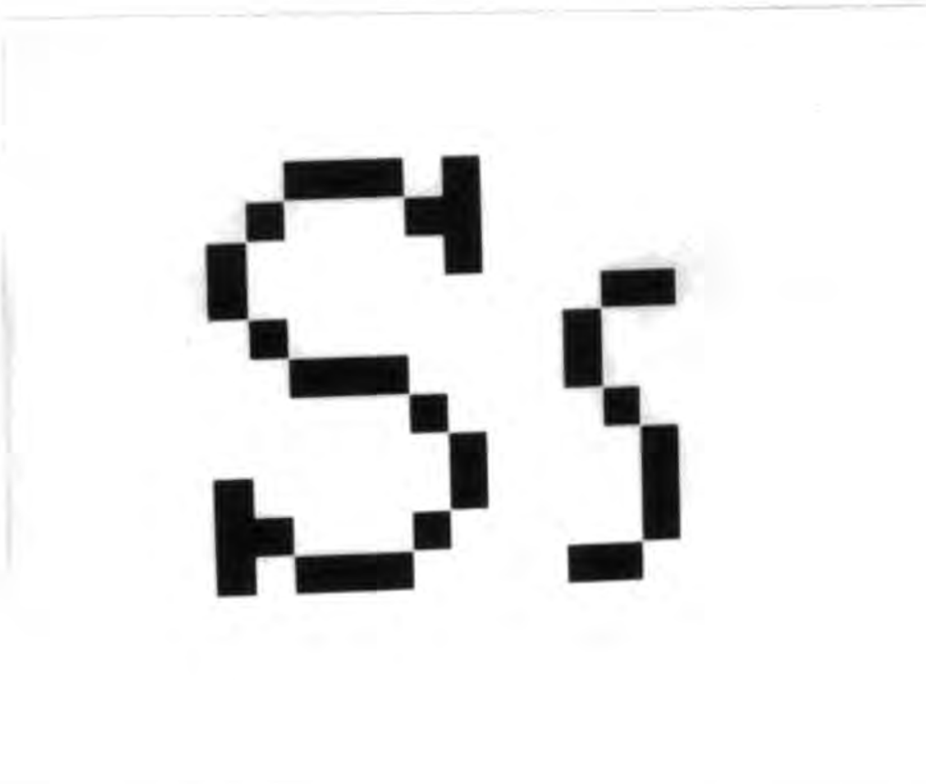


APPENDIX 8.2d: AN EXAMPLE OF THE FONTS USED IN EXPERIMENT 13

(letter characterised by horizontal/vertical strokes; 14 point binary)

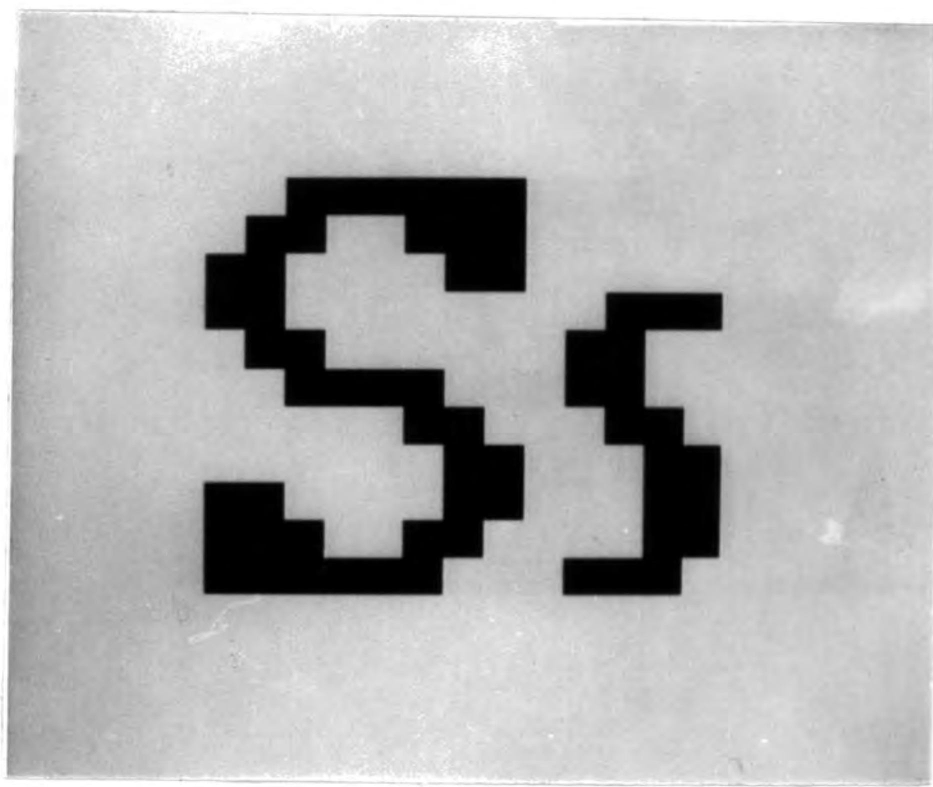
APPENDIX 8.2e: AN EXAMPLE OF THE FONTS USED IN EXPERIMENT 13

