

**THE PRINCIPLES OF SCREEN DESIGN FOR COMPUTER-BASED
LEARNING MATERIALS**

BY

Alan Clarke

**University of Sheffield
Department of Information Studies
Doctor of Philosophy (PhD) 1994**

TITLE PAGE

Title: The Principles of Screen Design for Computer-Based Learning

Author: Alan Clarke

Degree: Doctor of Philosophy (PhD)

Department: Department of Information Studies

Submission: November 1994

Acceptance: 1996

Acknowledgements

I would like to acknowledge the considerable help and support I received from a wide range of individuals and institutions. In particular, I would like to thank Christine Clarke, my sons, Nigel Ford, Colin King, Philip Barker, Michael Lloyd-Williams, the staff of the Employment Department's Learning Technology Unit, the staff of the Employment Department and Sheffield University libraries and all the individuals who agreed to become experimental subjects for this research.

CONTENTS

| | |
|--|----|
| Title | 1 |
| Title Page | 2 |
| Acknowledgements | 3 |
| Contents | 4 |
| | |
| List of Tables | 7 |
| List of Figures | 11 |
| | |
| Summary | 12 |
| | |
| 1. Introduction | 13 |
| | |
| 1.0 Introduction | 13 |
| 1.1 What is a principle? | 15 |
| 1.2 Colour | 16 |
| 1.3 Graphics | 17 |
| 1.4 Colour and Learning | 20 |
| 1.5 Graphics and Learning | 22 |
| 1.6 Design of Computer-Based Learning Materials | 25 |
| 1.7 Interactivity | 29 |
| 1.8 Significance of Computer-Based Learning | 32 |
| 1.9 Significance of Screen Design to Computer-Based Learning | 35 |
| 1.10 Current Investigation | 36 |
| 1.11 Conclusions | 38 |
| | |
| 2. Colour, Graphics and Text | 40 |
| | |
| 2.0 Introduction | 40 |
| 2.1 Use of Colour | 40 |
| 2.2 Highlighting and Density of Display | 43 |
| 2.3 Coloured Text and Highlighting | 45 |
| 2.4 Colour Preferences | 49 |
| 2.5 Use of Graphics | 52 |
| 2.6 Survey into the Use of Graphics in Computer-Based Learning Materials | 55 |
| 2.7 Representational Graphics | 56 |
| 2.8 Degree of Realism in Pictures | 57 |
| 2.9 Descriptive and Instructional Captions | 59 |
| 2.10 Analogical Graphics | 60 |
| 2.11 Logical Graphics | 61 |
| 2.12 Ability of Learners | 62 |
| 2.13 Nature of the Task | 64 |
| 2.14 Comparing Text, Colour and Graphic Displays | 65 |

| | |
|---|------------|
| 2.15 Structural or Graphical Organisers | 68 |
| 2.16 Charts and Tables | 70 |
| 2.17 Icons and Hotspots | 72 |
| 2.18 Animation | 74 |
| 2.19 Creation of Computer Graphics | 76 |
| 2.20 Use of Text | 77 |
| 2.21 Density of Text | 79 |
| 2.22 Reading Speeds | 80 |
| 2.23 Scrolling | 84 |
| 2.24 Justification | 87 |
| 2.25 Line Break-Points | 89 |
| 2.26 Upper or Lower Case Letters | 91 |
| 2.27 Comparison of Text Evidence | 91 |
| 2.28 Review of Experimental Design | 92 |
| 2.29 Survey of Computer-Based Learning Materials | 97 |
| 2.30 Conclusions | 106 |
| 3. Experimental Design and Resource Production | 107 |
| 3.0 Introduction | 107 |
| 3.1 Design of Experiments | 107 |
| 3.2 Design of the Colour Experiment | 117 |
| 3.3 Design of the Graphics Experiment | 122 |
| 3.4 Design of Colour Experimental Shell | 127 |
| 3.5 Design of Graphics Experimental Shell | 128 |
| 3.6 Experimental Resources | 140 |
| 3.7 Conclusions | 163 |
| 4. Experimental Results | 164 |
| 4.0 Introduction | 164 |
| 4.1 Characteristics of Subjects | 165 |
| 4.2 Colour Experiment | 172 |
| 4.3 Graphics Experiment | 195 |
| 4.4 Discussion of Results | 240 |
| 4.5 Conclusions | 242 |
| 5. Discussion of Results | 244 |
| 5.0 Introduction | 244 |
| 5.1 Experimental Design | 244 |
| 5.2 The Colour Experiment | 248 |
| 5.3 Changes to the Design of Experimental Shell | 251 |
| 5.4 The Graphics Experiment | 251 |
| 5.5 Operation of Tutorial | 273 |
| 5.6 Conclusions | 274 |

| | |
|---|------------|
| 6. Conclusions | 277 |
| 6.0 Introduction | 277 |
| 6.1 Use of Colour | 277 |
| 6.2 Complexity | 278 |
| 6.3 Learners | 279 |
| 6.4 Use of Graphics | 280 |
| 6.5 Additional Modules | 282 |
| 6.6 Research Literature | 283 |
| 6.7 Next Steps | 283 |
| References | 285 |
| Appendices | |
| 1. Main Program for Graphics Experiment | 303 |
| 2. Timescale of Research | 313 |
| 3. Publications and Presentations of Research Results | 314 |

List of Tables

Chapter 1

| | | |
|---------|---|----|
| Table 1 | Design Stages | 26 |
| Table 2 | Levels of Interactivity (Ayerst, 1987) | 30 |
| Table 3 | Levels of Interaction for Presentation (Gery, 1987) | 31 |
| Table 4 | Cost Comparison - Computer-Based Learning against Conventional Approaches | 33 |
| Table 5 | Training offered to Staff (NOP Social and Political, 1992) | 34 |
| Table 6 | Use of Open Learning (NOP Social and Political, 1992) | 35 |

Chapter 2

| | | |
|----------|--|-----|
| Table 7 | Highlighting and Density of Display | 44 |
| Table 8 | Student Attitudes to Coloured and Highlighted Text (Kelley, 1988) | 47 |
| Table 9 | Colour Preferences of Young Children (Reid and Wicks, 1988) | 50 |
| Table 10 | Colour Preferences of Young Children against Contrast | 51 |
| Table 11 | Guidelines on the Use of Pictures with Text (Levie and Lentz, 1982) | 53 |
| Table 12 | Comparison of Paper and Screen (Bork, 1983) | 54 |
| Table 13 | Comparing the Characteristics of Screen and Paper (Clarke, 1989) | 54 |
| Table 14 | Graphics in Commercial Courseware (Alesandrini, 1985) | 55 |
| Table 15 | Structuring Text (Tullis, 1981) | 67 |
| Table 16 | Mean Response Times (Tullis, 1981) | 67 |
| Table 17 | Use of Structural Displays (Patrick and Fitzgibbon, 1988) | 69 |
| Table 18 | Variables Investigated (Gould et al, 1987) | 82 |
| Table 19 | Correct Responses in Comprehension and Reading Rates as a Function of Window Size and Jump Length (Chen et al, 1988) | 87 |
| Table 20 | Review of Experimental Design | 93 |
| Table 21 | Summary of Survey | 99 |
| Table 22 | Analysis of Full and Part-Screen Images by Type | 101 |
| Table 23 | Analysis of Text (Foreground) Colours | 102 |
| Table 24 | Analysis of Use of Icons/Hotspots in Courseware | 102 |
| Table 25 | Number and Type of Icon/Hotspots | 103 |

Chapter 3

| | | |
|----------|---|-----|
| Table 26 | Comparison of Computer-Based Learning Survey and Graphics Experiment | 117 |
| Table 27 | Advantages and Disadvantages of using the Same or Different Subjects. | 124 |
| Table 28 | Screen Designs Graphics Experiment | 132 |

| | | |
|----------|---|-----|
| Table 29 | Complexity Measure (Tullis, 1983) | 138 |
| Table 30 | Examples of IBM Graphics Standards (Jackson et al, 1994) | 148 |
| Table 31 | Units of Colour Experiment | 159 |
| Table 32 | Units in Graphics Experiment | 160 |

Chapter 4

| | | |
|----------|--|-----|
| Table 33 | Characteristics of Subjects (Colour Experiment) | 165 |
| Table 34 | Gender and Qualifications of Subjects (Colour Experiment) | 166 |
| Table 35 | Highest Qualifications (Colour Experiment) | 166 |
| Table 36 | Highest Qualification in Great Britain (Labour Force Survey, 1988) | 167 |
| Table 37 | Characteristics of Subjects (Graphics Experiment) | 168 |
| Table 38 | Characteristics of Expert Subjects (Graphics Experiment) | 168 |
| Table 39 | Characteristics of Intermediate Subjects (Graphics Experiment) | 169 |
| Table 40 | Characteristics of Beginner Subjects (Graphics Experiment) | 169 |
| Table 41 | Highest Qualification (Graphics Experiment) | 170 |
| Table 42 | Colour and Graphics Experimental Samples | 171 |
| Table 43 | Standard Error and Sample Size | 172 |
| Table 44 | Means and Standard Deviations for Six Colour Experimental Versions | 173 |
| Table 45 | Colour Experiment Means and Standard Deviations for each Type of Display | 174 |
| Table 46 | Frames Accessed and Mean Time (Colour Experiment) | 174 |
| Table 47 | Summary of Mann-Whitney Tests for Colour Experiment | 177 |
| Table 48 | Colour Experiment Test Scores from On-Line Test | 186 |
| Table 49 | Results of Attitude Questionnaire (Colour Experiment) | 188 |
| Table 50 | Routes (Colour Experiment) | 190 |
| Table 51 | Choice of Modules of Tutorial (Colour Experiment) | 191 |
| Table 52 | Access to Tutorial Modules (Colour Experiment) | 191 |
| Table 53 | Range of Use (Graphics Experiment) | 196 |
| Table 54 | Complexity Measurements (Text-Only Displays) | 199 |
| Table 55 | Complexity Measurements (Text and Graphics Displays) | 200 |
| Table 56 | Mean Complexity by Screen Type | 201 |
| Table 57 | Complexity of Screens compared to Text Screens | 205 |
| Table 58 | Predictions of Mean Time | 206 |
| Table 59 | A Comparison of the Actual Mean Time by Learner Group compared to Complexity Prediction | 207 |

| | | |
|----------|---|-----|
| Table 60 | Results of Attitude Questionnaire (Graphics Experiment) | 211 |
| Table 61 | Size of Illustrations - Preferences | 212 |
| Table 62 | Types of Graphics - Preferences | 212 |
| Table 63 | Examples of Types of Screen | 214 |
| Table 64 | Tutorial and Additional Modules | 219 |
| Table 65 | Graphics Experiment Modules Accessed | 220 |
| Table 66 | Choice of Modules of Tutorial (Graphics Experiment) | 221 |
| Table 67 | Tutorial Module Access (Graphics Experiment) | 221 |
| Table 68 | Cross-References (Examples Module) | 222 |
| Table 69 | Access to the Dictionary | 223 |
| Table 70 | Access to the Dictionary from Main and Pop-Up Menus | 223 |
| Table 71 | Access to More Information Module | 224 |
| Table 72 | Access to Quiz Module | 225 |
| Table 73 | Access to Quiz Module from Main and Pop-Up Menus | 225 |
| Table 74 | Learners Modes of Operating (Graphics Experiment) | 226 |
| Table 75 | Users Mean Times | 227 |
| Table 76 | Comparison of Various Screen Designs with Text-Only Screens | 228 |
| Table 77 | Comparison of Various Screen Designs with Text-Only Screens (Beginner) | 229 |
| Table 78 | Comparison of Various Screen Designs with Text-Only Screens (Intermediate) | 230 |
| Table 79 | Comparison of Various Screen Designs with Text-Only Screens (Expert) | 231 |
| Table 80 | Comparison between different Graphic Screen Displays (All Subjects) | 233 |
| Table 81 | Comparison between different Graphic Screen Displays (Beginner) | 234 |
| Table 82 | Comparison between different Graphic Screen Displays (Intermediate) | 235 |
| Table 83 | Comparison between different Graphic Screen Displays (Expert) | 236 |
| Table 84 | Beginners' Answers (Quiz) | 238 |
| Table 85 | Intermediates' Answers (Quiz) | 239 |
| Table 86 | Experts' Answers (Quiz) | 240 |

Chapter 5

| | | |
|----------|---|-----|
| Table 87 | Comparison of Degree of Significance against Learners Group for Text compared to Structured Text Screens. | 256 |
|----------|---|-----|

| | | |
|----------|--|-----|
| Table 88 | Comparison of Degree of Significance against Learners Group for Text compared to Graphics Screens (t test). | 257 |
| Table 89 | Predictions of Time based on Text and Representational Graphic Screens. | 258 |
| Table 90 | Comparison of Degree of Significance against Learner Group for Text compared to Representational Graphics (t test). | 258 |
| Table 91 | Comparison of Degree of Significance against Learner Group for Text compared to Full-Screen Representational Graphics (t test). | 259 |
| Table 92 | Comparison of Degree of Significance against Learner Group for Text compared to one Part-Screen Representational Graphics (t test). | 260 |
| Table 93 | Comparison of Degree of Significance against Learner Group for Text compared to two Part-Screen Representational Graphics (t test). | 261 |
| Table 94 | Comparison of Degree of Significance against Learner Group for Text compared to four Part-Screen Representational Graphics (t test). | 263 |
| Table 95 | Representative Graphics (Time Differences with Predictions) | 264 |
| Table 96 | Learners' Choice of Route through the Tutorial | 267 |
| Table 97 | Comparison of Examples and More Information Modules | 270 |
| Table 98 | Accumulative Effect of References | 271 |

List of Figures

Chapter 1

| | | |
|----------|--|----|
| Figure 1 | Diagram of Development Time/Cost against Presentation Features | 27 |
| Figure 2 | Illustration of a Single Frame | 27 |
| Figure 3 | Simple Branching Structure | 28 |
| Figure 4 | Interaction Cycle | 29 |

Chapter 2

| | | |
|----------|--|----|
| Figure 5 | Normal Line Breaks - Syntactic Line Breaks (Lefrere, 1984) | 89 |
|----------|--|----|

Chapter 3

| | | |
|-----------|--|-----|
| Figure 6 | True Experimental Designs | 113 |
| Figure 7 | Experimental Overview - Allows for up to T(n) Treatments | 119 |
| Figure 8 | Overall Structure (Colour Experiment) | 122 |
| Figure 9 | Tutorial Structure (Colour Experiment) | 123 |
| Figure 10 | Overall Structure (Graphics Experiment) | 125 |
| Figure 11 | Tutorial Structure (Graphics Experiment) | 126 |
| Figure 12 | Menu Screen (Colour Experiment) | 128 |
| Figure 13 | Tutorial Screen (Graphics Experiment) | 130 |
| Figure 14 | Visual Acuity | 139 |
| Figure 15 | Main Menu (Graphics Experiment) | 146 |
| Figure 16 | Pop-Up Menu (Graphics Experiment) | 147 |
| Figure 17 | Turbo Pascal Floppy Disc Procedure | 149 |
| Figure 18 | Rmaster Image of a Floppy Disk | 150 |
| Figure 19 | Iconnew Image of a Floppy Disk | 150 |
| Figure 20 | ReadXgf (Graphics Procedure) | 151 |
| Figure 21 | Turbo Pascal Procedures Results and Collector | 153 |
| Figure 22 | Exit Window | 157 |
| Figure 23 | Structure of Question Module (Colour Experiment) | 157 |
| Figure 24 | Structure of Input Module (Graphics Experiment) | 158 |

Summary

The Principles of Screen Design for Computer-Based Learning

Alan Clarke

The critical interface between learners and computer-based learning materials is the screen. If the display of learning is not effective then learning will be hindered.

Screen design is therefore an important element in the design of computer-based learning.

This research investigated the three fundamental screen design elements of text, colour and graphics. A review of literature, experimental design and a limited survey of computer-based learning materials provided the background for this research. The experimental materials reflected the results of the review and survey by using representative subjects, providing a learning focus and employing computer-based materials. Two experiments were undertaken. The Colour and Graphics experiments considered the effects of a number of variables on learners' behaviour which included: the use of colour; the size and type of graphics; the learner's prior knowledge of tutorial subject; and the complexity of the display.

The results of this research showed that colour is a powerful motivating force as long as it is not used excessively. This was identified as the use of more than seven colours. Graphics can be used more extensively in current computer-based learning materials and users preferred representational graphics occupying a quarter to a half of the screen. However, learners were not prepared to make the effort to either use analogical graphics to make links with their prior knowledge or to extract information contained in the structure and form of logical graphics. Subjects were motivated by representational graphics.

Learners' behaviour in relation to the various screen displays they encountered was affected by their prior knowledge of the tutorial content. This was apparent in their choice of options (additional modules) within the tutorial, their methods of interacting with the material and their responses to individual displays.

Chapter One

INTRODUCTION

1.0 Introduction

Computer systems are increasingly providing the information which is supporting society. Building societies, schools, libraries, airline reservation desks and many others are using interactive computer displays to provide information, learning, guidance and support. In the field of computer-based learning students use a screen as part of an interactive learning programme. Computer-based learning is used to simulate a working environment, act as a dynamic electronic tutorial or provide a database of reference material for a learner. The use of computer-based learning is already established in many sectors of the economy such as banks, building societies, insurance, retail, transport, energy and telecommunications (Hawkrige, Newton and Hall, 1988). Computer-based learning has grown rapidly over the last ten years (Race and Brown, 1994). This trend is likely to accelerate with the availability of equipment and the falling cost of delivering electronic information compared to conventional approaches.

During the period 1985 to 1993 there was a substantial growth in the availability of computers in British primary and secondary schools. In 1985 there was a ratio of 107 pupils to each computer in primary schools and 60 pupils to each computer in secondary schools (Department of Education and Science, 1991). By 1993 the ratio of pupils to microcomputer was 25:1 in primary and 13:1 in secondary schools. Primary schools are now spending on average £2600 per year on all aspects of information technology. Secondary schools are spending on average £15450 per year (Department for Education, 1993).

The key question to be answered is how to design computer-based learning to aid, enhance and improve learning. Several researchers (Galitz, 1985; 1993; 1994; Lucas, 1991; Sherr, 1993) have shown that good screen design is critical in bridging the gap between humans and machines. An effective interface between a learner and a computer can provide a bridge that enables effective communication to take place.

Bork (1983) states that: 'the vast majority of screen displays are probably far less effective than they could be'. This is not simply a matter of legibility but of designing the display to suit the reader's purposes and motivation. Spannaus (1985) feels that it is essential that every screen should be attractive, easy to understand and instructionally effective. Most designers appear to concentrate on the design of an overall lesson while neglecting the individual screens which make it up. The appearance of information on a screen is probably as much to do with the programming language or authoring system being used as any decision by the author as to the way it appears. This may well be due to the principles of good design not being understood. Barfield (1993) has compared the current state of the computer user interface to buying a knife without a handle. A well designed screen can increase the user's motivation, reduce errors and aid human processing of the information. A poorly designed screen could have the opposite effect.

The design of effective computer-based learning material involves an author in making a number of screen design decisions. An author must combine graphics, text and colour into a structure which will develop a dialogue with a learner in the form of information, questions, responses and feedback. The dialogue produced by the material will aim to assist the user to learn the information displayed. Other factors such as learning objectives are also involved in the creation of computer-based learning materials and will influence screen design.

Several researchers (Moran, 1981; Sweeters, 1985; Rivlin et al, 1990) have produced guidelines to assist authors with screen design. However, they have all stressed that their guidelines need to be applied in the light of the content, learner and environment in which the material will be used. The approach taken by most designers is systematic but personal experience and intuition play a large part in it. A wide range of rules are quoted about design but few seem supported by evidence. Many rules or guidelines contradict each other. Rivlin et al (1990) state that colour should be used sparingly and that two or three colours at a time are sufficient. In contrast, Long (1984) advises that between five and twelve colours may be used depending on the nature of the information displayed. The difference in the guidelines is probably due to the context each writer is considering and the lack of underpinning evidence.

There is a general assumption that the use of computer graphics and colour enhances computer-based learning. In fact the more colour and graphics used the better, providing it is aesthetically balanced or pleasing to the eye. However, the evidence for these assumptions is difficult to find. In general terms, the computer display should be designed to aid learners to achieve their goals in an efficient and effective way. The design of learning materials should be compatible with both the subject and characteristics of the learners.

Stoddard (1985) states that central to the design of effective human computer interfaces is an understanding of the user. Individual differences will be critical in how a user will react to a particular screen design. Users' previous experiences with computers may well relate to their success in using the computer-based learning materials.

1.1 What is a Principle?

This study seeks to identify the principles of screen design for computer-based learning materials. The main areas of interest are:

- (1) the use of colour; and
- (2) graphics.

Gardiner and Christie (1987) offer explanations of guidelines, principles and standards in the context of user-interface design. A design guideline is usually a straightforward recommendation which is often accompanied by examples of its use, exceptions and links to other guidelines. A design principle is generally a more abstract recommendation which is frequently expressed in psychological or conceptual terms. A principle will need more interpretation within the context of a particular design issue than will a guideline. In contrast to both guidelines and principles, standards are unambiguous statements of interface requirements. Standards are frequently formal statements of the requirements of a design which are imposed on designers.

1.2 Colour

A colour screen image is made up of coloured elements called pixels or dots. Each of these is made by combining different intensities of red, green and blue. These different colours are actually represented by a pattern of phosphors arranged in a triangle. Each one of these groups (called a triad) can then be activated to form a point on the screen. The larger the number of these triads that go towards the make up of the screen, the greater the resolution it will be capable of displaying.

The distance between successive triads is known as the dot pitch and the smaller the figure that is quoted for this, the better the monitor is likely to be. Typical figures are 0.31 mm and 0.28 mm, the latter giving the better sharpness. Obviously, there is more to making a monitor than simply arranging the phosphors on the screen, so this should not be taken as a final indication of the quality of the image that is produced. A printed page is normally shown only in two colours - black and white with black text on a white background. A screen offers a number of combinations of colours for the foreground and background. Many colours can be used in the way an author desires. The colours available are dependent on the computer system used but many computers offer thousands of different shades to users.

Colour is created on a visual display unit by the use of three electron projection tubes, one each for red, blue and green. Three corresponding phosphors are deposited on the inside of the screen. These are configured in groups of three dots, or three stripes, one of each type. The three electron beams are directed onto the appropriate phosphor dots or stripes by a 'shadow mask'. A 'shadow mask' is a thin metal sheet containing many holes or slots. Colour is created by using different combinations of the three colours and varying the intensity of the electron beam.