(letter characterised by circular strokes; 12 point binary)

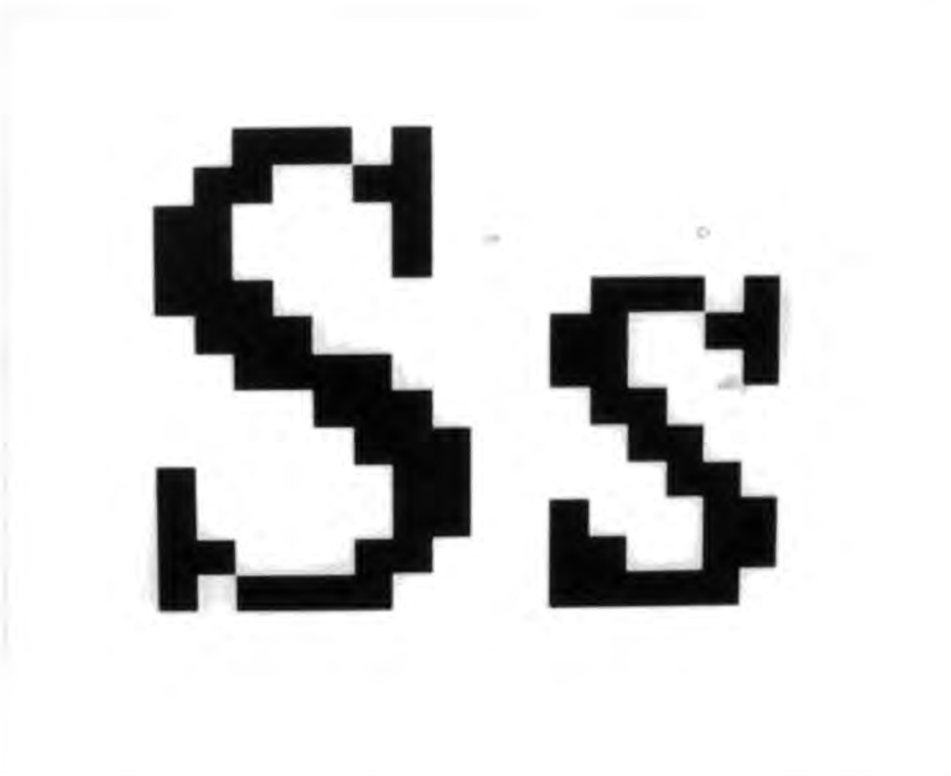


APPENDIX 8.2f:
AN EXAMPLE OF THE FONTS USED IN EXPERIMENT 13
(letter characterised by circular strokes; 12 point AA)