The purpose of designing computer-based learning materials is to create situations through which people will learn. Mayer (1982) stated that learning is the relatively permanent change in a person's knowledge or behaviour due to experience. Information can be displayed on a screen in many ways. The author of computer-based learning materials is seeking to design each screen of information in a way that

will enhance learning. Colour can be used in many ways to emphasise the display of information by showing the relationships between items, reflecting different types of condition or situation and strengthening the layout of information. The key to the successful use of colour is the understanding of its effect on the user of the display that you are creating. The display of information on a screen is mainly based on the judgement and experience of the author rather than sound design principles.

Barker (1987) has stated that:

'There is still much to be learned about the best ways to use colour in both alphanumeric displays and in high resolution graphics systems'.

1.3 Graphics

There are two main types of graphic display which are vector and raster. Vector graphics treat each image as a collection of line segments. All objects in the image are composed of line segments. The line segments are stored mathematically and all parts of a vector image such as squares and circles are mathematically defined.

Raster graphics represent each image as a bit map, that is by lighting up individual pixels on a grid. Raster graphics are capable of producing more detailed images because each pixel can be changed to reflect subtle differences in shading and colour.

The final element is the computer's graphic card and the monitor being used, since the resolution of different systems varies considerably.

The Acorn BBC B offers resolutions from 640 x 256 to 80 x 75 depending on which mode of operation is used, while the IBM PC offers a range of resolutions from 320 x 200 (CGA), 640 x 200 (EGA), 800 X 600 (VGA) to 1280 X 1024 (Ultra VGA) in colour systems and 720 x 348 in the Hercules monochrome system. The interesting comparison is that the printed page has a resolution of 2000 x 2000 addressable points. The resolution of a photograph (with a good lens and fine grain film) is 2000 x 3000 while an ordnance survey map (1:50000) has a resolution of 50000 x 50000 (Rivlin et al, 1990). This simple comparison of resolutions available

between the computer screen and other media illustrates that the legibility of a computer display may be a key variable. This would be especially true if the design guidelines of instructional text are simply transferred to a computer screen.

The type of monitor on which a graphic image is displayed is critical to the quality of the image. Two important factors which influence image quality are:

- (1) refresh rate; and
- (2) interlacing.

A monitor displays an image on the screen by painting a picture on the inside of the screen with an electron gun. The gun's electronic beam moves backwards and forwards across the screen. This causes the phosphors on the screen to glow. However, they glow only briefly and need to be refreshed at intervals to continue to glow. The rate at which the electron gun repaints the screen is called the refresh rate. A slow refresh rate may cause images to waver in intensity and users to have eye strain.

In addition to monitors having different refresh rates, some monitors and graphics cards use interlacing to speed up refresh rates. Interlacing works by skipping every other row when refreshing the screen. That means for example that on the first pass, alternative rows are refreshed. On the second pass the remaining rows are refreshed and the process is then repeated. Interlacing can cause eye strain over a long period of time.

There are many forms of graphics which can be used in computer-based learning materials such as pictures, diagrams, charts, graphs and animation. Graphics can serve many purposes. They can motivate, attract attention, excite, amuse, entertain and persuade (Murray, 1992) but they may also distract the learner (Rivlin et al, 1990). Graphics can contain a great deal of information in a small space compared to text. In order to use this information the user must interpret the graphics. Users may interpret graphics in different ways.

There are a number of ways of classifying graphical images. This thesis is concerned with screen design aimed at facilitating learning. It would therefore be appropriate to identify images in terms of their potential to aid learning; that is, how

an image conveys meaning and supports learning. The three main types of graphics are representational, logical and analogical (Knowlton, 1966; Alesandrini, 1984; 1987). Reviews of research finding have suggested that each of these types of graphic support learning (Levie and Lentz, 1982; Alesandrini, 1985).

Representational graphics are realistic images which are used mainly to convey information. However, they also serve to motivate and attract the attention of learners to key issues. Logical graphics take a variety of forms such as charts, diagrams and tables. Their purpose is to provide learners with spatial and structural information and depict complex relationships. Analogical graphics are images intended to encourage learners to make links with their prior knowledge and experience. An image of a race horse may allow a learner to associate it with prior understanding of speed, competition and other animals.

All three types of graphics are broad classifications including a wide range of images. This is particularly true of logical images which include charts, diagrams and tables all of which could be considered as separate types. What they have in common is that they are all concerned with structuring and conveying information in both the content and structure of the image.

All three types of graphic illustration require designers to make careful choices between illustrations. Many designers have been encouraged to use graphics to provide variety and interest in their materials (Lewis and Paine, 1985). There is also considerable evidence that many learners prefer visual to verbal methods of information presentation. The practical problems for designers of computer-based learning are how to choose which graphics to include; how to create the precise image required (especially if a designer has no graphical skills); when to use graphics and how to display an image.

Kearsley et al (1983) reviewing two decades of research into computer-aided instruction concluded that designers do not have a good understanding of the use of graphics in instruction.

1.4 Colour and Learning

A designer of computer-based learning materials must consider the effects of colour on learners. Several researchers have reported that learners prefer colour to black and white materials (Katzman and Nyenhius, 1972; Donahue, 1973). The fact that learners prefer colour in materials does not prove that colour aids learning. Descy (1981) considers that learning has three elements:

- (1) acquisition of central material presented to learners;
- (2) acquisition of peripheral material presented to learners; and
- (3) retention of the material that has been learned.

That part of the instruction which an author is trying to communicate to a learner is the central material while anything else is peripheral material. Central material acquisition is the main concern of designers of learning materials.

Many researchers have reported that colour does not significantly aid learning of central material. Kanner and Rosenstein (1961) compared the effectiveness of colour and monochrome television instruction and could not find any difference between the media in terms of learning. Katzman and Nyenhius (1972) used colour in titles, captions and labels but they were unable to find any improvement in the learning of central material. When comparing colour with black and white presentations they concluded that colour aided the recall of the material. However, other researchers have come to the opposite conclusion (Scanlon, 1967).

Chute (1980) and VanderMeer (1954) have both carried out research into the effects of colour in instructional films. VanderMeer (1954) tested his learners immediately after watching the film and six weeks later, using both verbal and non-verbal tests. He concluded that colour did not help in the immediate acquisition of learning but did help the retention of the material. Chute (1980) reported that "colour cueing may be more instrumental for enhancing long term retention".

This evidence does not clearly indicate any advantage of using colour in learning materials except that learners prefer colour and that colour aids learning retention. The confusing picture presented by the research may be explained by the

fact that different media were used - for example film, 35 mm slide and television and a wide variety of learners. The differences in media and learners may well account for the differences in results. Future research will need to consider both variables.

Colour is a potentially powerful tool to the designer but it is easy to misuse. It is common to see colour used as a pure gimmick in many computer games and some educational software. Colour must be used to aid learning not to hinder it.

Colour perception is a process of three interacting variables - hue, brightness and saturation. Hue is normally what most people think of as colour since red and green are different hues. Brightness is closely related to the intensity of light reaching the retina. The higher the intensity of the light source the more brightly coloured they appear. The retina is also sensitive to the wavelength of the colour in the light spectrum so that yellow is perceived as the brightest colour while red and blue are the least bright. Saturation is diminished when the interaction of hue and brightness is reduced by adding white light. For example, a fully saturated red becomes pink when white light is added. It would still be described as having a hue of red but with decreased saturation.

An important variable is contrast, or the relative brightness of foreground over background. In simple terms, the greater the contrast, the more readable the display. So red on black should not be as readable as yellow on black. This is also applicable to graphic images which will appear sharper the greater the contrast. If contrast is reduced, the detail of the image will be more difficult to see.

One crucial factor in considering how to use colour is the fact that 6 to 10 percent of the male and 0.05 percent of the female population are colour blind, that is, they have defective colour receptors and are not able to perceive certain colours. This at first seems a difficult issue to overcome but by combining shape and colour together, the colour blind people are provided with the necessary cues. This does mean, though, that two systems must be used instead of one and thus cost and production time will increase. Tullis (1981) used shape with a black and white display to code the image instead of colour with similar results to a colour display.

A key problem for designers is how many colours to use at any one time. The general guidance is to use colour sparingly. Usually missing from this advice however is the reason for the restriction. Miller (1956) has stated that the immediate

memory can only compare, retain and respond to a limited number of items at a time which is in the range seven plus or minus two. An example would be seven unrelated words or seven rule-related chunks of information. A rule-related chunk of information could be considered as an organised group of items of information linked by a rule. The long term memory seems to operate in a similar way in that it remembers chunks of information but it does not have an upper limit of chunks. Colour coding could provide the means of distinguishing a chunk of information.

People can distinguish a great many colours when comparing them directly but only a few colours can be recognised on their own (DeMars, 1975). It is possible to distinguish up to five colours almost perfectly (Smith, 1962). This is thus a base line with tasks that require a high degree of colour discrimination. Where a high degree of colour discrimination is not required then up to twelve colours could be used (Smith, 1963; Smith and Thomas, 1964). The evidence would suggest a range of five to twelve colours which could be used to design screens. The other variable which may affect the number of colours used is the nature of the display. If a display presents a great deal of information on a single screen then colour may well be a useful device to separate items. However, Cahill and Carter (1976) showed that the reverse is true and on displays of dense information, the number of colours should be restricted.

1.5 Graphics and Learning

Psychological principles of visual perception are essentially the principles which affect the communication between a human and a computer. The communication is frequently visual in that an individual reads a screen display and reacts to it by some action. Learners read a tutorial shown on a computer screen which may aid their learning. The communication is essentially visual and can be considered to be the principles of cognition applied to display.

Almost all the studies of human memory have shown that organised material is perceived, comprehended, retained and retrieved better than comparable but unorganised material. Miller's (1956) limit of items which immediate memory could

retain was mainly concerned with verbal information such as words, sentences and labels. Visual displays showing pictures or images alone seem to be processed as single items and do not seem to be as limited as verbal data. A large body of research (Standing, Conezio and Haber, 1970; Standing, 1973) is available which deals with recognition memory of pictures. Experiments are usually based on showing subjects a large number of pictures and testing their recall of the pictures. The experimental results have demonstrated that subjects can recognise a very high proportion of the pictures seen. Erdelyi and Stein (1981) showed that in a rote memory task with either words or pictures that both recall and recognition improved with repeated attempts to remember pictures but got worse with repeated attempt to recall words. If the subjects were asked to create a visual image of the word then repeated attempts to remember the word improved. It appears that the memory of pictures is not limited to a finite number of items. Each picture is a coherent chunk with its own organisation of form, colour, texture and size which makes it possible to perceive and store its information.

Even when the picture's features are minimal, the viewer can structure images provided that the picture is meaningful. Freedman and Haber (1974) presented adults with silhouettes consisting of black and white blotches. Some of the silhouettes were intended to show faces of people but not all of the subjects could see the faces in the pictures. The subjects were later tested for their recognition memory and had no memory at all of the blotches in which they did not see faces but easily recalled those they had seen as faces. It seems clear that if the subject can achieve a coherent organisation of the picture then the image is better retained.

The visual system (Haber and Wilkinson, 1982) is designed to produce organised perception. Information consisting of a variety of features such as size, shape, colour and position is structured by the mind to represent a visual scene. The visual scene is an object in which all parts are perceived together in a single construction. The individual features are not perceived as separate independent elements. If objects are shown together they are perceived in relation to each other - that is, near, far, behind, in front, to the right and to the left. The objects are not perceived in isolation from each other. A computer's display could therefore be considered as a single object or a group of related objects.

Visual systems attempt to interpret all stimuli reaching the eyes as if they are reflected from a real scene in three dimensions. This occurs even if the information is actually coming from a flat surface such as a computer display.

These principles seem to imply that visual information conveys structure without the need for the individual to know the meaning, shape, or content of the information being displayed. Users of visual information have excellent comprehension, retention and retrieval of the information compared to non-visual information.

A visual display in which all the features correspond to the information being communicated does not guarantee effective communication. The visual system may process the image in a way that changes the one-to-one relationship. However, the principles of visual perception do offer ways to display information on the screen which takes advantage of human visual ability.

Haber and Wilkinson (1982) considered tasks which were sequential in nature, that is, tasks which must be performed in a set order. The list of operations was represented as a picture so that the order of the operation corresponded to the spatial position of the image. The task chosen was the making of a sponge cake in which two groups of subjects were presented with instructions on how to make the cake. One group received the instructions in the traditional cook book style while the second saw the instructions presented as a pictorial flow chart. After they had studied the material both groups were asked to recall ingredients, sequence of steps and procedure, immediately after studying and twenty minutes later. The students who used the pictorial flowchart performed better than the other group. Two later experiments using the tasks of learning a car route and chemical testing confirmed this result.

Hartley (1980a; 1980b) showed that comprehension and retention are greater from text which has each semantic element marked by indentation compared to traditional printed formats. This can be explained by the fact that the display is visually organised in a similar way to the pictorial flowchart. There are a number of ways of visually organising information in addition to spatial position such as colour, intensity, texture, or shading. The devices are different means of increasing the

contrast between adjacent visual features. The relative merits of these devices is not clear and designers choose between them more through instinctive judgement than by any application of objective judgement.

1.6 Design of Computer-Based Learning Materials

There are many ways of designing and developing computer-based learning materials. The two extremes would be an individual teacher or trainer and a commercial team of people who have come together with the purpose of creating the material. An individual is probably working at home in the evenings and weekends with no fixed work plan, timetable or resources except personal skill and knowledge. A team will comprise a wide variety of specialists who may be working with several different production teams. A commercial computer-based learning team consists of instructional designers, graphic artists, technical specialists, subject matter experts and a project manager. This team will almost certainly be working to a fixed timetable with an associated work plan and an agreed completion date.

Both approaches could be successful in producing a creative, stimulating and error-free piece of learning material. The key factor is not the way the material is produced but the individual's or team's understanding of the design of computer-based learning materials. Kearsley (1986) states that 'The critical factor is not the approach used, but the extent to which the design principles are understood'.

The design process for computer-based learning materials can be divided into a number of key stages. Many researchers (Kearsley, 1986; Barker, 1987; Gery, 1987) have produced lists of the stages involved. Table 1 summarises the key stages involved. The design of the material can be considered at two broad levels which are:

- (1) macro - overall lesson plan of learning objectives; and
- (2) micro - design of each frame of material.

These levels are not totally separate from each other and the material relies heavily on their successful integration to be useful. A set of exciting and motivating screens

which do not support the learning objectives of the material are not going to enhance learning. Equally an excellent overall design which neglects the individual screens is also failing to maximise the learning potential of the material. The individual designer and the commercial team both need to consider both the macro and micro aspects to be successful.

Table 1 Design Stages

| |
|--|
| <ol style="list-style-type: none">1. Identify learning need2. Analyse the task3. Set learning objectives4. Select media5. Specify lessons/modules6. Script development7. Storyboard the material8. Design presentation of the material9. Assess material10. Revise the material |
|--|

The production of computer-based learning material is expensive both in time and money. Gery (1987) states that one hour of computer-based training material will require three hundred hours of development time. The Open University (Rowntree, 1992) experience suggests that two to four months of professional time is required to produce one hour of material, that is, 350 to 700 hours of development time for each hour of computer-based learning material. This development ratio will increase with the use of additional media or presentation features (Kearsley, 1986) in the material and the increase in interaction between the material and the user. Figure 1 shows the relationship between development time and presentation features. In simple terms, it shows that as extra presentational features are added the

development time and cost increases. Computer-based learning material has a number of features and components which make up its structure. The basic structural feature is a single frame, which is illustrated in figure 2.

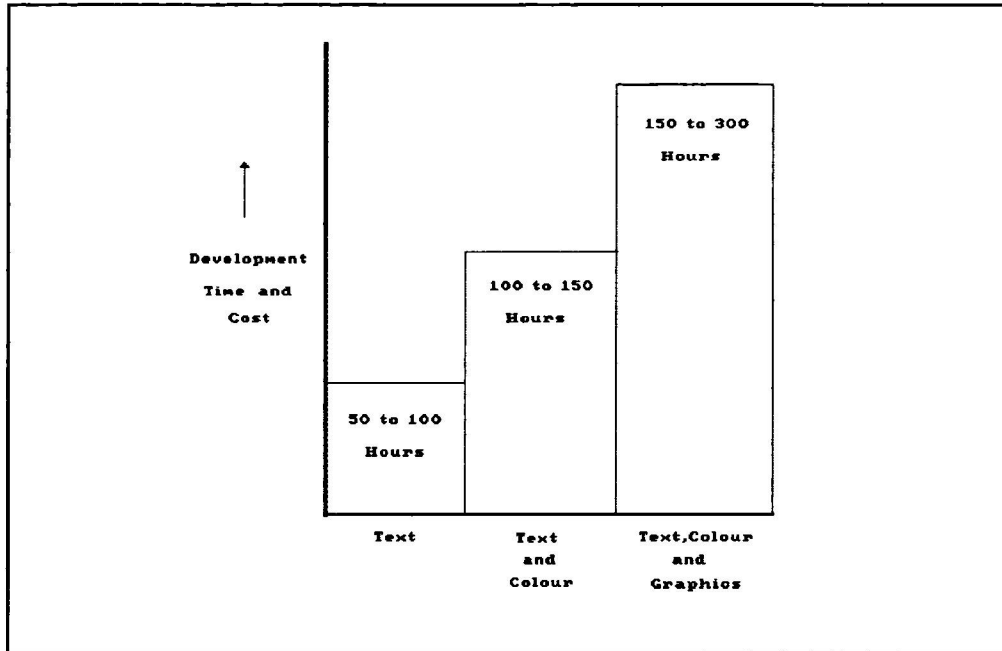


Figure 1 Diagram of Development Time/Cost against Presentation Features

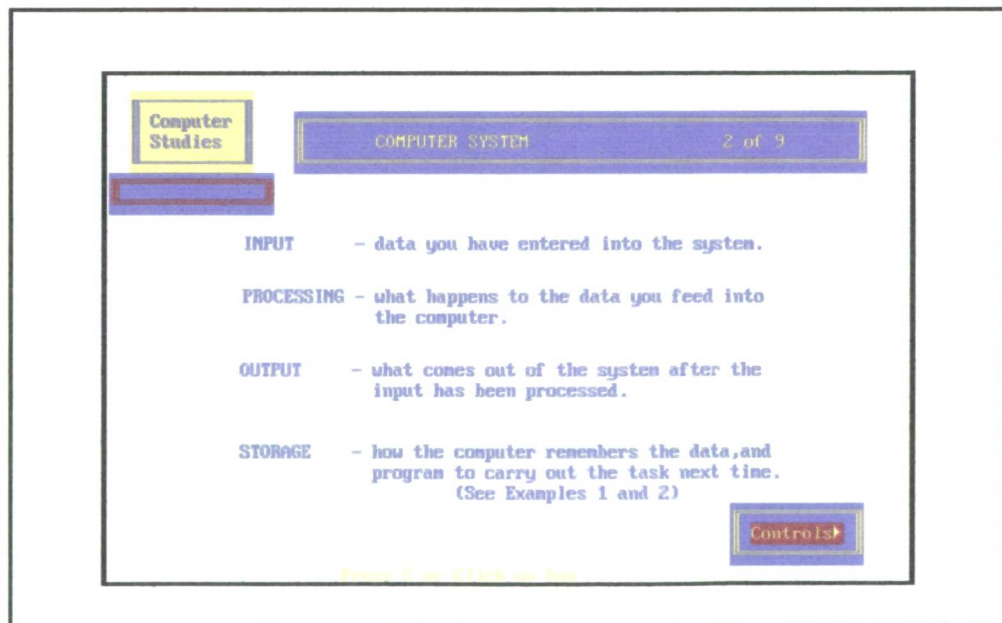


Figure 2 Illustration of a Single Frame

A frame is a single display of information; for example, a graphic picture which illustrates a point. In a computer-based package there will be many frames which provide the content of the material. A frame may be a static or dynamic display with overlaying information, scrolling text or animation.

The structure of the material is the way in which the frames are linked together. The basic structure is to link short sections of material together and to branch between the sections depending on the behaviour of the user. Figure 3 illustrates a simple branching structure. Designers are free to choose the routes available to a learner and how they are accessed.

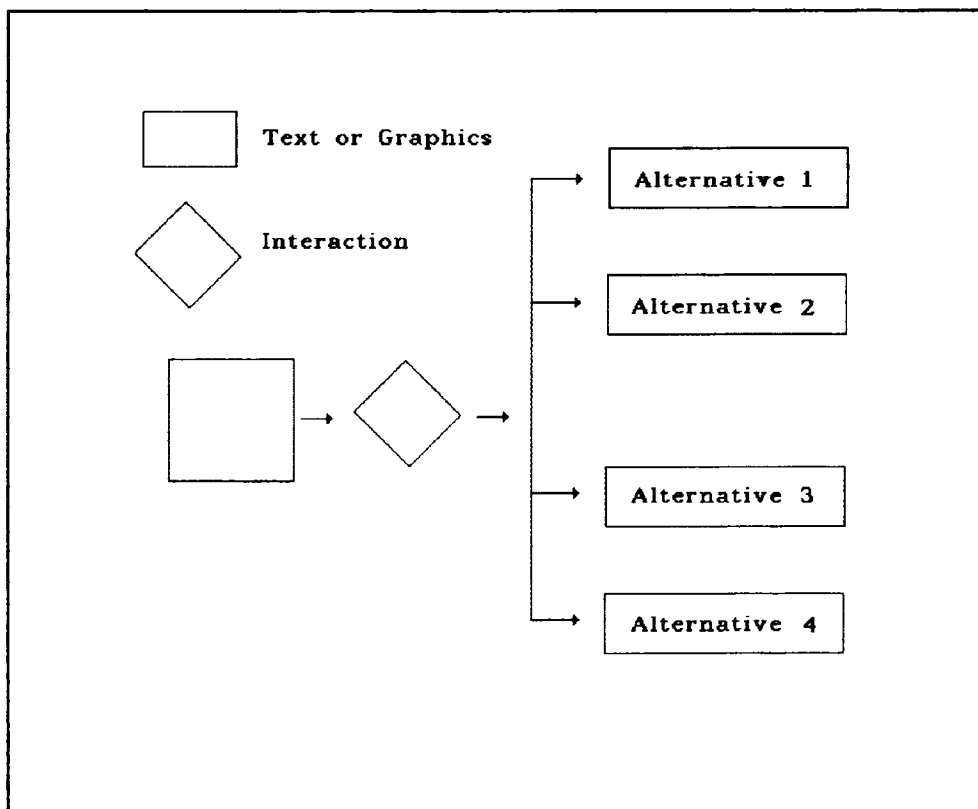


Figure 3 Simple Branching Structure

The control of the material can be provided in a number of ways such as menus, function keys, buttons, or a command language. A menu is simply a list of choices from which a user makes a selection. With function keys as a means of control, a user presses a key or combination of keys to make a choice. The keys are normally fixed in their purpose throughout the material. The use of function keys is common

in many software packages but users do need to understand their purpose. In learning materials, function keys are used in a limited way and their purpose is frequently illustrated on the screen to aid learners. A command language provides users with control of a program by entering commands which the software obeys. Command languages are a powerful means of giving control of the program to the user, but they do require users to learn them. Learners are normally new users. Expecting them to spend a long time absorbing command terms or function keys is therefore unreasonable. Menus are quicker to learn than the other control options and are therefore frequently used in learning materials (Kearsley, 1986).