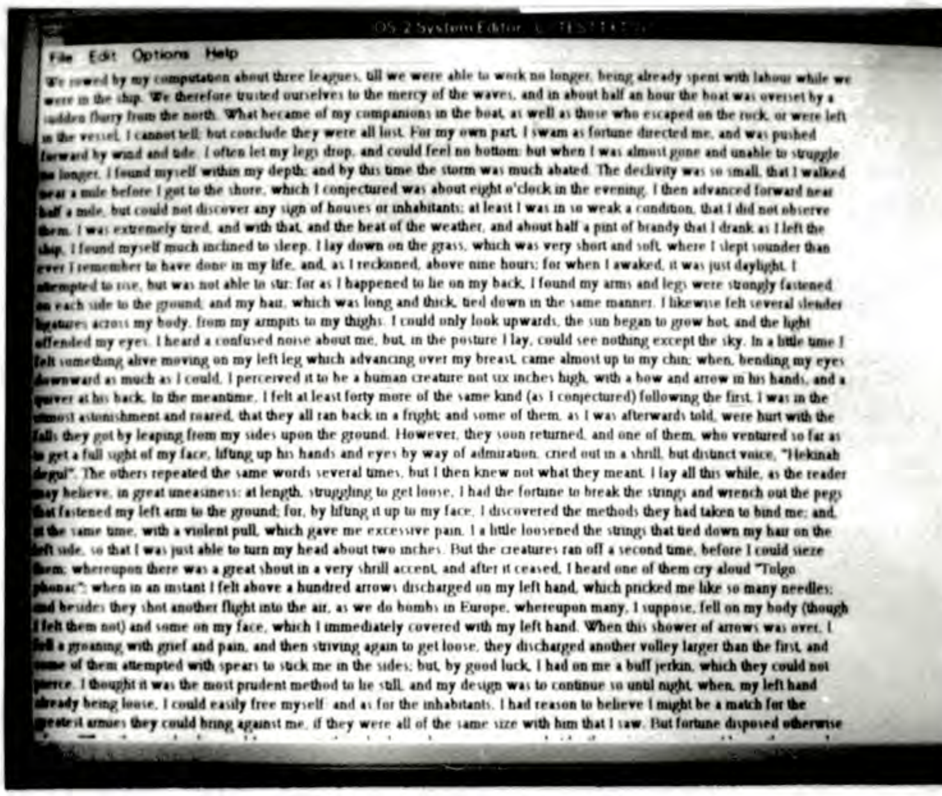
APPENDIX 8.2g:
AN EXAMPLE OF THE FONTS USED IN EXPERIMENT 13
(letter characterised by circular strokes; 12 point bold)



APPENDIX 8.2h:
AN EXAMPLE OF THE FONTS USED IN EXPERIMENT 13
(letter characterised by circular strokes; 14 point binary)



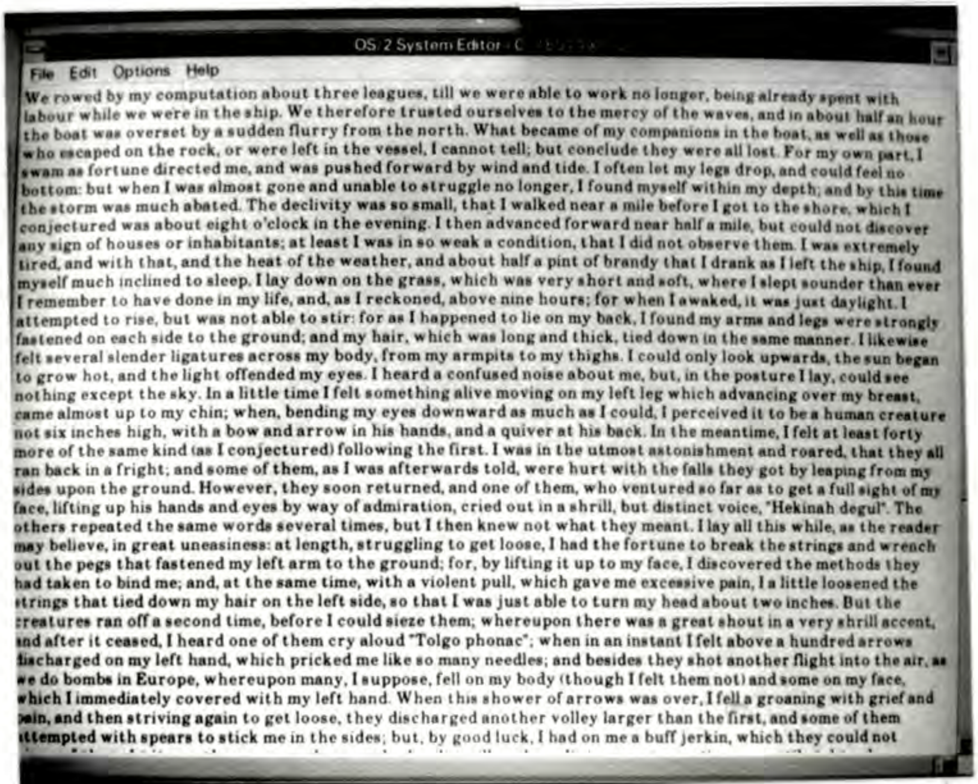
APPENDIX 8.3a: AN EXAMPLE OF THE READING MATERIALS USED IN EXPERIMENT 13 (12 point AA)