1.7 Interactivity

The dialogue between the user of the learning material and the computer is known as interaction. It is shown in figure 4 as a cycle of stimulus-response-analysis-feedback.

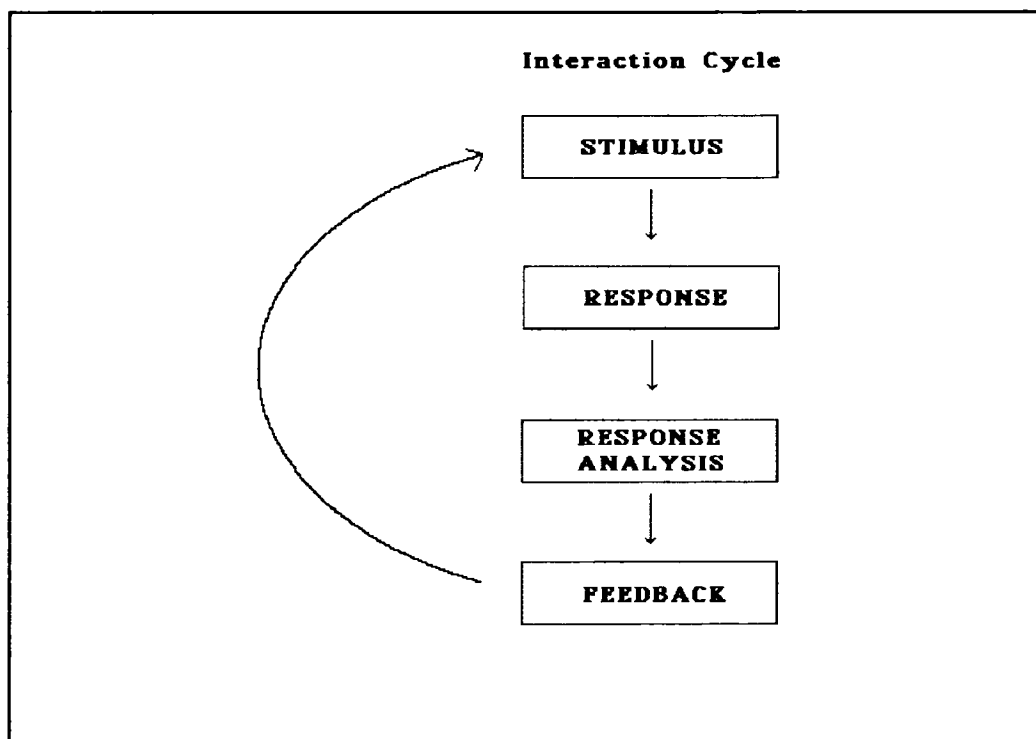


Figure 4 Interaction Cycle

The stimulus can take a variety of forms such as presenting learners with a choice, with a question or with information. The response to the stimulus is analysed by the system and a new stimulus displayed. Computer-based learning is created by developing a series of interactive cycles with the aim of assisting learners to develop knowledge and understanding.

The level of interaction varies considerably, from the passive reader of the display to the active participant where each action of the user causes the material to change. Ayerst (1987) and Gery (1987) have separately presented definitions of the various levels of interaction. Ayerst (1987) offers a general set of definitions which are shown in table 2.

Table 2 Levels of Interactivity

| Level of Interactivity | Description |
|-------------------------------|--|
| Passive | No interaction except to switch on |
| Spectator | Page turning (e.g Press Space Bar to continue) |
| Selective Spectator | Determining route through the material using menus or other options |
| Participant | As above but response given to answers e.g 'that is correct' |
| Directed Participant | As above but branching on the response |
| Fully Interactive | Full continuous interaction as in a simulation with changing parameters in response to student input |

Gery (1987) distinguishes between the type of material which is being designed. She has produced three sets of definitions for case studies, presentations and simulations. Table 3 provides her definition of different levels of interaction for

presentation type computer-based learning. Presentation type material is broadly the same as tutorial computer-based learning. This is the basis of the Colour and Graphic experiments which are considered later.

Table 3 Levels of Interaction for Presentation

| Level | Learner Outcome | Technique |
|--------------|--|---|
| Level 1. | Awareness Familiarisation Recognition | Learner selects from 2 options such as Yes/No True/False Feedback is correct answer with additional information |
| Level 2. | Understanding Knowledge Discernment Distinction Discrimination amongst alternatives | Learner selects from 3 to 5 options or fills in the blank type questions. Feedback is correct answer with additional information |
| Level 3. | Understanding Knowledge Discernment Discrimination between complex or closely related alternatives Application/Use | All the above plus extra information not previously included as part of the feedback |
| Level 4. | Analysis/Synthesis Advanced application | All above with many more branches |

Learners vary considerably in their needs, styles of learning and rate of information absorption. Heines (1984) suggests that the greater the level of interaction between the user and the program, the more the material can adapt to

meet the individual learner's differences. However, interaction can confuse and disorientate the learner (Rivlin et al, 1990). Interaction is the principle difference between computer-based learning methods and other approaches. It may stimulate learning in innovative and creative ways. It is therefore a factor that needs to be considered in the development of screen designs. Hamilton et al (1994) states that interactivity in computer-based learning is enhanced by interface design. There is therefore also a need to consider how screen design can influence interactivity.

1.8 Significance of Computer-Based Learning

Great Britain faces many challenges owing to increased economic competition from the other members of the Single European Market, traditional trading partners and the new developing nations. The success of many of these competitors has been underpinned by major investments in vocational education and training (Training Agency, 1990). New techniques in learning are increasingly contributing to the development of people by allowing them to acquire skills in ways which suit their individual needs (Employment and Education Departments and the Welsh Office, 1991). In schools, information technology is already a major factor in education with 28% of primary and 38% of secondary schools reporting in 1991 that it made a substantial contribution to learning (Department of Education and Science, 1991). The total training effort in Great Britain is large with 1.5 million person years of training undertaken in 1986/87 (Training Agency, 1989). This is equivalent to approximately 7% of the total person years worked and is valued at 33 billion pounds. Training is thus a major element in the economy. A small improvement in the efficiency of education and training is therefore likely to have a substantial cost-saving effect. Already a substantial number of learners have chosen to use more flexible approaches to learning. In 1992, 77000 students were enrolled with the Open University (Rowntree, 1992).

Computer-based learning has been identified as a key strategic weapon in the competition between corporate rivals (Stefanko et al, 1990). Abbey National (Rowntree, 1992) compared the cost of introducing computer-based learning with

conventional approaches to training. The comparison showed that the difference between computer-based learning and conventional approaches lay in the fixed and variable costs. Table 4 provides an example of the comparative costs of computer-based and conventional training from Abbey National.

Table 4 Cost Comparison - Computer-Based Learning against Conventional Approaches

| Costs | Computer-Based | Conventional |
|---------------|-----------------------|---------------------|
| Fixed Cost | £59000 | £2500 |
| Variable Cost | £6 per learner | £66 per learner |
| Break-even | 980 learners | Not applicable |

The computer-based learning had a large fixed cost due to the purchase of computers while its variable cost was small. Computer-based learning was thus cost effective if the number of learners was large enough. In the case of Abbey National, the break-even point was 980 learners.

However, this assumes the need to buy new computers in order to introduce computer-based learning. The spread of computers throughout education and commerce will result in the number of learners required to achieve a break-even point being reduced.

A survey of Training and Enterprise Councils and Industrial Training Organisations in 1993 showed that both groups anticipated an increase in the use of technology-based learning (Condor Group, 1993). Training and Enterprise Councils and Industrial Training Organisations anticipated that 72% and 59% of their companies, respectively, would increase their use of technology-based learning during the next year.

A national survey of employers in 1992 considered the use of Open and Flexible Learning in Great Britain (NOP Social and Political, 1992). All employers

were asked what types of training they provided for their staff. Table 5 shows the usage of the various types of training in descending order. Computer-based learning features as an important method with 32% using computer-based training, 12% using interactive video, 2% laser vision disks and 2% compact disks.

Table 5 Training offered to Staff

| Type of Training | Usage |
|--|--------------|
| On the job given by colleagues | 70% |
| On the job given by a supervisor | 66% |
| External course (Excluding Head Office) | 48% |
| On the job instruction given by a trained instructor | 41% |
| Text books | 39% |
| Videos | 37% |
| Computer-based training | 32% |
| Off the job training on site by trained instructors | 28% |
| External courses at Regional/Head Office | 26% |
| Training package where trainee has some control over the way in which he or she is trained | 22% |
| Audio tapes | 21% |
| Tutor support | 21% |
| Interactive video | 12% |
| Laser vision disks | 2% |
| Compact disks | 2% |
| No training provided | 21% |

The general trend is for companies to increase their use of Open Learning, including computer-based learning. The NOP survey attempted to compare their survey with earlier work in 1987 and 1989. This is shown in table 6.

Table 6 Use of Open Learning

| Use of Open Learning | 1987 | 1989 | 1992 |
|-----------------------------|-------------|-------------|-------------|
| Used Open Learning | 9% | 10% | 16% |
| Never used Open Learning | 91% | 90% | 84% |

The different sample construction of the three surveys does not allow detailed comparisons to be made but a general comparison is possible. This shows that in establishments with 11 or more employees that provide training, the use of open learning has increased from 9% in 1987 to 16% in 1992.

Computer-based learning is increasingly being used to support the education and training effort of the country. It can thus contribute to the future economic success of Great Britain.

1.9 Significance of Screen Design to Computer-Based Learning

A screen is the basic output device of a computer. It thus provides the critical interface between a learner and computer-based learning materials. If the display of learning on the screen is not effective then learning will be reduced. Screen design is therefore a critical element in the design of computer-based learning (Dean and Whitlock, 1988; Harrison, 1991).

A learner and a computer must interact in order to communicate with each other. The screen is the link between the two partners. The information displayed must aid the learning process by assisting learners to comprehend the learning material. The display must add to a learner's motivation and not reduce it. If learners find the information difficult to understand or to read because of poor contrast or an overcrowded display then computer-based learning material will fail to achieve its purpose. Screen design can therefore aid or hinder the learning process.

It is not difficult to identify a poor screen design. However, solutions are often less obvious. It is often difficult to solve interface problems without creating new ones. What is frequently required is a solution to a wide range of problems, but

such a solution is rare (Erickson, 1991). Interface design is a process of compromise. It is a multidisciplinary process combining the efforts of subject specialists, graphic artists, computer programmers and educationalists (Dean and Whitlock, 1988). Each member of a design team has individual priorities, perspectives, methods and criteria for success. It is therefore not surprising that their approaches may be in conflict with each other. So screen design faces three difficulties:

- (1) complete design solutions are rare;
- (2) design is a multidisciplinary process; and
- (3) solutions are often compromises based on trading one feature against another.

Screen design principles based on experimental results are relatively rare (Brown, 1986). General screen design principles to simplify interface design would have considerable value. Guidelines aim to assist designers to make well-informed choices, avoid mistakes and follow paths which other designers have found effective. They should also allow the occasional or inexperienced designers to improve their products (Thimbleby, 1990). This is important in the learning sector where many teachers and trainers produce the occasional item of computer-based learning.

1.10 Current Investigation

This study considered the screen design of computer-based learning materials. A number of factors and hypotheses were examined in relation to the use of colour and graphics. Two experiments were undertaken. The first experiment investigated the use of colour and the overall experimental approach while the second experiment considered the use of graphics.

Both the Colour and Graphics experiments employed a number of standard guidelines in relation to the use of text which were identified from a review of the research literature. The literature review also included an analysis of the

experimental design employed by other researchers. This provided a basis for the design of both experiments. In addition to the review of research findings, a survey of readily available computer-based learning materials was also carried out. This provided a baseline of current screen design practice against which comparisons could be made.

The general approach adopted in the experiments was to consider how screen design of computer-based learning materials encourages learners not simply to read the material but to consider it, relate it to their prior learning and to make informed choices. Computer-based learning can take a variety of forms such as a simulation, an interactive tutorial or a case study, though distinction between the different types of material is not precise. Both experiments adopted a tutorial style.

Effective learning is an active process and computer-based learning materials need to encourage learners to actively take part in the tutorial by interacting with the material. The Colour and Graphics experiments were designed to provide effective learning environments. The overall design of both experiments were not altered except to change screen design factors which were being investigated.

Colour and graphics may provide additional interest and motivation to learners. These features can be used to draw a learner's attention to key learning points or present information in an attractive and effective way. The Colour experiment was focused on two factors:

- (1) use of colour in computer-based learning; and
- (2) research methods to be employed in the experiments.

The following hypotheses were investigated in the Colour experiment:

- (A1) the addition of colour will increase the time each learner will spend on the individual frames of the tutorial;
- (A2) the addition of colour will increase the number of frames which a learner will explore during the tutorial; and

(A3) the addition of colour will increase the overall time each learner will spend on the tutorial.

The Graphics experiment was focused on the use of graphics and the complexity of the screen displays in computer-based learning. The following hypotheses were investigated in the Graphics experiment:

(B1) the time spent on each frame will increase as the complexity of the display increases;

(B2) the time spent on each frame will vary according to a learner's prior knowledge of the tutorial subject;

(B3) the route chosen by a learner through the tutorial will reflect the differences in that learner's prior knowledge of the tutorial subject;

(B4) the time spent on each frame will relate to the type and size of graphic being displayed; and

(B5) the additional modules which are cross-referenced to the tutorial will be accessed more frequently than the other additional modules.

1.11 Conclusions

For Great Britain to compete in world markets, it needs to consider the most effective methods of developing people. Great Britain's competitors have made major investments in vocational education and training in order to gain an economic advantage. Computer-based learning has the potential to contribute to the development of the workforce by meeting individual needs. In schools, there has been a substantial increase in both the availability of computers and the use of information technology in learning over the last ten years. In industry and commerce,

computer-based learning is already established in many companies.

The screen provides the critical interface between the learner and the learning material. Screen design is therefore a vital factor in computer-based learning. The value of illustrations and colour in learning materials is widely accepted. A number of guidelines have been published to assist designers. However, several researchers have concluded that much remains to be determined about how to design screens.

Chapter Two

COLOUR, GRAPHICS and TEXT

2.0 Introduction

This chapter provides a review of the research literature concerning the use of colour, graphics and text in screen design. This review includes an analysis of the design of experiments conducted in a range of media to consider if they support the transfer of their results to the design of computer-based learning materials. The final element is a survey of readily available computer-based learning materials to identify current screen design practice. The chapter thus provides an overview of research literature, an analysis of experimental design and a survey of computer-based learning materials.

2.1 Use of Colour

There have been a number of reviews of research into the use of colour in learning materials (Rudnick et al, 1973; Christ, 1975; Dwyer and Lamberski, 1983). In the main, the research has concentrated on media other than computer screens. There have been many studies into the effectiveness of colour in transparencies, slides, workbooks, film and television. The few studies which have concentrated on visual displays have been concerned with aircraft instrument displays, teletext and only occasionally with computer screens (Durrett and Stimmel, 1987). The effects of colour in both computer-based learning materials and other media will be considered since we are interested in the research into the use of colour to aid learning and comprehension. Research evidence from other media may well be transferable to computer displays.

Christ (1975) carried out a comprehensive review of 42 studies into the use of colour, published between 1952 and 1973. The review quantitatively analysed results to compare the effectiveness of colour and achromatic codes. In general terms, Christ (1975) concluded that colour is an effective aid to performance under some

specific conditions while under other conditions it can be detrimental. His results provided three main observations, which are:

- (1) when colour is used selectively to direct attention to novel material, increased learning of the material is likely to occur;
- (2) use of excessive amount of colour does not serve to direct attention and no improvement in learning is likely to be observed; and
- (3) coloured pictorial material is recalled more accurately than black and white material.

Christ's (1975) review concentrated on studies in which the subjects were mainly beginners and who were carrying out a single task. Students using computer-based learning materials are normally using the material for the first time and are usually only asked to concentrate on one aspect at a time. It should therefore be transferable to the computer-based learning environment.

Chute (1979) supports the results that show colour used in a superfluous way does not provide significantly better learning than black and white materials. Hativa and Teper (1988) compared the results of students working through a computer-based forty minute lesson on geometric concepts presented in three versions. The versions were identical in content and the only differences were provided by the use of colour. Version one was monochrome, version two used colour to provide functional cues to key items of information and the third version used colour indiscriminately. The experiments used 109 high school students randomly allocated to each of the versions. They were tested for their aptitude to geometric concepts and also for their comprehension of the lesson. The experiment showed that the students who used version two of the material performed significantly better when tested immediately after the lesson and one month later using pre- and post-tests. Colour cueing was therefore more beneficial than either the monochrome or indiscriminate use of colour. There was no significant difference between the monochrome and indiscriminate colour versions. This supports the view that using

colour in a superfluous way is not beneficial.

Several researchers have supported the view that colour used to provide cues to key learning concepts aids learning (Chute, 1979; Dwyer and Lamberski, 1983; Hannafin, 1985). Chute (1979) states that the beneficial effects of colour as a cueing device are due to colour facilitating the organisation and structuring of the learning task. Hannafin (1985) offers the explanation that colour aids learning by attracting and holding students' attention throughout the learning material. This is seen as a fundamental pre-requisite for the learner to process the information presented.

Kanner (1968) reviewed the effects of colour in learning materials presented on television and concluded that the value of colour as a cue for identifying key information decreases as the number of colours used increases. Chute (1979) offers the explanation for this effect by suggesting that colour adds more information than the learners can process. They thus miss some of the information presented to them.

Many studies have reported that students prefer coloured learning materials to black and white or monochrome media presentations. Baek and Layne (1988) reported that high school students found coloured computer-based learning material had more aesthetic value than the monochrome version of the material. Dwyer and Lamberski (1983) reviewed many research studies into the use of colour in teaching and learning to conclude that learners prefer colour to black and white versions of media presentations.

Christ (1975) did not report any connection between learners' responses to colour and their aptitude to the content of the learning material. However, other researchers have suggested that the use of colours as an attention getting and maintaining device appears to be particularly significant with less able students. Allen (1975) suggests that directing the attention of learners with a low aptitude for a subject is essential to aid their learning. Hativa and Teper (1988) reported that less able students benefited most from using computer-based learning materials which used colour to highlight key points.

2.2 Highlighting and Density of Display

A number of studies have considered the use of colour to highlight information. They have, in general terms, followed the same design of experiments in which subjects were asked to identify objects in a display of varying density comparing coloured and shaped cues. Smith (1962) asked subjects to carry out a visual search of a series of displays with varying densities using up to five colours. The experiment showed that as the density of display increased, so did the search time. If the subject knew in advance the colour used to highlight an object, the search time was significantly reduced compared with a monochrome display. However, if the subject did not know the colour of highlighting then there was no significant difference. This is an interesting finding in that it indicates the importance of a consistent use of colour in learning materials (for example red always used for key points, green always used for examples and so on).

Smith (1963) carried out a similar experiment but combined the use of colour in addition to shape coding in comparison with a black and white display. Subjects were asked to carry out a visual search and a counting task. The results again showed that as density increased so did search times. In addition, it was found that colour and shape produced a seventy-six percent reduction in errors and a sixty-five percent reduction in time for the counting task compared to the black and white display. Smith and Thomas (1964) further developed the theme of colour coding by comparing colour and shape coding. They reported that colour coding produced faster and more accurate results than shape coding. As with the previous two experiments the search times and errors increased as the density of display increased. The densities of display used in these experiments were twenty, forty, sixty, eighty and one hundred objects.

Cahill and Carter (1976) considered the use of more than the five colours used by Smith and Thomas (1964) in displays of varying densities. They found that up to ten colours could be used in low density displays (ten to thirty objects) without reducing performance but with higher density displays (forty to fifty objects) performance was reduced. The search time increased linearly with the density of the

display but showed a curvilinear relationship to the number of colours used. When only a few colours were used, search times were reduced but with increasing use of colours the search time increased.

These researchers were not primarily concerned with the design of computer-based learning materials but more with the identification of visually displayed information. There are however a number of findings which are important to designers of learning material. A summary of the findings relevant to learning materials are shown in table 7.

Table 7 Highlighting and Density of Display

1. Increasing the density of display will make the identification of information more difficult.
2. Colour is a powerful means of highlighting information compared to shape highlighting or no highlighting.
3. As the number of colours is increased so does the difficulty of identification of information.
4. There is an inverse relationship between the number of colours and density of display, that is high numbers of colours can be used but with low density displays. A smaller number of colours must be used with a more dense display in order to maintain the user's performance.
5. Consistency in the functional use of colour is important.

There are several ways of providing highlights on the computer screen other than by using colour or shape. These include inverse video, boxes and blinking. Several researchers have compared the different means of highlighting information. Coloured highlighting has been shown to reduce the search time to find information but the other devices have not produced similar results. Gamberg (1985) had subjects search for an individual target digit in a background of four distractor digits. Three highlighting devices were used (inverse video, boxing and blinking) and one standard condition. The target was highlighted fifty per cent of the time and a single

distractor the remaining time. Gamberg (1985) reported that subjects were slower on average to identify the target in all three highlighting conditions compared to the standard condition. It thus appears that highlighting with colour reduces search times while with the other three devices, search time is increased. Of course, the interesting question is why there should be a difference.

Fisher and Tan (1989) suggest that boxing and inverse video highlighting devices delay identification due to masking. The box and inverse video partially mask the digit being highlighted which slows identification compared to the digit which was not highlighted. The blinking digit is only highlighted during half the blink cycle since it is not highlighted during the off part of the blink cycle. The devices are thus delaying the location and identification of the highlighted digit. Coloured highlights appear to attract attention almost immediately. Fisher and Tan (1989) carried out two experiments to compare coloured highlights with inverse video, blinking and a no highlight condition. The results of the experiment showed that subjects could identify coloured highlighted targets more quickly than either those highlighted with blinking or inverse video. Tan (1987) in related research has shown that boxing causes similar undesirable delays in identifying a highlighted target.

2.3 Coloured Text and Highlighting

Kelley (1988) carried out an experiment to investigate the use of coloured and highlighted text and how it affected student performance. A group of students from Columbus State Community College were the subjects for the experiments. The group consisted of 51 men and 52 women with an average age of 26 and all taking an introductory computing class at the College.

The approach adopted was to randomly assign students to take one of four versions of a short computer-based training course on the use of dBASE II software. The course was taken on one of fifteen IBM personal computers with colour monitors all located in the same computer laboratory.