File Edit Options Help

We rowed by my computation about three leagues, till we were able to work no longer, being already spent with labour while we were in the ship. We therefore trusted ourselves to the mercy of the waves, and in about half an hour the boat was overset by a sudden flurry from the north. What became of my companions in the boat, as well as those who escaped on the rock, or were left in the vessel, I cannot tell; but conclude they were all lost. For my own part, I swam as fortune directed me, and was pushed forward by wind and tide. I often let my legs drop, and could feel no bottom: but when I was almost gone and unable to struggle no longer, I found myself within my depth; and by this time the storm was much abated. The declivity was so small, that I walked near a mile before I got to the shore, which I conjectured was about eight o'clock in the evening. I then advanced forward near half a mile, but could not discover any sign of houses or inhabitants; at least I was in so weak a condition, that I did not observe them. I was extremely tired, and with that, and the heat of the weather, and about half a pint of brandy that I drank as I left the ship, I found myself much inclined to sleep. I lay down on the grass, which was very short and soft, where I slept sounder than ever I remember to have done in my life, and, as I reckoned, above nine hours; for when I awaked, it was just daylight. I attempted to rise, but was not able to stir: for as I happened to lie on my back, I found my arms and legs were strongly fastened on each side to the ground; and my hair, which was long and thick, tied down in the same manner. I likewise felt several slender ligatures across my body, from my armpits to my thighs. I could only look upwards, the sun began to grow hot, and the light offended my eyes. I heard a confused noise about me, but, in the posture I lay, could see nothing except the sky. In a little time I felt something alive moving on my left leg which advancing over my breast, came almost up to my chin; when, bending my eyes downward as much as I could, I perceived it to be a human creature not six inches high, with a bow and arrow in his hands, and a quiver at his back. In the meantime, I felt at least forty more of the same kind (as I conjectured) following the first. I was in the utmost astonishment and roared, that they all ran back in a fright; and some of them, as I was afterwards told, were hurt with the falls they got by leaping from my sides upon the ground. However, they soon returned, and one of them, who ventured so far as to get a full sight of my face, lifting up his hands and eyes by way of admiration, cried out in a shrill, but distinct voice, "Hekinah degul". The others repeated the same words several times, but I then knew not what they meant. I lay all this while, as the reader may believe, in great uneasiness: at length, struggling to get loose, I had the fortune to break the strings and wrench out the pegs that fastened my left arm to the ground; for, by lifting it up to my face, I discovered the methods they had taken to bind me; and, at the same time, with a violent pull, which gave me excessive pain, I a little loosened the strings that tied down my hair on the left side, so that I was just able to turn my head about two inches. But the creatures ran off a second time, before I could sieze them; whereupon there was a great shout in a very shrill accent, and after it ceased, I heard one of them cry aloud "Tolgo phonac"; when in an instant I felt above a hundred arrows discharged on my left hand, which pricked me like so many needles; and besides they shot another flight into the air, as we do bombs in Europe, whereupon many, I suppose, fell on my body (though I felt them not) and some on my face, which I immediately covered with my left hand. When this shower of arrows was over, I fell a groaning with grief and pain, and then striving again to get loose, they discharged another volley larger than the first, and some of them attempted with spears to stick me in the sides; but, by good luck, I had on me a buff jerkin, which they could not pierce. I thought it was the most prudent method to lie still, and my design was to continue so until night, when, my left hand already being loose, I could easily free myself: and as for the inhabitants, I had reason to believe I might be a match for the greatest armies they could bring against me, if they were all of the same size with him that I saw. But fortune disposed otherwise

APPENDIX 8.3b: AN EXAMPLE OF THE READING MATERIALS USED IN EXPERIMENT 13 (12 point bold)

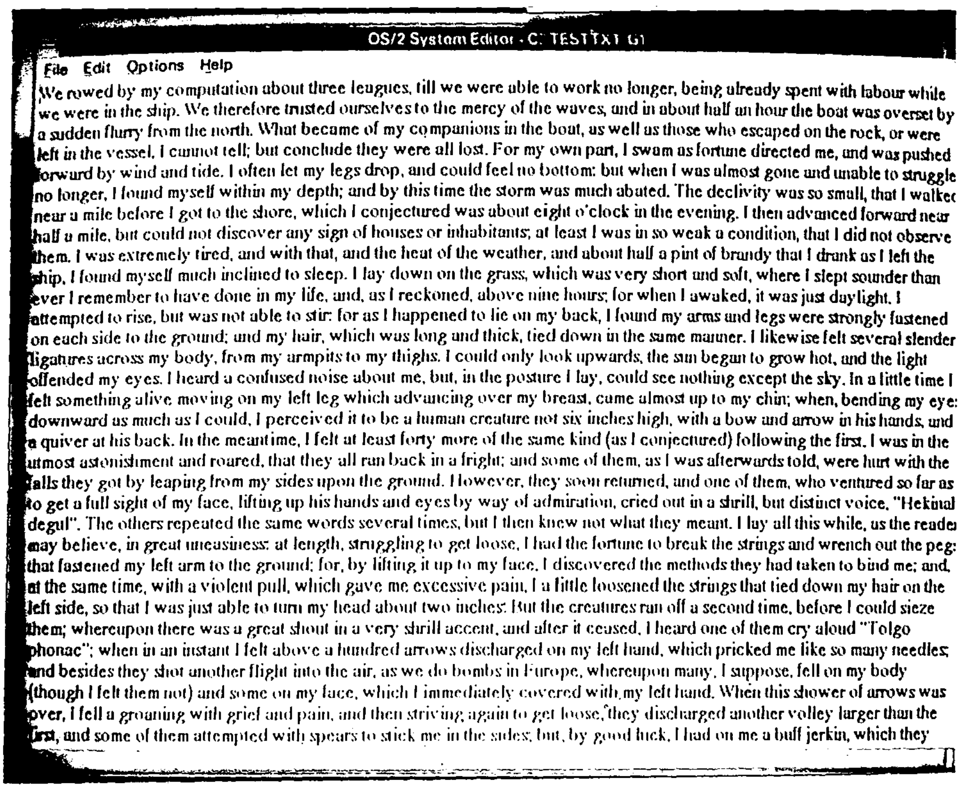


OS/2 System Editor

File Edit Options Help

We rowed by my computation about three leagues, till we were able to work no longer, being already spent with labour while we were in the ship. We therefore trusted ourselves to the mercy of the waves, and in about half an hour the boat was overset by a sudden flurry from the north. What became of my companions in the boat, as well as those who escaped on the rock, or were left in the vessel, I cannot tell; but conclude they were all lost. For my own part, I swam as fortune directed me, and was pushed forward by wind and tide. I often let my legs drop, and could feel no bottom: but when I was almost gone and unable to struggle no longer, I found myself within my depth; and by this time the storm was much abated. The declivity was so small, that I walked near a mile before I got to the shore, which I conjectured was about eight o'clock in the evening. I then advanced forward near half a mile, but could not discover any sign of houses or inhabitants; at least I was in so weak a condition, that I did not observe them. I was extremely tired, and with that, and the heat of the weather, and about half a pint of brandy that I drank as I left the ship, I found myself much inclined to sleep. I lay down on the grass, which was very short and soft, where I slept sounder than ever I remember to have done in my life, and, as I reckoned, above nine hours; for when I awaked, it was just daylight. I attempted to rise, but was not able to stir: for as I happened to lie on my back, I found my arms and legs were strongly fastened on each side to the ground; and my hair, which was long and thick, tied down in the same manner. I likewise felt several slender ligatures across my body, from my armpits to my thighs. I could only look upwards, the sun began to grow hot, and the light offended my eyes. I heard a confused noise about me, but, in the posture I lay, could see nothing except the sky. In a little time I felt something alive moving on my left leg which advancing over my breast, came almost up to my chin; when, bending my eyes downward as much as I could, I perceived it to be a human creature not six inches high, with a bow and arrow in his hands, and a quiver at his back. In the meantime, I felt at least forty more of the same kind (as I conjectured) following the first. I was in the utmost astonishment and roared, that they all ran back in a fright; and some of them, as I was afterwards told, were hurt with the falls they got by leaping from my sides upon the ground. However, they soon returned, and one of them, who ventured so far as to get a full sight of my face, lifting up his hands and eyes by way of admiration, cried out in a shrill, but distinct voice, "Hekinah degul". The others repeated the same words several times, but I then knew not what they meant. I lay all this while, as the reader may believe, in great uneasiness: at length, struggling to get loose, I had the fortune to break the strings and wrench out the pegs that fastened my left arm to the ground; for, by lifting it up to my face, I discovered the methods they had taken to bind me; and, at the same time, with a violent pull, which gave me excessive pain, I a little loosened the strings that tied down my hair on the left side, so that I was just able to turn my head about two inches. But the creatures ran off a second time, before I could sieze them; whereupon there was a great shout in a very shrill accent, and after it ceased, I heard one of them cry aloud "Tolgo phonac"; when in an instant I felt above a hundred arrows discharged on my left hand, which pricked me like so many needles; and besides they shot another flight into the air, as we do bombs in Europe, whereupon many, I suppose, fell on my body (though I felt them not) and some on my face, which I immediately covered with my left hand. When this shower of arrows was over, I fell a groaning with grief and pain, and then striving again to get loose, they discharged another volley larger than the first, and some of them attempted with spears to stick me in the sides; but, by good luck, I had on me a buff jerkin, which they could not

APPENDIX 8.3c:
AN EXAMPLE OF THE READING MATERIALS USED IN EXPERIMENT 13 (14 point binary)