The first version of the course displayed all text in green. The second version added a second colour by highlighting key words and phrases in white. Two more

colours were used in version three with dBASE data and examples being shown in yellow and the status line in grey. This version conformed to the accepted norm of the developers of computer-based learning materials, using no more than four colours and using them in a consistent manner (Faiola and DeBloois, 1988). In contrast to these three versions of the material, a fourth version used fourteen colours in a random way. The four versions therefore used a single colour, two colours, four colours and fourteen colours of text with a single consistent background colour.

The researchers recorded the students' performance during and at the end of the course, as well as the time taken to complete the course. The students also completed a questionnaire to determine their attitudes to the four versions of the material. The results of the research showed that there was no significant difference between the students' performance in the four groups. The on-line scores were 63% for group 1, 60% for group 2, 62% for group 3 and 60% for group 4 while the paper test produced a narrower range from 57% for groups 2 and 4 to 55% for group 1. The times taken were similar for all groups with group 4 averaging 36 minutes and group 3, 43 minutes. The researchers concluded that neither considered (group 3) nor indiscriminate (group 4) use of colour appeared to have an immediate effect on the students' performance.

The analysis of the student questionnaires (see table 8) revealed considerable differences in their attitude to the four versions of the course, in particular, the students' reactions to whether the version caused fatigue, eye-strain or was distracting. The students' responses show that they felt that version four using fourteen colours caused fatigue, eye-strain and was distracting. In contrast, version three using four colours was not distracting and did not cause fatigue or eye-strain. This evidence should be considered with care. Paterson and Tinker (1940) and Tinker (1963) observed in their research that although some variables seemed to have only modest effects, individual preferences can be changed to a far greater extent for the material by small changes to the same variables. A change of reading speed of 5% caused by a font change could result in changes of preference for the material of 300%.

Table 8 Student Attitudes to Coloured and Highlighted Text (percentages)

| Effects | One Colour | Two Colours | Four Colours | Fourteen Colours |
|-------------------|-------------------|--------------------|---------------------|-------------------------|
| Caused Fatigue | 34 | 32 | 14 | 55 |
| Caused Eye-Strain | 32 | 27 | 20 | 53 |
| Distracted | 20 | 18 | 10 | 57 |

The results of the research, although providing useful information still leaves many questions without an answer. The jump from four to fourteen colours leaves open the effects of colour combinations five to thirteen. Version four used fourteen colours in a variety of combinations, so no information can be gathered about the use of a large number of colours in a consistent way. The background colour was the same for all four versions and the primary colour for text in all four versions was green. It is only possible to speculate on the effects of using different coloured backgrounds and texts on student performance and attitudes. There are many variables which need to be considered in determining the use of coloured text and highlighting. These experiments only covered some of them. The major conclusion which can be reached is that learners feel that courseware using many colours inconsistently causes fatigue, eye-strain and distracts them, while courseware using four colours in a consistent way does not distract or cause fatigue or eye-strain.

The colours which appear on the computer screen are a product of combining the primary colours which a monitor's electron guns produce. If the electron guns are not perfectly focused, then coloured characters will not be sharp. For this reason, several researchers have argued that text should be restricted to these primary (monochromatic) colours since only a single gun is required to produce the colour. Isaacs (1987) states that 'Monochromatic colours should be used for text since images produced are likely to be sharper than those produced by several electron guns'.

On most microcomputers the three primary colours are red, green and blue.

Text should therefore be restricted to these colours. Reynolds (1979) however states that on a black background white, yellow, cyan and green (in that order) are the most legible and that magenta, red and blue are the least legible. She bases this view on the grounds of luminance (brightness) alone. The eye is at its sharpest with white and green light and its poorest with red or blue. The combination of the evidence shows that text should be written in green when on a black background. Isaacs (1987) suggests that given a reasonably well adjusted monitor then white, yellow and cyan will probably be as easy to read on a black background.

Reynolds (1979) suggests that coloured backgrounds for text should be avoided but does not provide any evidence for this conclusion. This study provides a clear guideline for designers but ignores the vast potential of the computer screen. Isaacs (1987) puts forward a general guideline that the strongest contrast in brightness between text and background colour should be the basis for selecting colours. This simple guideline is supported by a number of authorities. Faiola and DeBloois (1988) state that 'the better the contrast, the higher the degree of readability of text on a display'. They provide a number of examples such as white or yellow text on a grey background improves readability while grey on yellow lowers the contrast and thus reduces readability. Their list of effective combinations includes white on black, white on dark grey, white on dark blue, dark blue on light grey, black on light grey and black on dark grey. Isaacs (1987) also offers some general guidance which applies equally to highlighting, text and background colour that 'hot' colours should be avoided, a hot colour being one which appears to pulsate on the screen such as pink or magenta. He reports that these colours distract and reduce readability. Colours should be chosen to complement each other and not to clash.

Coloured text can serve a variety of useful functions. The screen designer can use it to gain the attention of the learner, emphasise key points, link pieces of information and add interest to the material. Kelley (1988) has shown that the indiscriminate use of a large number of coloured texts is not welcomed by the user. The use of four colours in a systematic manner does attract the most favourable attitudes from the students.

2.4 Colour Preferences

Reid and Wicks (1988) carried out an experiment to identify the colour preferences of young children between the ages of 3 and 6 years old using a BBC computer. A total of 243 children (122 boys and 121 girls) took part in the tests which consisted of a simple perceptual task of picking out similar shapes. Five types of shape were used: squares, oblongs, triangles, kites and stars in two sizes.

The test was carried out in two stages. Initially, the children were presented with a standard screen of all ten shapes in yellow on a black background. They were then asked to identify two shapes which were the same and then the other ones. This part of the experiment aimed to determine if the shapes themselves were influencing selection. The second stage consisted of the children being shown screens in all eight background colours. The shapes were randomly positioned on each screen and also randomly coloured (each pair of shapes were presented in the same colour). The choice of shape should have only been influenced by the colour alone. The results of the experiment are shown in table 9.

The experiment identified 8 preferred, 11 unsuitable and 37 neutral colour combinations. This compares with a study in Sweden (Nilsson et al, 1983) on 20 to 30 year olds to identify preferred text/background colour combinations. The Swedish tests produced 2 preferred, 6 unsuitable and 41 neutral combinations. Apart from minor differences, the results agree with those of Reid and Wick's (1988). It does appear that the age difference in the subjects did not affect the results.

The researchers concluded that the choice of colours was not simply a matter of taste or an arbitrary choice and suggest that further studies should be undertaken. It is perhaps worth considering other evidence against these results. The commonest advice given to designers when choosing colours is to aim for the maximum contrast between background and foreground. This should produce the highest degree of legibility. Reynold (1979) suggests that contrast is the critical factor in using colours on a computer. Table 10 shows the result of comparing the Reid and Wick's (1988) results with contrast. It shows that there is no alignment between a high contrast and

the preference for a colour combination. The preferred combinations are blue/black with a contrast of 11% while amongst the unsuitable there are combinations such as black/yellow with a contrast of 89%.

Table 9 Colour Preferences of Young Children

| Colour Preferences | | | |
|---------------------------|-------------------|-------------------|-------------------|
| Preferred | | Unsuitable | |
| Background | Foreground | Background | Foreground |
| Red | Black | Yellow | Green |
| Yellow | Black | White | Green |
| Blue | Black | Black | Yellow |
| Cyan | Black | White | Yellow |
| White | Black | Magenta | Blue |
| Magenta | Yellow | White | Blue |
| Yellow | Magenta | Green | Cyan |
| Green | White | Yellow | Cyan |
| None | None | White | Cyan |
| None | None | Cyan | White |
| None | None | Green | Yellow |

Bauer and Cavonius (1980) have carried out research into the effects of reversing contrast, that is, displaying text on a dark background or on a light background. They used 23 subjects aged between 23 and 47 years old with normal eyesight. The task involved users looking back and forth between a screen and a piece of paper in order to identify nonsense words displayed on the VDU. It is therefore similar to many learning packs which use both workbooks and a computer screen.

The results suggested that both the error rate and time to complete the task were significantly reduced by using dark letters on a light background. The users were also asked for their preferences and 22 subjects reported that they preferred

dark letters on a light background. Text was easier to read and more comfortable to use, while light letters on a dark background were described as tiring and requiring more effort.

Table 10 Colour Preferences of Young Children against Contrast (100% = maximum contrast Black/White)

| Colour Preferences | | | | | |
|---------------------------|-------------------|-----------------|-------------------|-------------------|-----------------|
| Preferred | | | Unsuitable | | |
| Background | Foreground | Contrast | Background | Foreground | Contrast |
| Red | Black | 30% | Yellow | Green | 30% |
| Yellow | Black | 89% | White | Green | 41% |
| Blue | Black | 11% | Black | Yellow | 89% |
| Cyan | Black | 70% | White | Yellow | 11% |
| White | Black | 100% | Magenta | Blue | 30% |
| Magenta | Yellow | 48% | White | Blue | 89% |
| Yellow | Magenta | 48% | Green | Cyan | 11% |
| Green | White | 41% | Yellow | Cyan | 19% |
| None | None | None | White | Cyan | 30% |
| None | None | None | Cyan | White | 30% |
| None | None | None | Green | Yellow | 30% |

Reid and Wick's (1988) results partially support these results in that black/yellow (dark on light) is a preferred combination while yellow/black (light on dark) is an unsuitable combination. Five of the preferred combinations used a dark figure on a light background with black being the common foreground colour. The remaining three combinations do not fit this connection especially yellow/magenta and magenta/yellow which are both preferred. Bauer and Cavonius's (1980) results would lead to the conclusion that yellow/magenta should be preferred while magenta/yellow should be unsuitable.

2.5 Use of Graphics

Alesandrini (1987) states that the effects of computer graphics on learning and motivation are only just beginning to be explored. Only a limited number of studies have investigated the use of graphics in computer-assisted instruction. A large number of studies have compared computer-based learning with traditional methods of learning. However, few studies have tested the usefulness of particular design strategies such as the use of graphics to assist learning. Peeck (1987) considered the current state of research into the use of illustrations in processing text. He concluded that the functions and conditions of illustrations were not clear and that much of the understanding of the use of pictures was speculative.

There has been a substantial quantity of research into the effects of illustrations in text on paper-based learning materials. Levie and Lentz (1982) reviewed 155 experiments comparing the effect of illustrated text against text alone involving 7182 individuals with an age range from primary school to adults. Only a few studies have involved the effects of graphics on a computer screen. The review of paper-based research on illustrated text offers the guidelines shown in table 11.

This large volume of research into the use of illustration in learning may well be transferable to a computer screen. There are, however, differences between the media which may limit the transfer process. Computer graphics are frequently dynamically drawn and interactively displayed in contrast to the static images upon which much of this research is based. Almost all the research evidence is concerned with how illustrations assist the learning of the associated text (Levin and Lesgold, 1978; Levin et al, 1987). This is a reasonable emphasis when dealing with media such as paper which is text-intensive and illustrations are easily inserted into the flow of words. A computer screen is not a text-intensive medium and graphic images will frequently appear on their own or with limited text.

Table 11 Guidelines on the Use of Pictures with Text

1. Pictures not covered by the information in the text are not likely to enhance learning of the text.
2. The presence of pictures relevant to the text are likely to assist learning.
3. The presence of pictures in the text will not aid the learning of the text which is not illustrated.
4. Pictures can help learners understand what they read and also to remember it.
5. Pictures can sometimes be used as substitutes for words or as providers of non-verbal information.
6. Learners may fail to make full use of complex illustrations.
7. Learners prefer illustrated text compared to text on its own.
8. Pictures may assist learners with poor verbal skills more than those with good verbal skills.

Several researchers have compared the two media: computer screen and paper. Bork (1983) saw four significant differences between them. These are summarised in table 12.

Duchastel (1988) categorised the differences between paper and screen into two types: those connected to visual effects and those linked to the way a learner could interact with the two media. These differences are by no means separate since to design the visual effect of paper or screen, it is critical to take into account how a learner will react and interact with the display. The two authors are broadly in agreement and Clarke (1989) in a review of the research evidence included and extended this analysis. Table 13 provides a comparison of the two media.

Table 12 Comparison of Paper and Screen

1. **Print quality:** Usually good in a textbook and poor on a computer screen.
2. **Space:** Expensive as print in a textbook but space is free on a screen.
3. **Display:** On a page it is static while it may be dynamic on a screen.
4. **Choice:** User is free to choose how to use a textbook but is often prompted when using a computer.

Table 13 Comparing the Characteristics of Screen and Paper

| Attribute | Paper | Computer Screen |
|------------------|-------------------------------|--|
| Density of text | Intense | Non-intense |
| Print Quality | Excellent | Poor but improving |
| Cost of space | Expensive | Free |
| Page size | Fixed | Flexible |
| Layout | Linear | Non-linear |
| Learning style | Passive | Interactive |
| Learners | High aptitude learner | Adaptable to high and low aptitude learners |
| Comprehension | Global comprehension problems | Devices to aid comprehension if used in moderation |
| Display | Static monochrome display | Dynamic colour display |
| Illustrations | Static | Animation and simulation |

It is extremely important to understand the nature of the two media in order to consider if research evidence can be transferred from paper to screen.

2.6 Survey into the Use of Graphics in Computer-Based Learning Materials

Instructional graphics have been classified on the basis of how they convey meaning into three types: representational, analogical and logical (Knowlton, 1966; Alesandrini, 1984; 1987). Alesandrini (1985) carried out a survey of American commercial educational courseware to consider the use of graphics in the material. She did not try to assess the quality of the images, only their use in the learning material. The survey covered four subject areas: mathematics, science, language and social studies. By far the commonest type of graphics used was the representational; none of the materials included analogical images. Logical graphics were extensively used in mathematics courseware (45%). However, in contrast, social studies materials did not use any of them. Only 14% of social studies material used any graphics whereas 65% of the mathematics courseware included them. In the science material where there are many opportunities to illustrate ideas with pictures, only a minority of material (43%) included illustrations. Table 14 shows the results of the survey.

Table 14 Graphics in Commercial Courseware

| Subjects | Mathematics | Science | Language | Social Studies |
|----------------------------------|--------------------|----------------|-----------------|-----------------------|
| Packages Surveyed | 20 | 14 | 12 | 14 |
| Amount of Graphics (few or none) | 35% | 57% | 67% | 86% |
| Amount of Graphics (some) | 25% | 36% | 25% | 7% |
| Amount of Graphics (many) | 40% | 7% | 8% | 7% |
| Representational | 40% | 71% | 25% | 21% |
| Analogical | 0% | 0% | 0% | 0% |
| Logical | 45% | 21% | 8% | 0% |

Overall the results of the survey (table 14) show a low use of graphics. The extensive use of representational pictures and total absence of analogical graphics seems to indicate a low level of understanding of the use of illustrations amongst designers. Although it is difficult to tell, the results may point to the view that designers see graphics as an optional extra and not a central feature of their design.

2.7 Representational Graphics

Representational graphics are those pictures which show a physical resemblance to the object or concept they are portraying. Research has mainly considered the use of illustrations to assist the learning of text associated with a picture. It has clearly shown that pictures facilitate the learning of the text actually illustrated. The pictures do not assist the learning of text which is not illustrated (Levie and Lentz, 1982; Anglin, 1986). Levin (1989), commenting on a generation of research into the effects of pictures in prose states that: 'Pictures interact with text to produce levels of comprehension and memory that can exceed what is produced by text alone'. He is also clear that the research has not provided the answers as to why or how illustrations aid learning from text.

There have been a range of studies to consider what kind of information may gain most from being illustrated. Haring and Fry (1979) have indicated from a study with children that pictures may assist learning best when they represent the main ideas in a passage. However, there is limited supporting evidence for this assertion. A range of researchers have suggested that pictures aid learning when they show spatial structural relationships in the text (Dwyer, 1970; Levie and Lentz, 1982). Another group of studies indicate that the effects of illustrations may depend on the difficulty and abstract nature of the content of the text they are illustrating (Duchastel, 1980; Dean and Enerok, 1983; Moore and Skinner, 1985). In other words, illustrations aid learning if the text is difficult to learn, or abstract rather than concrete.

The research evidence to indicate if there are benefits from particular positions for the pictures in text is slight. Koran and Koran (1980) suggest that a

picture at the end of a section of text may stimulate selective review and reorganisation of the previous processed information. Whalley and Fleming (1975) investigated the position of pictures in text using logical illustrations. They conclude that diagrams are studied longer if positioned immediately after the sentence referring to it. This would provide some support for Koran and Koran's (1980) findings but evidence is still slight.

2.8 Degree of Realism in Pictures

Graphics used in computer-based learning materials can vary considerably in realism. A photographic image of the object to be displayed can be scanned into the material, a detailed picture provided for learners or a simple line drawing used. The graphic images may be coloured or shown in monochrome. These variables may have a considerable effect on learners but research evidence is not easy to compare since there is little consistency between the pictures used. Studies investigating the use of illustrations in text have used pictures which vary in type, number, colour, size and density of information presented (Willows, 1980).

Dwyer (1970; 1978) and Dwyer and Joseph (1984) carried out a series of experiments (over 100 studies with more than 40000 adult learners) to consider the effects of realism. Four different pictures were produced to accompany a 2000 word passage on the human heart. The four pictures were:

- (1) a simple line drawing of the heart;
- (2) a detailed shaded drawing of the heart;
- (3) a photographic model of the heart; and
- (4) a photograph of heart specimens.

The images were either in black and white or colour and were designed to show the

same information only varying in the amount of detail they contained. The results of these studies (Dwyer, 1970; 1978; Dwyer and Joseph, 1984) show that the effectiveness of realism and complexity is dependent on the time and effort the reader is prepared and able to make to study the image. If the viewing time is short (externally controlled) then the line drawing containing the least degree of realism was most effective. However, when learners work at their own pace they take advantage of the additional information provided in the more realistic images. There is, however, some risk that the extra realism may distract attention from relevant and important learning points. Holliday and Thursby (1977) added to this basic view that the effectiveness of different types of picture depends on the ability of the learner to deal with the degree of detail contained in them. Dwyer and Joseph (1984) reported that to maximise student achievement, the most effective illustration would perhaps be a hybrid picture combining photographic detail and line drawing simplicity.

Two other studies have supported Dwyer's findings on the value of simple line drawings. Both Gorman (1973) and Borg and Schuller (1979) reported that the addition of detail to line drawings did not facilitate learning. However, Holliday (1975) and Rigney and Lutz (1976) reported that line drawing assisted students who were using computer-assisted instruction to learn science concepts. Other researchers have reported that learners prefer viewing pictures that contain more detail than a simple line drawing (Travers and Alvarado, 1970; Myatt and Carter, 1979).

The final point in considering the realism of graphics is the learning objective an author is aiming to achieve. A realistic picture may well gain the attention and interest of learners more easily than a basic drawing. Spaulding (1955) reported that students appreciated more realistic and detailed pictures. Thus, if an author is seeking to motivate students then a photograph may serve better than a line drawing. A simplified drawing may provide more effective learning when an author is seeking to explain factors hidden in a complex illustration or with learners who have a low aptitude for the subject. A realistic picture will aid this understanding.

The degree of realism is thus dependent on:

- (1) the time and effort a learner is willing and able to use;
- (2) the ability of learners to cope with the detail provided; and
- (3) the purpose of the illustration in the material.

2.9 Descriptive and Instructional Captions

Bernard (1990) considered two strategies for improving the effectiveness of illustrations in text. These were descriptive captions in which the text repeated the contents of the picture and instructive captions where the key features of the illustrations were highlighted. The content of the pictures did not overlap with the written content of the passage. It was thus very similar to a single screen display with a graphic image and some related text.

The subjects in the experiment were 143 student nurses in Western Ontario aged between 18 and 21. The experiment took place in a classroom. The descriptive captions consisted of a short paragraph explaining the main features of the illustration with the aim of duplicating as much of the information contained in the picture as possible. This type of caption was always placed after the picture. The instructive caption consisted of a set of directions on how best to learn from the illustration and was always placed before the picture. One group of students received both descriptive and instructive captions, two groups received either one or the other of the captions and a last group received a version without any captions.

After the learning material had been used the students completed a multiple-choice questionnaire, a recall test and a reading ability test. The results showed that both types of caption produced better results on the recall and comprehension tests than the illustration alone. There was no additive effect when both were used and the use of captions did not interfere with the recall of non-illustrated material. The difference between the two types of caption on recall and comprehension tests was

not large. However, the instructive caption did require far less space to use than the descriptive caption. This is important when using pictures on a single screen display where space will be limited.

2.10 Analogical Graphics

Analogical graphics aim to show a topic or a concept by showing something similar and implying a similarity. In order for an analogical graphic to be useful, learners must recognise or be able to comprehend the object used in the analogy. Analogical pictures should assist learners interpret the new information in the light of their prior knowledge. Several studies have shown that verbally presented analogies assist learning (Royer and Cable, 1975; 1976; Mayer, 1980).

There have been several studies to investigate the transfer of information gained from a picture analogy to understanding subsequent abstract information. Mayer (1975) provided adults who were learning to program either with instructions using a diagram of a computer labelled in familiar terms or with instructions without a diagram. The diagram used familiar objects as analogies for computer components. For example, input was likened to a ticket window while output took the form of a message pad. The learners who were provided with the diagram were found to be better at interpreting programs. Learners who did not see the diagram were found to be better at generating programs. In two subsequent studies, the analogy was found to be particularly helpful to low ability learners. Davidson (1976) reported that college students provided with an analogical image made fewer errors in identifying assertions based on a text. He asked students to read 'The Mat Maker', a chapter of Moby Dick with or without an analogical picture. The picture showed a sword loom which compared parts of the image with the abstract concepts in the text of the chapter.

2.11 Logical Graphics

Logical graphics do not resemble the physical things they represent but are a logical representation of them. For computer-based learning, these types of graphics are important since they have a key benefit of efficiently using space while providing detailed information. They are therefore a means of displaying information on a single screen which, if text alone was used, would require many screens of text to supply the same information. For a non text-intensive medium such as a computer screen, this is an important advantage. There are many studies which have considered the use of logical graphics. However, the vast majority do not use a computer screen but are paper-based.

Winn (1987) reviewed the research evidence for the use of charts, graphs and diagrams in educational materials. He concluded that the instructional effectiveness of graphics depends on many factors. One key factor is the ability of the student. In general terms, a student's ability determines the effectiveness of logical graphics. A more able student needs logical graphics less than a less able student. There are several examples of less able students being assisted by the presence of a diagram while more able students have not shown any improvement. However, there is evidence of lower ability learners being hindered by the need to process the extra information provided by the graphic.

Another factor which affects the usefulness of graphics is the nature of the task a student is being asked to carry out. Logical graphics appear to be most useful when learners have to study the relationship of the elements of the information and its pattern and structure.

Diagrams show the organisation and structure of the key points of a subject. Winn and Holliday (1981) reviewed research on learning from diagrams and concluded that diagrams aid learning by showing the connections between concepts. This helps learners to discriminate and to generalise by replacing words with arrows and lines. Holliday (1975) found that High School students learned 30% more from a science text passage when given a relevant diagram than from text alone. He compared text, text plus diagram and diagram alone and concluded that the diagram

on its own was the best. Holliday (1975) concluded that the diagrams represented a formidable decoding task for learners and if presented with text and a diagram, they concentrated on the text and ignored the diagram. Diagrams can aid learning but only if the information contained in them is actually decoded. Supporting text may well hinder the extraction of information from the diagram.

Dean and Kulhavy (1981) studied the effects of maps on learning. A group of students was given 15 minutes to read a 2190 word passage about an imaginary African tribe. Before reading the passage the students were either given a relevant map of the area, a relevant map plus instruction to ensure the information in the map was decoded, or a map of Africa unrelated to the passage. The result showed that students with the relevant map learned only slightly more than those with the map of Africa. However, the students who had instruction (to ensure the processing of the map's information) learned 60% more. The presence of an illustration is not in itself sufficient. The information that it contains must be processed by the learner.

Winn (1987) concluded that the research into logical graphics was uneven in its emphasis. Few studies have considered the use of graphs, while the main effort has been in the use of diagrams and charts. There are several examples of studies simply considering the presence of graphics without looking at the mechanisms for their effective use. He feels that future research should investigate the learning strategies that students employ when using logical graphics.

2.12 Ability of Learners

Holliday, Brunner and Donais (1977) studied the effects of flow diagrams to teach biological cycles to students. Two types of complex diagram were used to show carbon dioxide, oxygen, nitrogen and water cycles and how they interact with each other.

The diagrams were identical except that:

- (1) first diagram showed each element as a line drawing and as a verbal label (picture-word); and
- (2) second diagram labelled each element only (block-word).

Students were asked to study the diagrams and complete a multi-choice questionnaire to test their retention of the material. They were also tested to establish their verbal ability. The results showed that students studying the picture-word diagram performed significantly better than those who studied the block-word diagram. Students with low verbal ability performed significantly better with the picture-word diagram than with the block-word diagram.

The researchers explained the results in terms of dual coding theory (Paivio 1971; 1975) which suggests that pictures are encoded both as a visual image and as a verbal label while text alone is encoded only verbally. The picture-word diagram thus provides those learners with poor verbal skills with an alternative source of information to overcome their lack of skill.

Moyer et al (1984) reported similar findings from their experiments in which students were asked to study one of three descriptions of a mathematical problem. The treatments were a verbal description of the problem, verbal description with illustrations of the elements of the problem and a short terse description of the problem. Students who used the verbal and picture combination achieved significantly greater success in solving the problem than the other two groups. In particular, the verbal-picture description was most helpful to low ability students. These two studies suggest that students with low verbal abilities will benefit from the use of diagrams containing illustrations and labels.

In contrast to these findings Parkhurst and Dwyer (1983) provided undergraduates with four versions of a tutorial on the heart consisting of a picture and text. The four versions differed in that the pictures varied in the degree of realism from a simple line drawing to a photograph. They reported that increasing realism was linked to improved performance for learners with high intelligence.

However, the performance of low intelligent users decreased with increasing realism. This can be explained by Allen's (1975) suggestion that low ability students have difficulty processing instruction which is rich in detail.

The evidence suggests that learners with poor verbal skills will be assisted by diagrams combining graphic pictures with verbal labels. There is however a limit in that the same learners may find it difficult to process images that are too rich in detail.

2.13 Nature of the Task

Kulhavy et al (1983) asked students to study three types of map showing a plan of a fictitious town. The maps had:

- (1) features identified by verbal labels only;
- (2) features identified by realistic drawings; and
- (3) features identified by arbitrary geometric symbols.

The students were tested by asking them to fill in the features of a blank map and to state which features were nearest or farthest from a given point. In both tests, the groups who studied maps 1 and 2 out-performed the group that studied map 3. The researchers explained this result by suggesting that symbols which are not congruent with what they represent interfere with learning. The results also indicate that realistic labels are more effective in recall and recognition tasks.

This interference with learning caused by the use of arbitrary symbols may not be true for all tasks. Winn (1986) reported that symbols might well interfere with the recall of elements in a pattern but not if the elements were in a sequence. His experiment involved a group of high school students studying two circuit diagrams. In the first circuit diagram, each element was represented by a conventional symbol and a letter. The letters were cross-referenced to the component's name at the

bottom of the page. In the second circuit diagram, each element was represented by a letter in a small box. The components were also numbered in order across the page from the top left to the bottom right corner.

The students were asked to draw the circuit diagrams and recall and list the components in the correct order. The group who saw the diagram using symbols did best at recalling the components in their correct order. The group who had seen the diagram without the symbols did best at drawing the diagram with the components in the correct places. Both these tasks require different types of cognitive processing. Drawing the diagram requires simultaneous processing while listing components requires successive processing.

Simultaneous processing requires the learner to grasp the whole picture; this becomes harder the more elaborate each element of the picture is. Successive processing requires the learner to remember each element separately in relation to the other elements.

The nature of the task which a learner is undertaking obviously affects the use of graphics. It should not be treated in isolation since earlier evidence also suggests that the ability of the learner is also a key factor. Increasing the detail of an element in a diagram may well aid in some tasks; however, too much detail may restrict some users. There is, however, too little evidence on the relationship between ability and task to provide reliable guidelines.

2.14 Comparing Text, Colour and Graphic Displays

Tullis (1981) carried out a comparison of using text, colour and graphic information displays. All the displays portrayed the same information with the only variables being colour, structure and graphics. The information was part of a computer-based diagnostic system for telephone engineers to find faults on telephone lines. The aim of the experiment was to determine the best way of showing the information to the engineers.

Four types of display were compared:

- (1) Narrative - a text only display;
- (2) Structured - key information was displayed in a box at the top of the screen and all the categories of information were logically separated;
- (3) Black and White Graphics - a monochrome graphic display using shape coding to represent values; and
- (4) Colour Graphics - colour replaced the shape coding but everything else remained the same.

All the subjects in the experiments were experienced employees of the telephone company who were used to operating computers (minimum experience of 9 months). The employees were trained on the fault diagnosis system using ten units of instruction. The first six were paper-based and the last four represented the four test formats. The order of the four computer displays was balanced across the group. After each unit the trainees were tested using multiple-choice questions shown on the computer screen. The total instruction time was seven hours and at the end, the subjects were asked to complete a questionnaire comparing the four displays. To study the effects of practice the test for each format was given in two sessions. The first test was administered immediately after the training for an individual format, the second after the subject had already been trained and tested at least once on all the other formats.

The results of the experiment showed that there was no significant difference in accuracy between the formats. However, the response times were significantly shorter for both the graphic displays (nearly twice as fast) compared to the narrative display in both of the sessions. The structured text showed similar results to the narrative in the first session but response times improved in the second session to equal those of the graphical displays. The subjects were asked for their preferences

and seven out of the eight trainees preferred the colour display while the monochrome format was equally preferred with the structured text.

The main conclusions are that graphics can enhance the performance of users and in particular those users who have been able to practice. The users clearly prefer colour to black and white although performance remains similar even after practice. The most interesting result is that structured text can produce similar results to colour for practiced users and that users like it as much as monochrome graphics. The method employed in structuring the text is given in table 15 and response times in table 16.

Table 15 Structuring Text

| Step | Structuring Method |
|-------------|--|
| 1 | Key information is presented in a prominent position |
| 2 | Related information is grouped together and separated from other types of data |
| 3 | Information is presented in a fixed tabular form |
| 4 | Information is presented concisely |

Table 16 Mean Response Times

| Display | Session 1 (seconds) | Session 2 (seconds) |
|--------------------------|--------------------------------|--------------------------------|
| Narrative | 9.5 | 8.2 |
| Structure | 8.3 | 5.0 |
| Black and White Graphics | 7.0 | 5.1 |
| Coloured Graphics | 6.5 | 5.0 |

2.15 Structural or Graphical Organisers

Organisers are similar to diagrams in that they are schematic representations of the relationship between concepts. They are a way of highlighting the key points and activating a learner's relevant prior knowledge. Tullis's (1981) research (discussed earlier) compared the use of three different forms of organised display using structured text, monochrome graphics and coloured graphics and showed all forms were better than text alone.

Patrick and Fitzgibbon (1988) studied the effects of presenting a structured display on the assimilation of information while undertaking a learning task, in this case, how to use a computer-based editor. The structured display was a diagrammatic representation of the major functional states or computer displays involved in the editing task and the routes between them. The experiment involved presenting the organiser to the learner at different points in the learning of the task. Three conditions were studied:

- (1) organiser presented for 2.5 minutes prior to taking the computer-based training modules;
- (2) organiser presented for 2.5 minutes after taking the computer-based training modules; and
- (3) only computer-based modules used, no organiser.

After each condition, the subjects performed the computer-based editing task. The learning modules consisted of four sections each relating to the four functional states/displays of the editing process. The material contained an overview, simulations of the displays, tests and summaries of the learning.

30 subjects took part in the experiment. None of these were computer users and none had any experience of this particular computer system. 10 subjects performed under each of the three conditions and their performance was measured.

The measurements taken were accuracy, speed and errors committed by the users while carrying out the editing task. Table 17 shows the results of the research.

The analysis of the research findings showed that 9 subjects completed the editing task using the prior display condition while only 5 each completed the task under both the other conditions. The editing was divided into 17 sub-tasks and in the prior state, a mean of 15.9 tasks was completed compared to 14.1 in the post display state and 11.0 in no display condition. The differences just failed significance at the 5% level for completing the whole task but showed a significant difference comparing prior display to the no display conditions considering the sub-tasks.

The errors committed during the editing task showed not only a significant difference between prior and post states but also between post and no display conditions. In the prior display condition, a mean of 2.7 errors were made while 7.7 and 11 errors were made in post and no display conditions, respectively. This shows that the structured display used prior to using the learning material significantly reduces errors.

Table 17 Use of Structural Displays

| Subjects | Prior Display | Post Display | No Display |
|--|----------------------|---------------------|-------------------|
| Subjects completing all tasks | 9 | 5 | 5 |
| Sub-tasks completed (mean) | 15.9 | 14.1 | 11.0 |
| Errors committed | 2.7 | 7.7 | 11.0 |
| Speed (seconds) All subjects (mean) | 913.07 | 1336.2 | 1778.02 |
| Speed (seconds) Subjects completing all tasks (mean) | 875.9 | 1051.19 | 1469.34 |

Finally, subjects in the prior display condition were significantly faster than the post display subjects in completing all the editing tasks. The post display state subjects were also significantly quicker than the subjects in the no display condition.

If only the speeds of the subjects who completed all the tasks are considered, then the prior display condition is significantly faster than the no display condition.

The experiment showed that the use of a structured display prior to studying the learning material resulted in less errors, more successful completions of the task and a quicker performance. The researchers offered two possible explanations for the results. Firstly, that the structured display provides learners with a conceptual framework which assists the organisation and association of subsequent information which has to be assimilated. This is similar to Ausubel's (1968) subsumption theory; the display can be regarded as an advanced organiser which provides learners with a relevant learning framework. Secondly, that the structural display itself provides learning material which aids learners in assimilating display related material faster and allows them to concentrate on the new material in the learning modules. The display related to 30% of the contents of the learning modules and did not contain any material not in the modules. However, the researchers were unable to suggest which explanation was correct.

2.16 Charts and Tables

Charts and tables of information convey meaning by organising categories of concepts into rows and columns. The row and column headings combined with the separation of divisions by vertical and horizontal lines means that they are dependent on spatial organisation. The meaning of a chart or table is a combination of the distance separating different elements and the organisation of the different rows and columns.

Bruce and Foster (1982) considered the use of a variety of formats for presenting a table of information on a videotext screen. The researchers aimed to test out the results of a variety of other workers who had suggested a number of layout adjustments to improve the efficiency of tables. Tinker (1960) had come to the conclusion that the addition of extra space between blocks of rows aided efficiency. This was supported by Sutherland (1980) who noticed that inserting a space every 3 or 5 lines aided the use of monochrome tables. Wright and Fox (1970) had come to the conclusion that using colour to differentiate columns in a currency

conversion table hindered the user. They had suggested this was due to poor legibility of the colour combination used (black on red). The experiment used four colour formats:

- (1) monochrome (all characters in yellow);
- (2) row alternation (alternating the colours in successive rows of yellow and green);
- (3) row banding (alternating the colour of successive blocks of three rows of yellow and green); and
- (4) column coding (colour coding of the table columns - first column cyan with alternating columns of yellow and green with a white last column)

In addition, a blank row was inserted between successive blocks of three rows, which the researchers called spatial blocking, to bring the total number of formats to eight. The table used consisted of 6 columns and 13 rows, with the first column containing a name of a town and the other columns numbers. A blank space was left between each column.

The subjects were all students at Manchester Polytechnic aged between 18 and 22 with normal colour vision. A total of 96 students took part in the experiment and they were randomly allocated to the different formats. Before the test, the subjects were given a practice session using printed material and allowed to read the instructions for the test. Only after this period was the test run. The users were asked to find items of information from the table and the time taken to locate the item was measured.

The results of the experiment suggested that:

- (1) the colour coding of columns slowed down the users performance; and

- (2) row alternation and banding were not significantly different from the monochrome layout - although they did differ from each other.

Row banding was found to be better than row alternating. The insertion of a blank line had a negligible effect on performance despite the reports of other researchers. The overall conclusion was that tables should be presented in monochrome or using row banding and that colour coding of columns should be avoided.

It is doubtful if the results of this experiment could be transferred across a wide range of tables presented on computer screens. The experiment was limited to a single table of a mixture of words and numbers. In the vast array of possible layouts of tables such as all numbers, all words, information presented on a grid and the range of sizes of tables, this is a very small sample. However, the experiment does indicate two points: that research based on printed materials may not be transferable to a screen and that colour does not automatically improve design.

2.17 Icons and Hotspots

One of the commonest control features of computer-based learning materials is the use of icons or hotspots. These are small areas of the screen which can provide links with other parts of the learning material or provide access to additional functions. A user clicks on the icon or hotspot with the mouse pointer. Icons and hotspots are usually of three types: a word or phrase, a picture, or a mixture of words and pictures.

Guastello, Traut and Korienek (1989) considered the meaningfulness of the three types of icon or hotspot to users; that is, allowing learners to make the correct choices of route through the material. Learning will not be enhanced if learners are not clear of the meaning of an icon or hotspot. They may find themselves in the wrong part of the material and perhaps be unable to find their way back to the correct place.

The researchers tested four hypotheses concerned with the effectiveness of the different types of icon and hotspot.

These were:

- (1) pictorial icons are more meaningful than word hotspots;
- (2) meaningfulness would be dependent on factors such as long versus short abbreviations and industry standards versus pictures;
- (3) meaningfulness would be dependent on the familiarity of the users with the content of the material; and
- (4) icons consisting of a mixture of words and pictures would be more meaningful than icons or hotspots using only words or pictures.

Two experiments were conducted. The first used icons/hotspots for objects found in the building automation system environment and were rated by 187 system operators. The second experiment used icons/hotspots from engineering, computer systems and finance environments and were used by 139 undergraduates.

The results of the experiments showed that mixed word and picture icons/hotspots are more meaningful than either word or picture icons/hotspots on their own. This was apparent across a wide range of domains. Other findings include a preference amongst users for the use of long rather than short abbreviations and that it is possible to develop icons/hotspots which are more meaningful than the industrial standard symbols.

The researchers explained their findings in terms of the parallel processing theory of cognition (Egeth, Jonides and Wall, 1972). This theory considers that a human receiving a stimulus of mixed verbal and pictorial content divides the stimulus between the two cerebral hemispheres. One hemisphere contains the verbal processing centre and the other the spatial processing centre. An individual may be skilled in verbal or spatial processing. When the mixed stimulus is processed both centres work simultaneously and the centre with the best developed set of connections finishes first. The use of mixed word and picture icons thus allows the faster mechanism to be brought into use every time.

2.18 Animation

Computer screens are not simply a means of displaying static pictures - they are a dynamic medium. A screen can show animated pictures and video. Even at a simple level, designers can control the clearing of screens by wiping the display from the top, bottom, middle, left or right of the display. A graphic image can be scrolled, granulated or moved around the screen. This environment is therefore a rich creative one for authors of learning material. These functions can be used to create interesting and motivating material for a learner.

Alesandrini and Rigney (1981) asked students to review learning material by either rereading the material or by interacting with a simulation using animated graphics. The simulation group performed better in a pictorial post-test than those who reviewed the material by reading it again. The simulation group also had a more favourable attitude to the subject than the other group. However, there was no difference between the two groups when they were given a verbal post-test.

Baek and Layne (1988) investigated the use of colour, graphics and animation in a computer-assisted learning tutorial on mathematics. High school students were randomly assigned to one of six groups. Each group used a different version of the tutorial. These consisted of two black and white, two colour, a graphics and an animation tutorial. The material was identical except for the addition of graphics, colour and animation. Each subject carried out a twenty minute lesson on mathematics. The subjects were asked to complete a pre- and post-test using pencil and paper test and the time spent on the tutorial was also measured. 119 students took part in the experiment of whom 90% were computer users.

The experimental results showed that the students' performance improved with the addition of graphics and animation. The graphics group achieved a higher score than the text-only group while the animation group resulted in the highest of all. The effect of colour was not detected except that the students reported that the coloured versions had more aesthetic value than the black and white lessons. Several studies have shown positive effects of animation in learning materials (Reed, 1985; Tritz, 1986; Zavotka, 1987; Rieber, 1990; 1991).

Avons et al (1983) carried out two experiments with 9 to 11 year old children to introduce them to journey graphs using a computer simulation. The children were shown a display of a moving car with a graph which plotted the car's position at regular intervals. The learners were encouraged to discover the relationship between the car's movement and the graph. The children controlled the speed of the car by using a sliding control with five settings.

The display initially performed eight journeys showing only the simulation of the car and then for the next nine trips a graph was drawn alongside the simulation. The link between the simulation and the graph was controlled by fixed rules which the researchers hoped the children would learn. A post-test of the users showed they had not learned the rules. This may have been caused by several factors such as an insufficient number and types of example, or the display being too complex to enable the learning of the underlying principles.

The children felt that the key feature of the display was the speed of the simulation. They related the speed of the car to the slope of the graph but did not proceed to discover the other information contained in the display. The researchers concluded that there were two reasons for this result. Firstly, that the children stopped looking for other relationships once they had discovered the speed and slope. Second, that the learners' concentration on the simulation and dynamic display interfered with their ability to perceive and explore detailed relationships. Other studies have also reported that animation has little or no additional learning value compared to text or still graphics (Rieber and Hannafin, 1988; Rieber, 1989; Rieber, Boyce and Assad, 1990).

The evidence is mixed for the effectiveness of animation in aiding learning and indicates that it has potential but with the warning that one should not assume effectiveness in all circumstances. The indiscriminate use of animation has been reported as being distracting for learners (Tritz, 1986). The research findings do provide a reasonable basis for a series of guidelines for the use of animation in learning.

The main points are:

- (1) use simple and short sequences of animation (Rieber, 1989; 1991);
- (2) provide a verbal narration to accompany the animation (Mayer and Anderson, 1991);
- (3) emphasise key points with animated prompts and other aids (Rieber, 1989; 1991);
- (4) varying the speed of the animation can provide emphasis (eg slow, frame by frame or in reverse under the learner or program's control) (Eaker and Jacols, 1982);
- (5) animation is appropriate when the content involves motion or related issues (Rieber, 1989; 1991);
- (6) animation should be restricted to explaining the key points or features of the tutorial (Tritz, 1986); and
- (7) use animation to reveal actions which normally would be invisible to the human eye (Simone, 1992).

The research evidence concerning the effectiveness of animation in learning is mixed. However, there are a number of studies which do provide support for its use and some guidelines for animation in computer-based learning.

2.19 Creation of Computer Graphics

Computer graphics offer several advantages over more traditional methods of producing pictures. Graphics software provides users with a range of tools such as

straight lines, circles and squares which allow them to produce high quality images. Students do not need to have any artistic ability in order to create good quality pictures. The systems are flexible and adaptable so that learners can build on their efforts or elaborate on a provided outline. A wide range of colours is available to be used and revision of the graphic images is straightforward. Graphics software thus provides a tool for learners and teachers to exploit.

Many researchers are interested in the way students represent information in the form of mental models. A mental model is a dynamic mental representation that allows people to make predictions about future states, make inferences, and imagine situations that have not been experienced before (Preece et al, 1994). It does appear that making learners draw diagrams of information may be a good way to develop mental models. Students, however do need assistance to create this type of image (Alesandrini, 1987). More able students may well be able to create mental models of the concepts without the need to draw diagrams (Winn, 1987). There is little evidence to support these ideas; however, several projects are underway.

2.20 Use of Text

Text serves a range of purposes but it has four main functions which are to inform, persuade, guide and instruct readers (Tillman and Glynn, 1987). An author of instructional text is trying to provide a learning experience to assist learners in performing a task, gaining knowledge or changing their attitudes. A great deal of computer-based learning material is text-based and does not exploit the use of graphics, colour or video. Learning materials are however tending to use all the potential of the computer and not simply its ability to display text. The earlier comparison of a computer screen and a page shows that a computer screen is not a text-intensive medium. To limit learning materials to text only would be to fail to capitalise on the potential of the computer. This section will concentrate on considering text in isolation while focusing on its main design features.

A number of researchers have produced lists of the design factors which influence the effectiveness of learning. They all differ in their definition of learning

and this may account for the differing range factors they report as significant. Another important reason for the different lists is simply that there is no common agreed names for many variables. Hathaway (1984) refers to the density of text, Heines (1984) talks in terms of break points, and Isaacs (1987) asks how much text should appear on the screen. All three researchers are essentially discussing how many words should be on each line and how many lines should be on a single screen. This use of different systems of names for the same or similar variables does not aid the development of a set of screen design principles.

Designing the layout of an individual screen requires consideration of the local issues of how much text to present, how to justify it and which colours to use. Designers are aiming to maximise the learning opportunities of a screen and attempting to enhance local comprehension, that is, the comprehension of the text displayed. Designers are also concerned with aiding learners gain an overall comprehension of a tutorial. The presentation of the individual screen will have considerable effects on both the users' comprehension and their speed of assimilating the material.

The focus of this section is on how to present text so that it aids learning, in other words how to present stimulating information that will motivate learners and assist them in retaining and recalling the information. This is clearly stated by four of Gagne's (1985) model of instructional events. These are:

- (1) presenting the stimulus material;
- (2) gaining attention;
- (3) enhancing retention; and
- (4) stimulating recall.