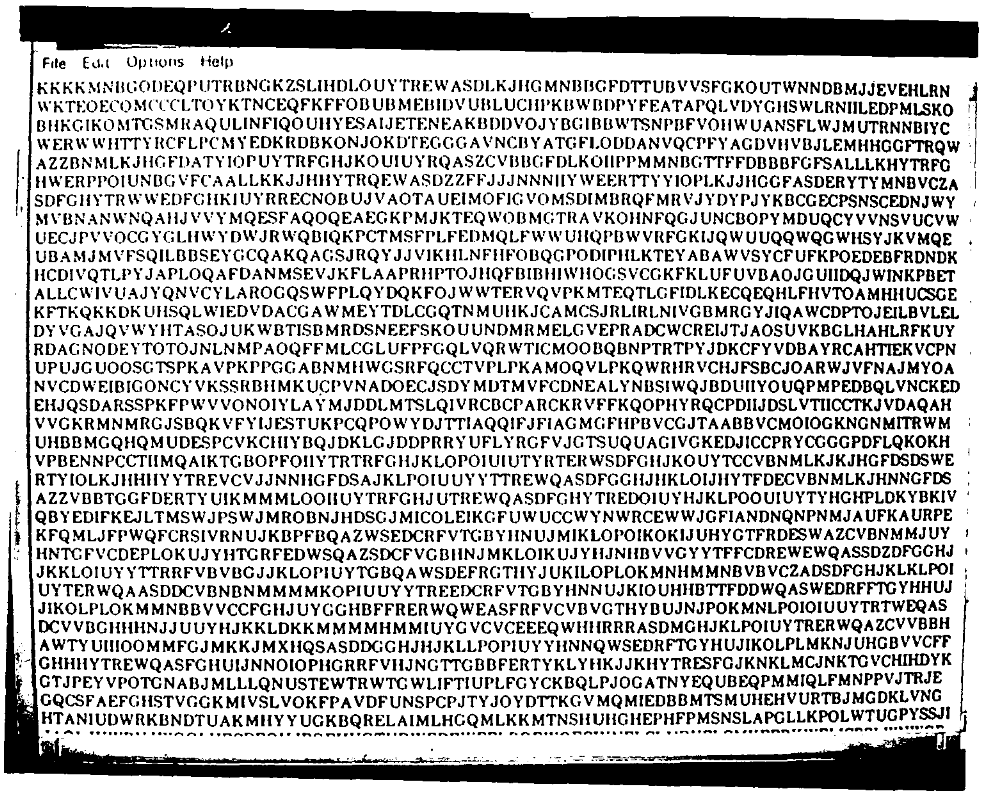
OS/2 System Editor - C: TESTTXT G1

File Edit Options Help

We rowed by my computation about three leagues, till we were able to work no longer, being already spent with labour while we were in the ship. We therefore trusted ourselves to the mercy of the waves, and in about half an hour the boat was overset by a sudden flurry from the north. What became of my companions in the boat, as well as those who escaped on the rock, or were left in the vessel, I cannot tell; but conclude they were all lost. For my own part, I swam as fortune directed me, and was pushed forward by wind and tide. I often let my legs drop, and could feel no bottom: but when I was almost gone and unable to struggle no longer, I found myself within my depth; and by this time the storm was much abated. The declivity was so small, that I walke near a mile before I got to the shore, which I conjectured was about eight o'clock in the evening. I then advanced forward near half a mile, but could not discover any sign of houses or inhabitants; at least I was in so weak a condition, that I did not observe them. I was extremely tired, and with that, and the heat of the weather, and about half a pint of brandy that I drank as I left the ship, I found myself much inclined to sleep. I lay down on the grass, which was very short and soft, where I slept sounder than ever I remember to have done in my life, and, as I reckoned, above nine hours; for when I awaked, it was just daylight. I attempted to rise, but was not able to stir: for as I happened to lie on my back, I found my arms and legs were strongly fastened on each side to the ground; and my hair, which was long and thick, tied down in the same manner. I likewise felt several slender ligatures across my body, from my armpits to my thighs. I could only look upwards, the sun began to grow hot, and the light offended my eyes. I heard a confused noise about me, but, in the posture I lay, could see nothing except the sky. In a little time I felt something alive moving on my left leg which advancing over my breast, came almost up to my chin; when, bending my eye downward as much as I could, I perceived it to be a human creature not six inches high, with a bow and