2.21 Density of Text

Hathaway (1984) has stated that: 'the density or compactness of displayed text is a factor that impacts on the speed with which readers comprehend material'. Kolars et al (1981) studied how text density affected the efficiency of reading from a screen. The researchers measured reading speed and accuracy under a variety of conditions. Two densities of text were displayed with two different line spacings and using five different rates of scrolling. The subjects used for the tests were twenty students (mostly females) who had normal vision without corrective lenses. Each person was tested individually at the same time each day over a three day period. They were given twenty noncontiguous passages of about 300 words in length to read followed by a comprehension test to ensure they had read the passage. The results were that the smaller character text (80 characters per line) was read faster than the larger character text (40 characters per line) although there was no difference in comprehension. The addition of double spacing improved reading speeds and marginally improved comprehension. The general conclusion reached was that 80 character double spaced displays produced faster reading times and greater accuracy. The problem with this generalised statement is that the research did not consider other combinations of text density and spacing. Perhaps 50, 60, 70 or 90 characters would have produced better results or triple spacing of the lines even higher comprehension.

A designer may well conclude using common sense reasoning that larger more widely spaced characters would be easier to read (40 character display). A study of eye fixations in reading large characters challenges this view. When a person is reading, the eye does not continuously scan the line of print. It stops and fixates for a short time and then rapidly jumps to another place on the line. Readers are usually unaware of this stop/start movement. The perception of words only occurs during these eye fixations. Individuals use more fixations on large characters without any increase in comprehension to compensate. They therefore need longer to read and occupy more space on the screen to no effect. Smaller characters are the key to good text displays on the screen. The Kolars et al (1981) study provided

limited clues to the optimum size of characters.

Heines (1984) has suggested that after text style, the most important factor in readability is the length of the text line. He states that the shorter the line, the easier it is to read. Long lines need excessive eye movement and make it difficult to move smoothly from the end of one line to the start of the next. With small character sizes this difficulty is increased. This problem has been solved in the newspaper and magazine industry by using short lines and a multi-column layout. Type size is small but short lines maintain readability. A short line of text can allow a reader to scan vertically without the need for the return sweep of the eye trying to find the next line. Heines (1984) has offered the guideline that lines should be between 8 and 10 words long.

Although there are many gaps in the evidence, it is possible to draw some simple conclusions. The smaller 80 character text should be used in preference to 40 character. Each line should be short with 8 to 10 words and the lines should be double spaced. This will result in a screen display which maximises readability and comprehension.

2.22 Reading Speeds

Muter et al (1982) conducted a study comparing reading speed, comprehension and discomfort of reading from a book and a computer display. The results of their tests showed that there was no statistically significant difference between the two groups in terms of comprehension and discomfort but the computer screen was read 28.5% slower than the page. Several other researchers have reached similar conclusions about reading speed. (Kak, 1981; Gould and Grischkowsky, 1982; 1984; Wright and Lickorish 1983; Mills and Weldon, 1984; Heppner et al, 1985). Two studies have however reported no differences in reading speed between the printed page and a screen (Cushman, 1984; Switchenko, 1984).

The original experiments on the different reading speeds of people reading from paper and a screen were carried out by Gould and Grischkowsky (1984). Twenty-four people proof-read extracts from newspapers and magazines on paper

one day and on a screen the next day. Each day involved six hours of proof-reading. Proof-reading was chosen because it was a visually intensive task and likely to be sensitive to the differences in display. The format of both displays was identical, each using 23 lines of 62 to 65 characters. The fonts, polarity and colour of the media differed, the paper showing dark characters on a white background while the screen had greenish characters on a black background.

The results of the experiment showed that 22 out of the 24 subjects read slower from the screen than the paper. The difference began in the first 45 minutes and continued all day at the same rate. The proof-reading speed was 20% to 30% slower depending on how the calculation was made. The readers slowed to 78% of their paper reading speed on the screen or they read 29% faster from paper. The accuracy of detecting spelling mistakes was not significantly different with few errors and with a consistent performance all day. It is interesting to note that although the vast majority of subjects read more slowly from the screen than paper, 12 of the 24 subjects preferred proof-reading from the screen display.

Gould et al (1987) carried out 10 experiments and several analyses in an attempt to explain the difference in reading speed. They considered a wide range of variables which are shown in table 18.

The first experiment compared proof-reading speeds between good quality print and very poor quality print on paper, in order to determine the effects of print quality on proof-reading. The poor quality print was read 20% slower than the good quality print with a similar accuracy of detecting spelling mistakes. The result is in the same range as the reading speed difference between paper and screen display. It could be concluded that print quality was the main factor in determining the speed difference. However, the target group was very small. It should perhaps only be regarded as one indication.

The second experiment compared proof-reading speeds between groups of experienced and inexperienced computer users. The result was almost identical to the original work of Gould and Grischkowsky (1984). 11 out of the 12 subjects read the screen more slowly than the paper. The experienced users read faster from the display than the inexperienced users but were still 20 to 30% slower than their paper reading speeds.

Table 18 Variables Investigated

| |
|---|
| <p>Task variables</p> <ul style="list-style-type: none">Horizontal versus vertical paper orientationTime - when does the difference occurType of reading proof-reading versus comprehensionReal life reading distancesVisual anglesEye movements <p>Display variables</p> <ul style="list-style-type: none">Dynamic characteristics of screen displaysDifferent CRT displaysSelf-luminous versus reflective displaysDisplay contrastScale of layoutCharacter height/line width ratioPolarityAspect RatioSame font <p>Personal variables</p> <ul style="list-style-type: none">Experience at reading from screen displaysAge |
|---|

The third experiment considered the effects of proof-reading against reading for comprehension. The conclusion reached was that users read significantly faster from paper whether they are reading for comprehension or proof-reading.

Differences in reading speed should be reflected in eye fixations as explained previously, either in the number of fixations or their duration. Reading time is obviously the multiple of the number of fixations and the mean duration of the fixation. Experiment six considered the number and duration of eye fixations on paper and the screen during proof-reading. The result was that participants made 15% more fixations during screen reading than on reading paper, that is, an extra fixation per line of text displayed on the screen. The mean duration of fixation was

the same for both media and the increase in fixations converts to a reading speed difference of 25%.

Gould et al (1987) carried out several other experiments considering the effects of different visual display units, reading distance, polarity, fonts and visual angles. The experiments did not isolate any individual display variables which would account for the reading speed difference. Polarity appears to account for a small amount of the difference, as does the scaling difference between paper and the screen. Some of the experiments indicated that some of the difference could be accounted for by the distinction between self-luminous and reflective displays. Text on a printed page is read by the light reflected from the paper whereas a screen is self-luminous.

The results of all the experiments showed that 71 out of 79 subjects read faster from paper and this occurred on a variety of displays, reading tasks and materials and order of performance. The participants fell into two age groups with mean ages of 23 (range 20-27) and 48 (range 40-61). The differences within the two groups were greater than between them. Only 8 people read faster from the screen than paper.

The researchers concluded that the image quality of the characters rather than the task or user variables was the most likely factor to account for reading speed differences. Paterson and Tinker (1940) and Tinker (1963) reviewed decades of reading research on typeset materials. They showed that most variables have only a slight effect on reading speed even when varied widely. However, the result of combining several factors could result in reading speed differences of 20%. The effect of the variables appears to be accumulative on paper and if this is true of screen text then it would support the results of Gould et al (1987).

For screen designers, the reading speeds of users is an important factor but others may be more critical. The emphasis should perhaps be on making the text readable and maintaining the motivation of users to learn the material. Designers need to know that even small changes to design variables may have significant effects on reading speed. The other key result of this work is that there is no significant difference in comprehension between reading from a book and reading from a computer screen.

2.23 Scrolling

There are two main types of scrolling possible on a computer screen - vertical and horizontal. Vertical scrolling moves the display upwards or downwards at some set speed. The top or bottom line of the display disappears and a new line of text appears at the bottom or top of the screen. Horizontal scrolling moves the display along a line of text so that the leading characters disappear to be replaced by new characters. Horizontally scrolling text can move both to the right and to the left. Vertical scrolling can involve the movement of the whole screen or text scrolling in a particular window. Scrolling is one of the dynamic characteristics of computer screens. It should perhaps add interest to the display of text.

Kolers et al (1981) were interested in studying how fast the text should scroll across the screen to maximise reading speed and comprehension. They used five rates of scrolling ranging from zero, that is a static display to 20% faster than the individuals' preferred speed. Two types of scrolling were used: the jump where the text moves up a line at a time; and a smooth transition of the lines over a period of time.

The results of the experiments indicated that as the scrolling rate increased so did the individual's performance, with best results occurring at 20% faster than preferred rate. The performance on the static page was equal to the performance during the 10% faster rate of scrolling. However, even after three sessions, 40% of the participants still preferred the static display. The subjects always preferred a slower scrolling rate than their optimum reading rate even after reading sixteen scrolled pages. The researchers concluded that if the users were free to choose the rate of scrolling it would always be less effective than the static page. Comparing the two sorts of scrolling produced distinctly one sided results. In the jump mode readers made many mistakes, often lost their place and disliked this type of display. The smooth scrolling mode was concluded to be much more effective than the jump mode.

Many researchers and designers of computer-based learning materials have reported that users find it difficult to keep track of scrolling text. Oleron and Tardieu

(1978) suggested that it was more difficult for a person to reorganise and understand a scrolling text than a static display. This is obviously critical if the text being presented is designed to facilitate learning. The conclusion that can be drawn is that static displays should be used for the majority of computer-based learning materials but that scrolling text could be used for scanning material to select suitable sections. An example would perhaps be the examination of a database of examples to illustrate concepts in the main body of a tutorial. The rate of scrolling must be set to be at least 20% faster than the preferred rate or the static page is more effective. The jump mode of scrolling should not be used considering the evidence from Kolers et al (1981).

Bork (1983) reported that users might adjust rates of display to suit individual differences. Students often used rates of less than 30 characters per second which were much slower than they could read. This was interpreted as their desire to feel they were in control of the computer. This is supported by the findings of Schwarz et al (1983) and Oleron and Tardieu (1978) and leads to a considerable doubt about the use of vertical scrolling in computer-based learning materials. Scrolling should be used primarily for searching for material by the learner and not for the presentation of information. It is probably best considered as an aid to browsing material to find a relevant section in the same way pages of a book can be skimmed to see if they are relevant.

The other main type of scrolling is horizontal where the text is advanced across the screen from side to side in jumps of one or more characters. Horizontal scrolling can involve the whole display moving across the screen or a single line moving horizontally. This type of scrolling has a restricted use but can aid screen designers when space is limited for single line menus, feedback to questions and error messages. A graphical display could be labelled with a number of windows each using horizontal scrolling which would be activated by the user. This mode of display allows information to be displayed over the top of the main display in order to enrich it. The horizontal format needs only a small amount of the screen and can display a significant message.

Only a few researchers have considered the use of horizontal scrolling. Granaas et al (1984) investigated the readability of vertically scrolling text with a

constant display rate of 272 words per minute. The main emphasis of their work was on varying the jump lengths of the text and relating this variable to the comprehension of users. They concluded that a jump length of 1 or 2 characters resulted in poorer comprehension than with larger jumps of 4 or more characters. In their opinion, the main factor in reading comprehension was jump length; however, this is difficult to support in that they did not consider the effect of varying the display rate.

Chen and Tsoi (1988) repeated the experiments but also considered the effects of display rates, two different line lengths (20 and 40 character windows), and three jump lengths on readability. They agreed with Granaas et al (1984) that jump length had a significant effect when the display rate was 200 words per minute but at 100 words per minute it did not effect reading performance. With the 20 character window, the medium (5 character) jump length produced the highest comprehension, but there was no significant difference using the 40 character window display. The overall conclusion is that at higher display rates, jump length is a key variable on readers' comprehension but it has no effect at slower speeds.

Chen et al (1988) continued these experiments, adding the extra factor that users could now control their own rate of display of text. Eighteen undergraduates with normal eye sight were the subjects of the tests. Text was displayed in a mixture of upper and lower case at display rates of 45 to 5500 words per minute. Each subject had two sessions of thirty minutes with a five minute interval. During the first session, text was displayed in a twenty character window and in the second session, in a forty character window. The speed of display was started at 45 wpm and each subject was asked to read as quickly as possible from the display. A comprehension test was applied after each test. Table 19 provides a summary of the experimental results.

If the results shown in table 19 are considered from a screen designer's viewpoint, several conclusions can be drawn. The optimum mix of factors is the use of a 20 character window with a jump length of 5, which will result in good comprehension and reading speed. The set of factors to avoid appears to be a 40 character window with a single character jump length. This results in the lowest reading comprehension and second lowest reading speed. In contrast, the same set of

Table 19 Correct Responses in Comprehension and Reading Rates as a Function of Window Size and Jump Length

| Window Size (characters) | 20 | | | 40 | | |
|---|-----------|----------|----------|-----------|----------|----------|
| Jump Length (characters) | 1 | 5 | 9 | 1 | 5 | 9 |
| Reading Rate (words per minute) | 90 | 128 | 139 | 91 | 128 | 144 |
| Comprehension Score (percentage correct) | 81 | 77 | 68 | 66 | 72 | 71 |

variables with a 20 character window display produces the highest comprehension score. The overall variation in comprehension is not great while there is a marked reduction in speed with the one character jump condition. This effect is supported by the researcher's earlier work and also by Granaas et al (1984). Users had to read very slowly to maintain their level of comprehension in the one character jump state. There are several factors not tested in this work, principally the effect of other window sizes.

2.24 Justification

When margins are aligned they are said to be justified, while margins that are not aligned are said to be ragged. This page is 'left justified' in that the left margin is straight and the right margin is ragged. Books and magazines are often presented with both right and left justification and are not generally difficult to read.

Text printed in a book has three characteristics which make the double justification of text pleasing to readers' eyes. It allows for variable letter width so readers' eyes are trained to take in different amounts of information in similar horizontal distances. Typeset text has kerning which allows certain letter combinations such as AV and LY to overlap to maintain a common letter spacing. These two characteristics combined with the ability to add small spaces between adjacent letters means that left and right justification can be achieved without noticeable changes.

These characteristics of typeset text have not in the past been easily available to screen designers. The latest versions of wordprocessing software now offers the option to adjust the space between characters. Kerning is thus becoming available to screen designers (Simpson, 1994). In the past, designers have had to achieve right justification by the addition of extra spaces between words. This produces a poor visual image for readers and due to variation in spacing, a reduction in readability. An earlier conclusion was that a line length of 8 to 10 words was to be recommended for screen design. Right justification of short lines is particularly difficult since fewer spaces between words are available to designers.

Muncer et al (1986) set out to test the view that justified text created by adding spaces randomly to create equal line lengths will result in a decrease in readability. He used 56 students at an American college in the age range 18 to 23 with a wide span of reading ability. Each subject was classified as either a good or a poor reader using similar material to that used in the exercise.

The results of the research showed that double justified text was disruptive to reading and that the effect was greater with good or average readers. Poor readers were not greatly affected by double justification. The researchers suggested the difference in results was due to different reading styles. Effective (fast and fluent) readers tend to read in phrases and are therefore more disrupted by randomly added spaces. Poor (slow with reduced literacy skills) readers read a word at a time and so are not as disrupted by the extra spaces.

Gregory and Poulton (1970) have reported that poorer readers have a significant reduction in reading comprehension with text doubly justified. This is the opposite conclusion to that reached by Muncer et al (1986) but it is probably explained by the way the text was justified in each case. Gregory and Poulton (1970) used hyphenated justification which disrupts the spacing of the text less than the addition of random spaces. Muncer et al (1986) reported that all types of reader were disrupted so the two studies are probably not contradictory.

The overall conclusion is that double justification of computer text is not to be recommended if it is produced by the random addition of extra spaces. Most microcomputers use this type of justification unless specialist software is used to emulate typeset print. Screen designers should use left justification with a ragged right margin. This is especially important if the learners are highly effective (fast and fluent) or average readers.

2.25 Line Break-Points

The issue of justification is complicated by the question of where to break the text into lines. Several authors have suggested that the effects reported as due to justification may be in fact due to choice of points at which to end lines. Hartley (1980a; 1980b) has proposed that line breaks should follow the syntax of the text. Bork (1982) supports this view and claims that poor readers are aided more by this layout than good readers. Heines (1984) feels that good readers can cope with a wide range of layouts while poor readers generally benefit from a syntactic line break layout. Poor readers are assisted to extract the meaning of text presented to them if it is presented to maximise readability and comprehension. Figure 5 shows a comparison of text presented in a standard layout with text presented in syntactic line breaks (Lefrere, 1984).

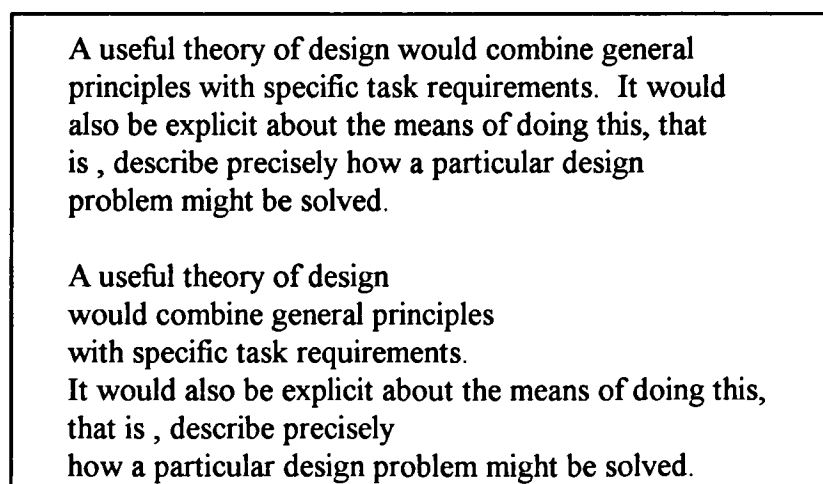


Figure 5 Normal Line Breaks - Syntactic Line Breaks

A number of researchers have tried to improve the readability of text by using different layouts. Andrews (1949) presented text in a 'square span' format. Text was presented in double line blocks, each block representing a single unit of thought. The square block layout improved reading speed but that could have been because of either the thought unit or double line vertical presentation. Later researchers came to a variety of conclusions ranging from there being no great advantage with unit format but shorter units better than longer ones, to the idea that reading improvement was due mainly to the unit layout (North and Jenkins, 1951; Klare, Nichols, and Shuford, 1957; Graf and Torrey, 1966; Carver, 1970). This confusion is probably a result of different reading practices allowed by the researchers. Some allowed re-reading and others skimming, so comparisons are difficult to make.

Carver (1970) experimented with a variety of formats. The subjects used were mature readers and the tasks were normal reading ones. The results showed that there was no advantage to the mature readers of the unit presentation. However, he concluded that poor readers may have benefited. Several other researchers have reported evidence to support this conclusion (Frase and Schwartz, 1979). It does seem that the presentation of meaningful units of text assists the poor or average reader.

Faiola and DeBloois (1988) have suggested that interactive video screen designers should aim for text which is easy to read and that holds readers' interest. They propose that this is achievable by dividing large elements of the material into smaller pieces. This will improve the material's clarity and assist the users' retention of the material. Caldwell (1980) has also put forward a similar view of screen design. He states that the learning material should be broken down into individual paragraphs and units. These should be double spaced and key points highlighted in colour.

The presentation of text on computer screens is limited. Computer screens are not a text intensive medium. The presentation of material in meaningful units is however easy to support in this type of environment. Users are able to control the presentation of each unit.

2.26 Upper or Lower Case Letters

The vast majority of text that readers will come across is a mixture of upper and lower case. It is therefore in keeping with readers' experience that computer displays should use both unless some special factors suggests one is more legible than the other.

Faiola and DeBloois (1988) report a study in which it was shown that lower case is more legible than upper case text. Heines (1984) feels that the use of lower case letters adds considerably to readability of the text and that a mixture of upper and lower case should always be used.

Henney (1981) tested the effects of varying upper and lower case text on speed and accuracy of reading. Her results showed that text using both cases was read faster than upper case on its own while the users were more accurate reading upper case text alone. The differences were so small (one second per paragraph faster) that they do not offer strong evidence for designers.

With little definite evidence to influence designers, it is wise to reflect the wider world of books, magazines and newspapers and display text which use both cases. This option also allows the added flexibility that words can be displayed in upper case to emphasise an important issue.

2.27 Comparison of Text Evidence

These text design factors were largely arrived at in isolation from each other, therefore it should not be assumed that simply combining all the elements will produce effective design. The interaction of the different elements is likely to have quite the opposite effect on occasions. Paterson and Tinker (1940) and Tinker (1963) have described the effect of changing a single variable which results in minor changes to the users' behaviour but when combined with other variables causes major effects.

A comparison of the different experimental methods used by the various researchers also leads to the conclusion that the results should be used with caution.

A number of researchers have used comprehension tests which should provide an element of consistency between their work. However, there are considerable differences between the comprehension tests employed. Multiple choice questions, summarising into a tape recorder and indicating if a question can be answered on the basis of the information contained in the passage have all been used as tests of comprehension.

Gould et al (1987) asked subjects to summarise what they had read into a tape recorder but did not analyse what they had recorded. It is therefore not a good test of comprehension or even that the passage has been read and not merely skimmed. Kolers et al (1981) presented their subjects with ten questions at the end of each passage. The subjects had to choose which questions could be answered using the information contained in the material. This was intended as a simple test of whether the passage had been read and it serves that purpose but it is not an effective test of comprehension. Chen and Tsoi (1988) asked their subjects to answer four multiple choice questions after reading each of eighteen paragraphs (one question on each 50 words presented). They therefore tested the comprehension of the individuals in a detailed and specific manner. These examples serve to illustrate the difficulties of applying the results of an individual researcher as part of an overall set of design principles.

These experiments are mainly concerned with the immediate issues of presenting single screens of text to learners. To produce a full model of screen design needs more evidence than is currently available.

2.28 Review of Experimental Design

Table 20 shows a comparison of a range of experimental designs in terms of subjects, media, learning focus and integration of the material. Computer-based learning materials can be characterised as using computers, aimed at assisting learners and involving learners using many different displays (integrated display) in a sequence which they choose. Table 20 compares a number of experiments in terms of these characteristics.

Table 20 Review of Experimental Design

| Researchers | Subjects | Media | Learning Focus | Integrated Material |
|---------------------------------------|--|--------------------|-----------------------|----------------------------|
| Alesandrini and Rigney, 1981 | Students | Computer | Yes | Yes |
| Anglin, 1987 | Students | Paper | Yes | Not Applicable |
| Avons et al, 1983 | Children | Computer | Yes | No |
| Baek and Layne, 1988 | High School Students | Computer | Yes | Yes |
| Bauer and Cavonius, 1980 | Adults ages 23 to 47 years old | Computer | No | Yes |
| Bernard, 1990 | Student Nurses ages 18 to 21 years old | Paper | Yes | Not Applicable |
| Booher, 1975 | Adult Navy Personnel | Electronic Display | Yes | No |
| Brooke and Duncan, 1981 | Adults | Computer | No | No |
| Bruce and Foster, 1982 | Students | Computer | No | No |
| Cahill and Carter, 1976 | Adults | Slides | No | Not Applicable |
| Campbell, Marchetti and Mewhort, 1981 | Students | Paper | No | Not Applicable |
| Chen et al, 1988 | Students | Computer | No | No |
| Davidson, 1976 | Students | Paper | Yes | Not Applicable |
| Dean and Kulhavy, 1981 | Students | Paper | Yes | Not Applicable |
| Fisher and Tan, 1989 | Students | Computer | No | No |

| | | | | |
|--|----------------------|--------------------|-----|----------------|
| Gould and Grischkowsky, 1984 | Adults | Computer and Paper | No | No |
| Gould et al, 1987 | Adults | Computer and Paper | No | No |
| Guastello, Traut and Korienek, 1989 | Students | Paper | No | Not Applicable |
| Hativa and Teper, 1988 | Students | Computer | Yes | Yes |
| Holliday, 1976 | School Students | Paper | Yes | Not Applicable |
| Kelley, 1988 | Adults | Computer | Yes | Yes |
| Kolers et al 1981 | Students | Computer | No | No |
| Kruk and Muter, 1984 | Students | Computer | No | Yes |
| Luder, 1984 | Adult research staff | Slides | No | Not Applicable |
| Macdonald and Cole, 1988 | Adults | Aircraft Display | No | Not Applicable |
| Marschalek, 1988 | Children | Computer | No | Not Applicable |
| Muncer, Gorman, Gorman and Bibel, 1986 | Students | Paper | No | Not Applicable |
| Patrick and Fitzgibbon, 1988 | Students | Computer | Yes | Not Applicable |
| Pond, 1982 | Adults | Computer | No | No |
| Reid and Wicks, 1988 | Children | Computer | No | No |
| Rigney and Lutz, 1976 | Students | Computer | Yes | Yes |
| Ross, Morrison and O'Dell, 1988 | Students | Computer | Yes | Yes |
| Scanlon, 1967 | Adults | Television | No | Not Applicable |

| | | | | |
|-----------------------------------|----------------------------|-----------------------|-----|----------------|
| Schwarz, Beldie and Pastoor, 1983 | Adults | Computer | No | No |
| Seddon and Shubber, 1985 | Children | Paper and 35mm Slides | Yes | Not Applicable |
| Sewell and Moore, 1980 | Students | Paper and Audio Tape | Yes | Not Applicable |
| Smith, 1962 | Adults | Colour Slides | No | Not Applicable |
| Smith and Thomas, 1964 | Adults | Colour Slides | No | Not Applicable |
| Surber and Leeder, 1988 | Children | Computer | Yes | No |
| Tullis, 1983 | Adult Television Engineers | Computer | No | No |
| Willows, 1978 | Children | Paper | No | Not Applicable |

Subjects

Most researchers find it difficult to obtain sufficient suitable subjects. This problem is illustrated in the type of subjects used by many researchers. Often subjects for research into screen design have been drawn from the population of students who are more readily available to take part in experiments than other groups. This immediately raises the question, 'Are these subjects representative of the overall population who would be using computer-based learning?'. In many cases they are not since they are drawn from a narrow age group, have higher qualifications than the general population and often have no experience other than the world of education. These are generalisations and are not true of all the experimental designs. However, as table 20 shows, half (21 out of 41 designs) of the experiments used students and frequently undergraduates.

Media

The screen design literature, and in particular the advice available to designers, often uses sources of evidence which are based on research on media other than the computer. Frequently, the research into instructional text, colour television, slides and electronic instrument displays has been used to support screen design guidelines. The validity of transferring results from other media to computer screens has not been attempted on any systematic basis and often authors do not state the sources of their evidence. Table 20 reveals that 18 of the 41 experiments were based on media other than a computer screen.

Learning Focus

The review of the literature reveals that the majority of the evidence available from research comes from experiments which were not focused on learning, that is, the experimental subjects were not trying to learn from the materials used. Only 17 of the 41 experiments compared have a focus on learning.

Subjects who are trying to learn from a display are unlike other users of computers. Their motivation for using the material is different. Designers of a wordprocessor know that users will initially need to be able to input, save and print text. The more advanced features will normally be assimilated over a period of time. It is unlikely that they will be needed on the users' first visit to the software. However, learners do not normally use learning materials over an extended period of time and so do require an understanding of all the features of the system on their first visit to the material. Many learners will only use a learning tutorial once or twice. It is rare that learners use a tutorial many times over an extended period.

Integrated Materials

Computer-based learning materials consist of multiple screens of tutorial with additional features such as glossaries of terms, tests and additional information or help systems. It is thus an integrated environment of many different screen designs

through which learners move. The learners using the materials will have a number of choices available to them about where to go to next and which functions to activate. Screen design is not simply about a single display which learners study, but rather multiple screen displays which learners interact with and from which they make choices. Learners are faced with numerous displays. Screen design should be considered in the context of multiple displays which may interact with each other rather than a single isolated display. A comparison of experimental designs in table 20 shows in fact that only 8 designs from 41 used an integrated display.

Summary and Conclusions of Review of Experimental Design

The ideal experimental design to consider screen design of computer-based learning materials should have:

- (1) subjects who are drawn from the representative population;
- (2) appropriately designed computer-based materials;
- (3) a learning focus; and
- (4) integrated displays.

Only 6 of the 41 experiments compared are based on a computer, using an integrated display and with a learning focus.

2.29 Survey of Computer-Based Learning Materials

Research evidence will often present an image to authors of computer-based learning which is idealised in the sense that it will only consider the screen design factors and not the limitations which an author is normally working under. Few authors will have unlimited resources and thus their designs will be influenced by financial and other restraints. During this research project, computer-based learning materials were designed in order to collect evidence of how learners react to particular screen designs with the aim of identifying results which will assist designers. It is thus important to be aware not only of the results of research but also the standard of

screen design demonstrated by existing computer-based learning material. In a sense, it is necessary not only to review the research literature published in the form of articles and books but also the practical outcomes in the form of computer-based learning materials.

A survey of all the computer-based learning material in use in the Employment Department's Learning Resource Centre at Sheffield was carried out. This is a specialist resource centre which aims to show the range of computer-based learning approaches in use rather than providing a range of subject-specific learning products to users. It contains computer-based learning material which represents the results of development projects supported by the Employment Department. The computer-based learning material covered in the survey is thus a sample of new or leading edge material rather than a cross section of all products produced to date. The survey is not comprehensive but illustrates a sample of the material currently available in Great Britain. The aim of the survey was to identify the use of graphics in a sample of computer-based learning material. A total of thirty-four packages was surveyed in the areas of industrial, commercial, basic education and trainer/teacher training.

The survey was based on an analysis of individual screen displays. A single graphic image of any size on a screen (except ones acting as icons) counted as a graphic display. Thus, a part-screen single graphic or a part-screen with multiple graphic images could also contain some text. Only a full-screen graphic display was likely to contain little or no text. The results of the survey are displayed in tables 21, 22, 23, 24 and 25. Table 21 provides a summary of the survey.

Graphics and Text

The overall analysis (table 21) of the learning packages shows that text-only screens represent 64% of the total, with graphics contributing the remaining 36% of screen displays. The balance is even more pronounced when the graphic displays are considered, since part-screen graphics are by far the largest component. These part-screen graphic displays contain a large element of text. The commonest device to convey learning is text and in the main, coloured text with 86% of the text shown

using two or three colours. Only 10% of screens use a single text colour.

Representational graphics are the commonest employed (almost two thirds), followed by logical graphics (almost one-third) and a rare analogical graphic. These results are in line with those resulting from Alesandrini's (1985) survey of American courseware.

Table 21 Summary of Survey

| Type | Industrial Training | Commercial Training | Basic Education | Trainers and Teachers | Total |
|----------------------------|----------------------------|----------------------------|------------------------|------------------------------|--------------|
| Packages surveyed | 10 | 13 | 5 | 6 | 34 |
| Amount of text (total) | 45% | 67% | 70% | 67.5% | 64% |
| Amount of graphics (total) | 55% | 33% | 30% | 32.5% | 36% |
| Representational (total) | 33% | 15% | 16% | 22.5% | 22% |
| Analogical (total) | 2% | 3% | 1% | 2.5% | 2% |
| Logical (total) | 20% | 15% | 13% | 7.5% | 12% |

Industrial training materials use graphics extensively with 55% of the displays employing them. The emphasis in the industrial displays is on the use of representational graphics (33%) and logical graphics (20%). Analogical graphics contribute only 2% of the total screen displays. The other three subjects use a minority of graphic screens: commercial training 33%, trainer/teacher 32.5% and basic education 30%.

Size of Graphic Used

Graphics which occupied only part of the screen are the commonest employed and these graphics occupy less than half of the screen.

Part-screen graphics represent:

- (1) 86% of the representational graphics; and
- (2) 58% of the logical graphics.

Only a small number of graphic displays use the whole screen and few displays use multiple part-screen graphics (21% of representational graphics). Even in those displays employing multiple images, graphics occupy less than half the screen. A stereotype graphic display would consist of a small image with supporting text using two or three colours. Table 22 provides a detailed analysis of the type and size of graphics identified.

Use of Colour in Text

The survey of the thirty-four computer-based learning packages shows that the majority of the material used multicoloured text. The overall emphasis in the material surveyed is to restrict the use of coloured text to four or less colours. This is perhaps a reaction to the general advice on the use of colour within computer-based learning which emphasises limiting the use of colour. Table 23 provides an analysis of the number of colours used to display text.

Icons and Hotspots

Icons and hotspots were commonly used in all four types of learning material although frequently only a single icon or hotspot was employed. This single icon or hotspot normally represented the page forward switch. At the other extreme, one package employed twenty-three icons on each screen display throughout the material. Table 24 provides an overview of the use of icons and hotspots in the different types of learning material.

Table 22 Analysis of Full and Part-Screen Images by Type

| Type | Industrial Training | Commercial Training | Basic Education | Trainers and Teachers | Total |
|--|----------------------------|----------------------------|------------------------|------------------------------|--------------|
| Representational (full-screen) | 1.5% | 3.5% | 1.0% | 17.5% | 3.0% |
| Representational (part-screen single image) | 24.0% | 7.5% | 14.0% | 5.0% | 15.0% |
| Representational (full-screen multiple images) | 0.5% | 0% | 0% | 0% | 0% |
| Representational (part-screen multiple images) | 7.0% | 4.0% | 1.0% | 0% | 4.0% |
| Representational (Total) | 33.0% | 15.0% | 16.0% | 22.5% | 22.0% |
| Analogical (all types) | 2.0% | 3.0% | 1.0% | 2.5% | 2.0% |
| Logical (full-Screen) | 7.5% | 7.0% | 1.0% | 6.0% | 4.0% |
| Logical (part-screen single image) | 11.0% | 7.0% | 12.0% | 1.5% | 7.0% |
| Logical (full-screen multiple images) | 0% | 0% | 0% | 0% | 1.0% |
| Logical (part-screen multiple images) | 1.5% | 1.0% | 0% | 0% | 0% |
| Logical (Total) | 20.0% | 15.0% | 13.0% | 7.5% | 12.0% |

Table 23 Analysis of Text (Foreground) Colours

| No. of Colours | Industrial Training | Commercial Training | Basic Education | Trainers and Teachers | Total |
|-----------------------|----------------------------|----------------------------|------------------------|------------------------------|--------------|
| 1 | 9.0% | 9.0% | 10.0% | 9.0% | 10.0% |
| 2 | 75.0% | 67.0% | 58.0% | 22.0% | 58.0% |
| 3 | 12.0% | 24.0% | 20.0% | 62.0% | 28.5% |
| 4 | 5.0% | 0% | 12.0% | 3.5% | 3.0% |
| 5 | 0% | 0% | 0% | 0% | 0% |
| 6 | 0% | 0% | 0% | 0% | 0% |
| 7 | 0% | 0% | 0% | 0% | 0% |
| 8 | 0% | 0% | 0% | 3.5% | 0.5% |

Table 24 Analysis of Use of Icons/Hotspots in Courseware

| Type | Industrial Training | Commercial Training | Basic Education | Trainers and Teachers | Total |
|--------------------|----------------------------|----------------------------|------------------------|------------------------------|--------------|
| Words | 65.5% | 10.0% | 44.0% | 33.0% | 36.0% |
| Pictures | 23.5% | 75.0% | 56.0% | 51.0% | 51.0% |
| Words and Pictures | 11.0% | 15.0% | 0% | 16.0% | 13.0% |

The limited research evidence available (Guastello, Traut and Korienek, 1989) shows that mixed word and picture icons are more meaningful than either word hotspots or picture icons on their own. This preference is apparent across a wide range of domains. The survey, however, reveals that they are the least used with only 13% of material using word and picture icons. Picture only icons are the commonest with 51% of material using them while 36% used word only hotspots. This is perhaps an example of designers following their own experience rather than research evidence. Table 25 provides a breakdown of the survey and illustrates the

range of icons and hotspots used in the sample. It shows the type of icon/hotspot used in displays employing different numbers of icons/hotspots (frequency) on each display.

Table 25 Number and Type of Icon/Hotspots

| Frequency of Icons/ Hotspots | Word | Picture | Word and Picture | Total Number of Displays |
|-------------------------------------|-------------|----------------|-------------------------|---------------------------------|
| 1 | 39 | 161 | 1 | 201 |
| 2 | 30 | 21 | 13 | 64 |
| 3 | 36 | 1 | 0 | 37 |
| 4 | 17 | 0 | 0 | 17 |
| 5 | 5 | 1 | 0 | 6 |
| 6 | 8 | 2 | 20 | 30 |
| 7 | 1 | 0 | 0 | 1 |
| 8 | 3 | 4 | 0 | 7 |
| 10 | 1 | 0 | 0 | 1 |
| 11 | 3 | 0 | 0 | 3 |
| 12 | 2 | 0 | 1 | 3 |
| 13 | 6 | 0 | 0 | 6 |
| 16 | 1 | 0 | 0 | 1 |
| 17 | 0 | 0 | 16 | 16 |
| 23 | 0 | 0 | 23 | 23 |

Control Approaches

There are a range of approaches to the user control of computer packages. The main methods are menus, command languages, graphic interfaces (these are characterised by the use of a mouse as a pointing device and icons as links to other parts of the material) and combinations or variations of these main approaches.

The survey of computer-based learning material in use in the Employment Department's Learning Resource Centre confirmed that current learning material tends to use a main menu combined with a graphic interface (iconic links). This was the control approach adopted in the Colour and Graphics experiments.

Implications of Survey

In the survey of computer-based learning materials there are a number of implications for the design of experiments. These are:

Graphics

- (1) all three types of graphic image are employed in current learning materials;
- (2) representative graphics are the most frequently used of the three types followed by logical and analogical images, respectively;
- (3) both full- and part-screen images are used, with single part-screen graphics being the most frequently used;
- (4) graphic screens are not used to the full extent of their potential, in particular analogical images are only occasionally used;

Text

- (5) text is the most commonly employed device to convey learning material;
- (6) computer-based learning material uses up to four colours in text displays;

Icons/Hotspots

- (7) icons are often used in computer-based learning;
- (8) although the number of icons varies considerably, the most frequently used is a single icon or hotspot;
- (9) icons and hotspots take three main forms - picture only, word only and picture and word combinations;
- (10) picture only icons are the most frequently employed, and

Control of Computer-Based Learning

- (11) the main control approach in computer-based learning is a combination of menu and icons using a pointing device such as a mouse.

For the design of the Colour and Graphics experiments, these implications were interpreted to form the basis of the material and also to extend the use of graphics in a systematic way. The experimental material therefore used:

- (1) a control approach combining a menu and icons using a mouse as a pointing device;
- (2) all three types of graphic image;
- (3) analogical images more extensively in order to explore their use in computer-based learning since potentially they are powerful aids to learning; and
- (4) text to support the learning in an appropriate way.

2.30 Conclusions

There is a considerable body of research literature related to screen design. However, the material is not evenly balanced between graphics, colour and text. Only text has an overall substantial well-supported body of evidence. A great deal of the research is based on media other than computer screens.

Analysis of experimental design shows that many of the experiments were not based on a computer screen, nor did they employ representative subjects. Equally important, few experiments were related to learning or used an integrated display. Only approximately 14% of the experiments considered met these criteria (computer screen, learning focus and an integrated display). These factors suggest that the transfer of evidence to a computer screen will be limited.

The survey of the computer-based learning material provided a basis for the design of the Colour and Graphics experiments. In a number of areas the survey suggested that designers limited their use of colour and graphics. Only four colours were used in text-only computer-based learning materials and graphics were under-employed compared to their capacity to aid learning. In particular, analogical graphics were only occasionally employed although they do appear to have the potential to assist learning.

Chapter Three

EXPERIMENTAL DESIGN AND RESOURCE PRODUCTION

3.0 Introduction

This chapter deals with the design and resource production of the Colour and Graphics experiments. It discusses how to reduce the threats to experimental validity and the various design choices before considering the production of learning material. The Colour experiment was used to test the research methods while the robustness of both experiments was considered. The outcomes of the review of experimental design and survey of computer-based learning material reported in chapter two were used as the basis for the design of the experimental resources.

The experimental variables are given and in particular the approach to the measurement of screen complexity is explained for text-only, graphics-only and text and graphics combined screens.

Finally, the development of software tools and experimental shells in order to produce the resources for the Colour and Graphics experiments is discussed. Examples of the approach are provided.

3.1 Design of Experiments

The overall approach to the research was to investigate the three fundamental screen design elements of text, colour and graphics. The review of literature and survey of computer-based learning materials provided the background on which to construct the experiments. The key features of the experiments were:

Text: A number of guidelines relating to the use of text on computer screens were identified and applied to the experiments.

Colour: The number of colours used on a display and how it influenced the learners' behaviour, in particular the time learners spent on the tutorial and their exploration of the tutorial. The findings of the Colour experiment were applied to the design of the Graphics tutorial (how colour was used in the learning material).

Graphics: The size and type of graphics and their influence on learners' behaviour in terms of the time spent on each screen display.

Learners: The learners' prior knowledge of the tutorial subject and how it influenced their choice of route through the tutorial and time spent on each screen display.

Complexity: The relationship between time spent on each screen display and the complexity of the display.

Experimental: To ensure that the experimental approach was both valid and reliable.

Methods

Experimental Design: During the Colour experiment, the design of the software was also tested and the results applied to the design of the Graphics experiment.

It was the policy that both experiments were designed to follow the characteristics of effective screen design research discussed earlier in the review of experimental design; that is, computer-based, with a learning focus, using a range of displays and with a broad range of subjects.

The experimental hypotheses are given in full in section 1.10. Two experiments were designed. The first (Colour) was to explore the issues connected with the experimental methodology and colour and the second (Graphics) to consider the use of graphics, prior knowledge and complexity.

The outcomes of the Colour experiment were used to develop the design of the Graphics experiment in a number of ways:

- (1) structure of the experimental shell;
- (2) development of the software; and
- (3) use of colour within the learning material.

Both experimental tutorials adopted identical screen design guidelines for the common features, such as the use of text.

The main method was to develop the computer-based learning material complete with software to collect the desired information. The experimental subjects would then work through the tutorial and their behaviour would be automatically recorded on a computer disk. The information measured is given in sections 3.4 (Colour experiment) and 3.5 (Graphics experiment). In addition to the automatic data collection, the subjects were asked to complete a questionnaire. This provided basic information about them as well as their attitude to the design of the tutorial.

In judging the design of a piece of research, the validity of the experiment is of prime importance. The experimental design must be capable of answering the questions that are posed. Validity can be divided into two types: internal; and external.

An internally valid experiment will provide evidence to prove whether a variable makes any difference to the specific experimental instance. An example could be 'does the change of background colour improve the retention of the learning material displayed?' An internally valid experiment could answer the question but that does not mean that the results can be applied in other circumstances. The experimental design would need to be externally valid in order for the results to be transferable to other situations for the same or similar groups. The design of a research experiment should ideally be both internally and externally valid.

Influences of Internal Validity

Campbell and Stanley (1963) identified eight influences which might threaten internal validity. These are:

1. History

Influences between measurements could affect the results of experiments. In general, the longer the interval between measurements, the greater the probability that the experiment will be invalid.

2. Maturation

The subjects themselves will change during the experiment (they become tired, bored, hungry and older).

3. Testing

If the subjects take a pre-test before the experiment, it can lead to their obtaining a higher score in the post-test owing to familiarity with the test.

4. Instrumentation

If changes in measurement criteria or observation (new observers or change in scoring criteria) are made during the experiment, this can effect results.

5. Statistical Regression

With groups who achieve extreme results, there is a tendency to score closer to the average on subsequent attempts.

6. Selection

The method used to divide subjects into different groups may result in a bias which could obscure experimental results.

7. Experimental Mortality

This occurs when members of the comparison groups withdraw at different rates. This leads to groups which were comparable at the start of the experiment changing to non-similar groups by the end of the experiment.

8. Interactions among Factors

The individual influences may act together to affect the result of the experiment.

Influences of External Validity

Campbell and Stanley (1963) identified four influences of external validity. These are:

1. Reactive or Interactive Effects of Testing

The pre-test may sensitise the subjects to the experimental variable.

2. Interaction of Selection

The selected individuals may be unusually susceptible to the effects of the experimental variable. Thus it may not be possible to generalise the results to the larger population.

3. Reactive Effects of Experimental Arrangements

Findings observed in a laboratory situation may not be readily transferable to the real world.

4. Multiple Treatment Interference

If the subjects have experienced multiple experimental treatments, then their effects may be accumulative and the individual effects will be difficult to separate from each other.

Experimental Designs

There are a number of research designs which have been used in experiments. These may be classified into two main types which are pre-experimental and true experimental designs. The pre-experimental designs have difficulty in controlling the threats to internal validity to such an extent that they have little value in research work. The true experimental designs offer strong internal validity but may have some problems with external validity. For the purpose of this research, the choice of experimental design must rest with the true experimental designs because they control the threat to internal validity. The three designs are:

1. Pre-test and Post-test Control Group Design (Figure 6)

In this design, two or more groups are selected randomly. The groups are tested prior to the experimental treatment. The control group does not take part in the experimental treatment but is tested again once the other groups have experienced the treatment. This design controls the problems threatening internal validity but still leaves the threats to external validity to be overcome.

2. Post-test-only Control Group Design (Figure 6)

In this design, two or more groups are selected randomly. The groups are not tested prior to the experimental treatment. The control group is tested with the experimental groups once they have experienced the treatment. By not pre-testing the groups, this design overcomes the issue of interaction between pre-test and experimental variable. The other influences on external validation are still to be considered.

3. Solomon Four Group Design (Figure 6)

This design is essentially a combination of the previous two designs. Two experimental and two control groups are used. Pre-tests are given to two of the groups but not the other two (one control and one experimental group). All four groups are tested once the two experimental groups have experienced the treatment. This design has the advantage over the other two designs in that it does provide more information. It is possible to determine if there is an interaction effect between the pre-test and the treatment results.