arrow in his hands, and a quiver at his back. In the meantime, I felt at least forty more of the same kind (as I conjectured) following the first. I was in the utmost astonishment and roared, that they all ran back in a fright; and some of them, as I was afterwards told, were hurt with the falls they got by leaping from my sides upon the ground. However, they soon returned, and one of them, who ventured so far as to get a full sight of my face, lifting up his hands and eyes by way of admiration, cried out in a shrill, but distinct voice, "Hekinal degul". The others repeated the same words several times, but I then knew not what they meant. I lay all this while, as the reade may believe, in great uneasiness: at length, struggling to get loose, I had the fortune to break the strings and wrench out the peg that fastened my left arm to the ground; for, by lifting it up to my face, I discovered the methods they had taken to bind me; and, at the same time, with a violent pull, which gave me excessive pain, I a little loosened the strings that tied down my hair on the left side, so that I was just able to turn my head about two inches: But the creatures ran off a second time, before I could sieze them; whereupon there was a great shout in a very shrill accent, and after it ceased, I heard one of them cry aloud "Tolgo phonac"; when in an instant I felt above a hundred arrows discharged on my left hand, which pricked me like so many needles; and besides they shot another flight into the air, as we do bombs in Europe, whereupon many, I suppose, fell on my body (though I felt them not) and some on my face, which I immediately covered with my left hand. When this shower of arrows was over, I fell a groaning with grief and pain, and then striving again to get loose, they discharged another volley larger than the first, and some of them attempted with spears to stick me in the sides; but, by good luck, I had on me a buff jerkin, which they

APPENDIX 8.4a:
AN EXAMPLE OF THE VISUAL SEARCH MATERIALS USED IN EXPERIMENT 13 (12 point AA)

APPENDIX 8.4b: AN EXAMPLE OF THE VISUAL SEARCH MATERIALS USED IN EXPERIMENT 13 (12 point bold)



APPENDIX 8.4c: AN EXAMPLE OF THE VISUAL SEARCH MATERIALS USED IN EXPERIMENT 13 (14 point binary)