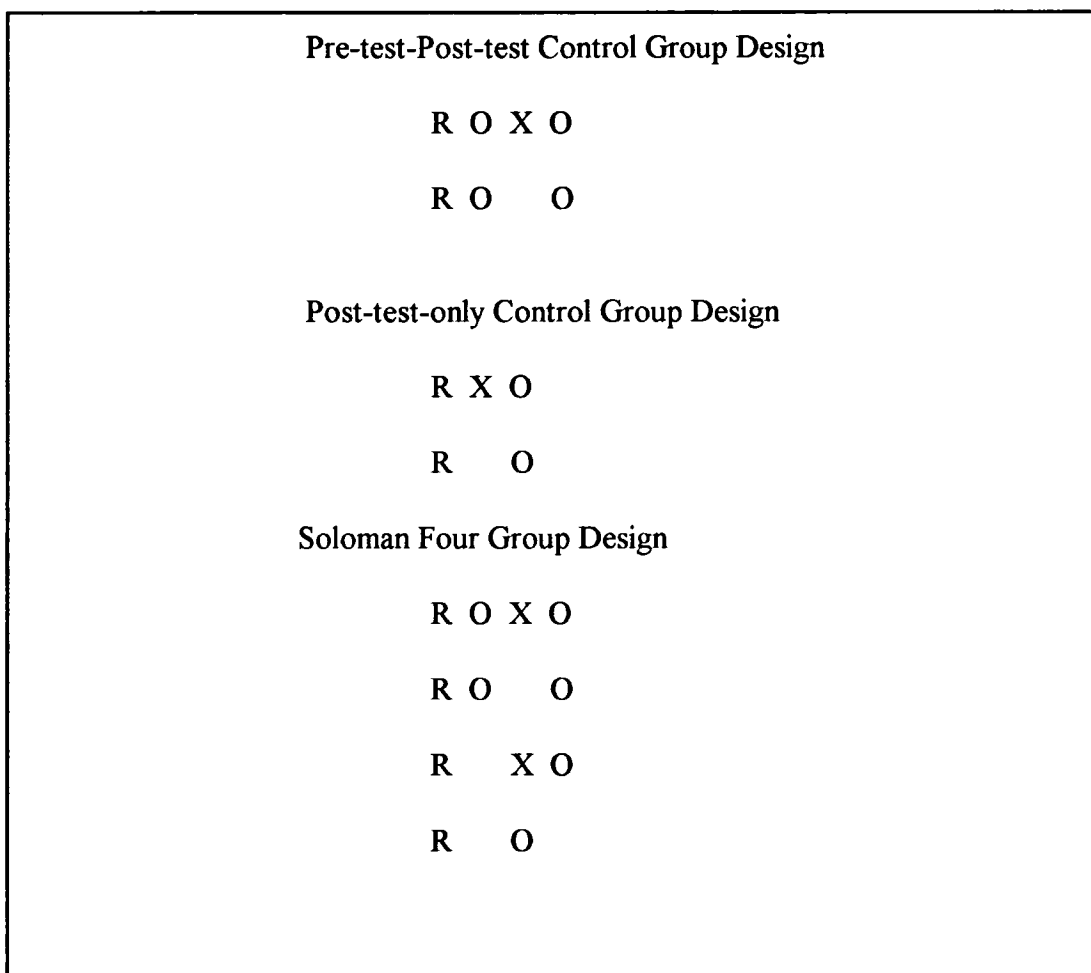


Figure 6 True Experimental Designs

R = Random assignment of subjects to different experimental treatments
O = Observation of the subjects
X = Experimental treatment

Subjects

A critical limitation to any experimental design which involves subjects using computers is their own attitudes to the computer. A range of research activity has concentrated on identifying the factors which lead to a negative attitude to the use of computers. An experiment which chooses as its subjects a group who were always going to be negative or positive to computers is going to produce results which are skewed.

Richey (1991) investigated the attitudes of adults to learning from a computer-based system compared to other methods. She found that only 6.6% of her subjects preferred computer-based learning to the other methods while 25% rated it as their least favoured option. The most preferred methods were:

- (1) instructor/video 58.5%;
- (2) lecture/discussion 29.9%;
- (3) computer-based learning 6.6%; and
- (4) workbooks 5.0%.

In contrast to these results, Riding (1992) reported that staff working in the hotel and catering industry preferred computer-based learning to individual instruction (by their line manager), watching a video or reading a manual. The difference between the two studies is probably due to the methods to which the computer-based approach was compared and the experience staff had of the other methods. The study did not quantify the preference in terms which allowed for a comparison with the Richey (1991) results. However, a large proportion of the trainees still preferred other methods.

A number of studies have considered which factors relate to a positive or negative attitude to computers. The main factors are age, educational achievement, gender, economic status and degree of computer literacy.

The studies (Lloyd and Gressard, 1984; Flynn, 1989; Mahmood and Medewitz, 1989; Morris, 1989) showed that:

- (1) age and educational achievement were not consistently found to be factors in determining computer attitudes (Lloyd and Gressard, 1984; Morris, 1989);
- (2) there was no evidence that negative attitudes were associated with gender or socio-economic status (Flynn, 1989); and
- (3) positive attitude was consistently linked to the amount of computer experience the individual possessed (Lloyd and Gressard, 1984; Mahmood and Medewitz, 1989).

The design of the Graphics experiment specifically targeted subjects who were experienced computer users. Even the beginners' group were computer literate and regular users of computers. The subjects were therefore likely to have a positive attitude to computers and possibly to computer-based learning.

Review of Experimental Design

The review of the literature (section 2.28) revealed that there were four limitations to experiments considering screen design issues. These were subjects (51% of experiments used students/undergraduates), media (44% based on media other than computers), learning (59% not focused on a learning task) and integration (80% did not use multiple screen displays). Only 15% of the experiments compared were computer-based, had a learning focus and employed an integrated system of multiple screen displays.

The Graphics experiment reduced these limitations by drawing experimental subjects from the working population who were all regular computer users. The experimental subject group represented ages from 16 to 64, educational achievement from no qualifications to post-graduate awards and both men and women. Overall,

the experimental subjects were better qualified and older than the working population. This may be a restriction in transferring the results of these experiments. However, analysis of the students of the National Extension College (Force 9, 1993) showed that typical users of Open Learning tended to be urban, white collar workers with higher income than the average. This suggests that the sample is perhaps representative of users of Open Learning which includes computer-based learning.

Both the Colour and Graphics experiments were based around the computer screen, using an IBM compatible computer since it is the most available computer in the world. Although the IBM compatible computers offer a choice of several resolutions in which to display materials on the screen, the experiments used only one display resolution - 640 x 480.

The focus of both experiments was learning. They provided tutorials on study skills and computer science. There are however a number of approaches to learning which were not included in the experiments, such as simulation and experiential learning. The results of the experiments are therefore valid when considering tutorial style approaches but other learning designs have not been investigated.

The Graphics experiment used a fully integrated approach, with graphics and text screen designed to form part of the whole tutorial. Different types of graphics were integrated together so that the results reflected the complex design of computer-based learning materials.

Typical Computer-Based Learning Materials

The survey of computer-based learning materials undertaken in chapter two showed that, on average, most material used a majority of text and a minority of graphics. The Graphics experiment was designed to produce a balanced use of text and graphics. The comparison of the survey results and the experiment is shown in table 26.

Table 26 Comparison of Computer-Based Learning Survey and Graphics Experiment

| Display | Analysis | Experiment |
|-------------------------|-----------------|-------------------|
| Text | 64% | 47% |
| Graphics | 36% | 53% |
| Representative Graphics | 22% | 39% |
| Analogical Graphics | 2% | 9% |
| Logical Graphics | 12% | 5% |

The Graphics experiment was therefore not representative of the average package as presented by the survey. It contained more graphics and used a different balance of the different types of graphics.

3.2 Design of the Colour Experiment

The development of computer-based learning material is a considerable task. It has been shown to take between 150 to 4000 hours to produce one hour of learning material (Gery, 1987; Rowntree, 1992; Marshall et al, 1994). The two experiments (Colour and Graphics) used in this research produced approximately one hour and three hours of material, respectively. The computer-based learning material was not designed simply to help individuals to learn, but aimed at investigating the variables of screen design. It was therefore important that the experimental variables and learning content could be efficiently and effectively adjusted. The normal solution is to use a software authoring tool which is specifically designed to develop computer-based learning material. Unfortunately, authoring tools always impose some constraints on design in order to provide an efficient system.

In order to provide a suitable system which is both efficient and free of constraint, a software system or shell was created for the research. This shell has some of the facilities of a computer-based learning authoring tool to provide an

efficient development system combined with the specific features to allow experimental variables to be adjusted. The shell is a software framework in which learning material can be inserted or changed without the need to amend the underlying structure. The shell is essentially free of learning content and provides a framework or structure for the learning content to fit into. The structure can also be adjusted by simply moving parts of the framework around. The text is separated from the shell so that it can be created, changed or amended quickly thus providing an efficient tool. In the Colour experiment, the variables were separate from the learning material so that they could be varied by a single change to the software. As part of the standard structure, the shell recorded the experimental information automatically (route taken, time taken on each frame and version of material).

The design of the Colour shell was based on the post-test-only control group design. The assignment of subjects to experimental treatment was random and carried out by a program within the shell. The subjects forming the various experimental groups were therefore randomly selected. The experiment was carried out in the natural environment of the subjects rather than a laboratory setting. Many of the experiments reviewed in chapter 2 revealed that experimental subjects were frequently asked to undertake learning or other tasks in laboratory settings. This may cause reactive effects between the experimental arrangements and the variables being studied. By undertaking the screen design experiments in a subject's own environment, this risk was eliminated.

The subjects were assigned randomly to an experimental group. Each experimental group undertook one of the tutorials which was testing a chosen design strategy. Figure 7 illustrates the approach taken. Once they had completed the tutorial, the subjects were asked to complete a questionnaire. There were several different design strategies being tested and the variation of the subjects (age, sex, and learning ability) meant that several variables were interacting simultaneously. The design of the experiment was a factorial design in that the variables were manipulated simultaneously in order to study the effect of each variable on the dependent variable as well as the effects due to interactions among the several variables. The experiment was designed to assess the individual and combined effects of the variables, such as a particular screen design strategy against the learning ability of a group.

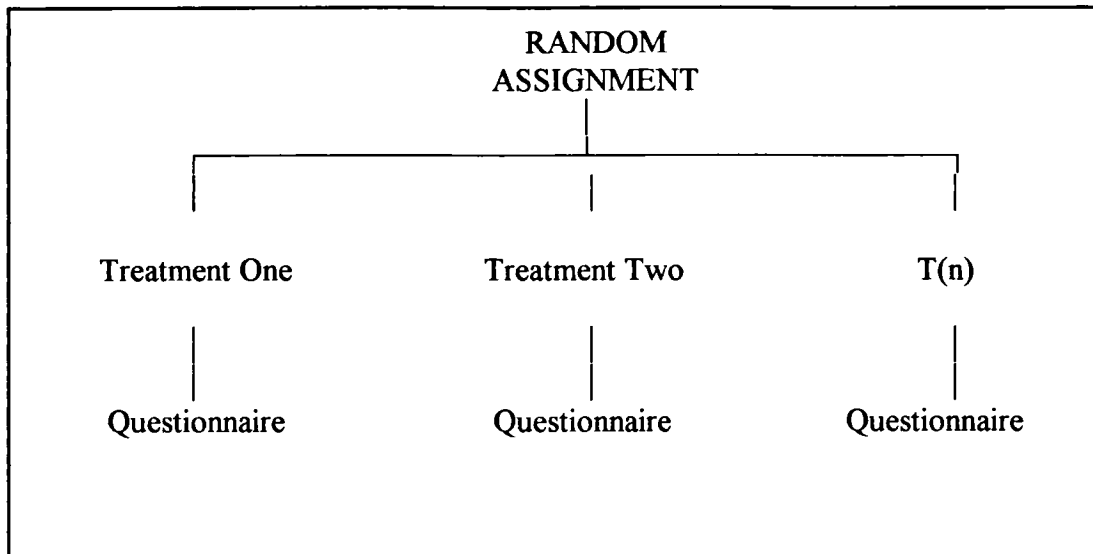


Figure 7 Experimental Overview - Allows for up to T(n) Treatments

The post-test-only design should ensure the internal validity of the experiment. However, the threats to external validity do need to be controlled and removed. There are four main threats to external validity:

1. Reactive or interaction effects of testing. This should not be a problem since the experiments will not be using a pre-test. The subjects will not therefore be sensitised to the experimental treatment.
2. Interaction of selection and experimental variables. This still remains an issue which the design will control by using a range of learners drawn from a variety of ability groups. The system will be sited in the normal learning environment of the subjects. The program will then randomly assign them to the version they will use.
3. Reactive effects of experimental arrangements. This was a possible problem, but the structure of the experiment eliminates this threat by two methods. Firstly, by carrying out the experiments in a subject's own environment (not in a laboratory) and secondly by the subjects' essentially undergoing a piece of training and not being aware of the experimental treatment.

4. Multiple treatment interference. This should not be a threat since the subjects will only have one experimental treatment applied to them during the experiment. The interference arises when subjects experience multiple treatments which are applied sequentially. Nevertheless, the order of the material which a subject encounters working through the tutorial may cause this effect.

Grabinger and Albers (1988) designed an experiment to investigate the effects on learning of different screen designs amongst 140 students. Their approach serves to illustrate the experimental design for this research. Their subjects were drawn from three schools in Lincoln, Nebraska and consisted of 71 females and 69 males who were all computer users. All the subjects were volunteers who had received parental permission to take part in the experiment.

Two educational programs were written - Order and Orbit. Each was available in several versions but all with identical content, questions and learning activities. The text was designed in accordance with the criteria published by Hartley (1980). It was therefore double spaced, left justified, maximum line length of 8 to 10 words, with left and right margins and contained only a single concept on each screen.

The researchers sought to measure three dependent variables. These were average time per screen (total time taken divided by the number of screens), immediate recall and delayed retention. The average time spent on each screen was intended to indicate if a student was encouraged to spend time studying, organising, reflecting or memorising material on the screen.

The procedure for the experiment was limited by the fact that the students could only be released for twenty minutes each. The students were selected by their teacher and assigned to the programs in an ordered manner to ensure a balance of subjects for each treatment. The students used the program in a special room with a researcher present who explained only the basic keys needed to help the students begin. Once the student had completed the version, they were asked to take a post-test of nine questions and a retention test was also taken two weeks later.

An examination of their experimental design, which aims at investigating a similar area to this research, should inform the validation of the research design.

Grabinger and Albers' (1988) experiment had a number of weaknesses. The subjects were not randomly selected and were allocated to different experimental versions. The experiments were undertaken in a special (laboratory) environment and a fixed time limit was given for each subject and version of the experiment. These weaknesses cast doubt on both the external and internal validity of the experiments (external validity: interaction of selection and reactive effects of experimental arrangements; internal validity: selection of subjects).

This analysis would seem to cast doubts over the generalisability of the experimental results of Grabinger and Albers' (1988) work. The experimental design for the shell should be significantly more generalisable due to the random allocation of subjects and the use of the experimental material in a familiar learning environment. The experiment was carried out in a learning resource centre.

The variables of this experiment are similar to those chosen by Grabinger and Albers (1988) but they have been improved especially with regard to the measure of the average time spent on each screen. The experimental design shell increases the effectiveness of this measurement by capturing the time spent on each screen and also by mapping the route each learner takes through the material. It therefore allows for more detailed analysis of the influence of the screen design on the individual learning.

Methodological Issues (Colour Experiment)

The Colour experiment was also designed to test the methods to be employed during the research. The key factors to be investigated were:

- (1) post-test-only design;
- (2) design of tutorial;
- (3) automatic recording of subject responses;
- (4) technical issues; and
- (5) robustness of software.

The outcomes of the Colour experiment contributed to the design of the Graphics experiment. Figures 8 and 9 illustrate the overall and tutorial structure used in the Colour experiment.

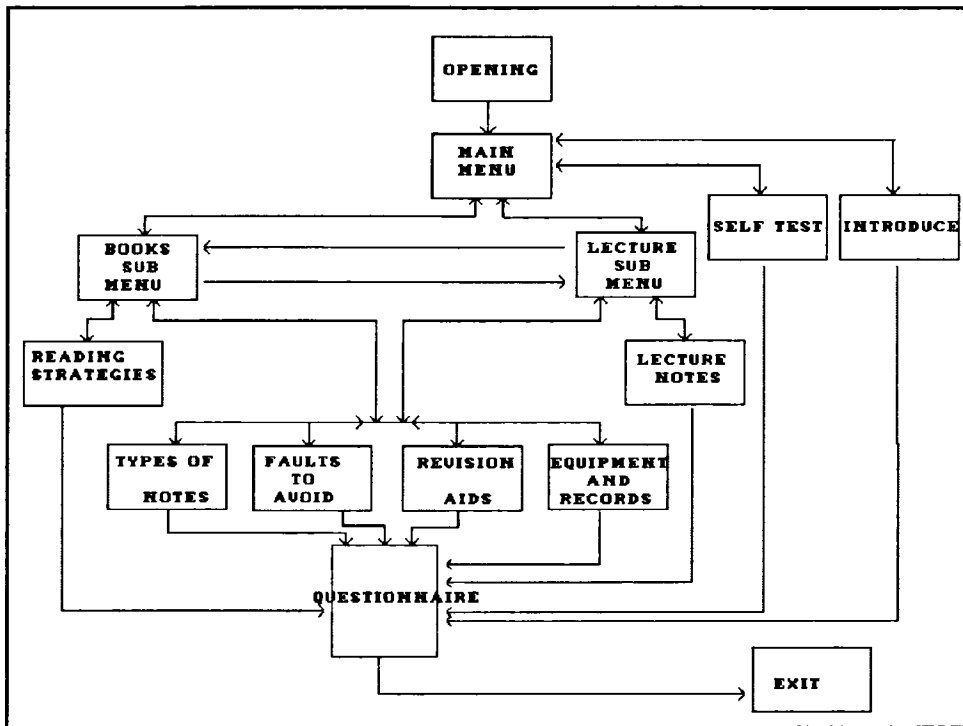


Figure 8 Overall Structure (Colour Experiment)

3.3 Design of the Graphics Experiment

The design of the Graphics experiment was again based on the post-test-only control group design. The subjects, however, were not randomly assigned to the experimental treatment since the same subjects were used in each experimental treatment. This has a major advantage over using different subjects in that it eliminates the individual differences which the subjects bring to each treatment. Table 27 compares the advantages and disadvantages of using the same and different subjects. These individual differences may well influence how subjects tackle the experimental task.

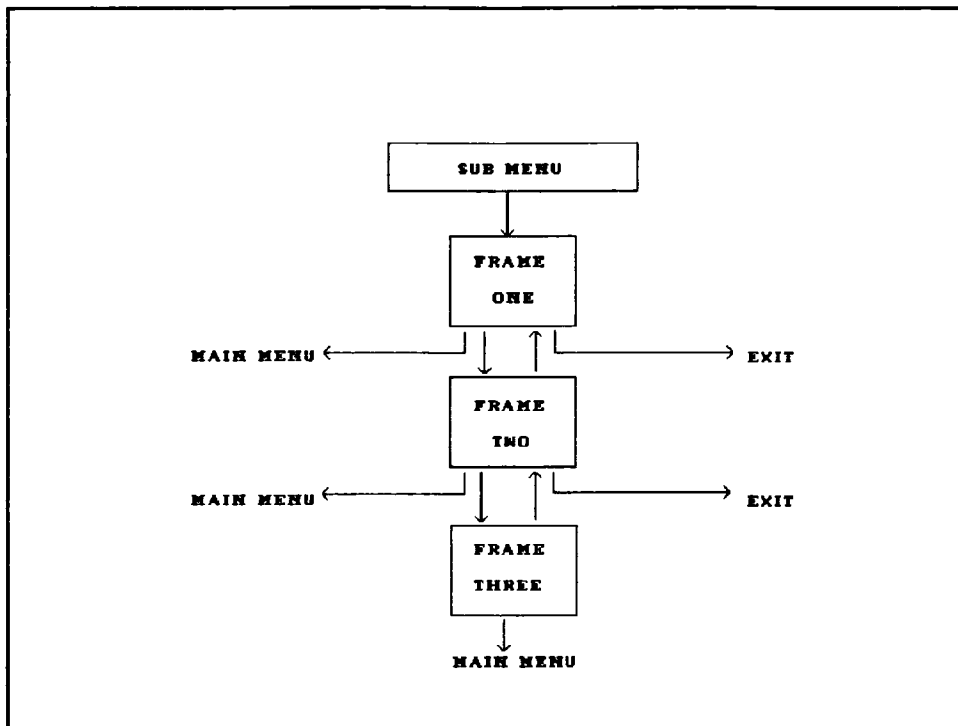


Figure 9 Tutorial Structure (Colour Experiment)

The individual subjects were free to choose the order in which they tackled each treatment. This would reduce and perhaps eliminate the effects of the subjects having to work through the treatments in a set order. The subjects could choose to start with the tutorial on 'Computer Input'. In making the choice, they had no information from the system giving them any indication of the nature of the screen design in that section of the tutorial. Within each section of the tutorial there were a range of screen design treatments:

- (1) text-only;
- (2) structured text;
- (3) full-screen representative graphics;
- (4) part-screen representative graphics
- (5) full-screen analogical graphics;
- (6) part-screen analogical graphics; and
- (7) full-screen logical graphics.

Table 27 Advantages and Disadvantages of using the Same or Different Subjects

| Subjects | Advantages | Disadvantages |
|--------------------|---|--|
| Same Subjects | Eliminates individual differences between experimental conditions | <p>Problem of order effects because subjects have to carry out the conditions in a certain order.</p> <p>Cannot be used when subjects have to be different (eg men or women). May result in each subject having to perform a lot of experimental conditions.</p> |
| Different Subjects | No order effects | Different subjects introduce an unknown amount of individual difference in the performance of the different experimental conditions. |

Due to the effects of the subjects choosing their own routes through the material and each tutorial containing a range of screen designs, the problem of ordering should have been minimised. Figures 10 and 11 illustrate the overall and tutorial structure of the Graphics experiment, respectively. The double-headed arrows in figure 10 leaving each of the modules 1 through 10 indicate that the tutorial allows learners the choice of many routes through the learning material. The arrows pointing into and out of Frame Four in figure 11 show that tutorial modules are not limited to four frames and the structure illustrated is repeated. Each tutorial module contains different numbers of frames.

The personal characteristics of the experimenter, such as age, gender and anxiety for the experiment to work may influence the behaviour of the experimental subjects. For example, the experimenter's briefing of the subjects may effect their behaviour during the experiment. This was reduced by standardising the briefing, providing each subject with an identical set of written instructions.

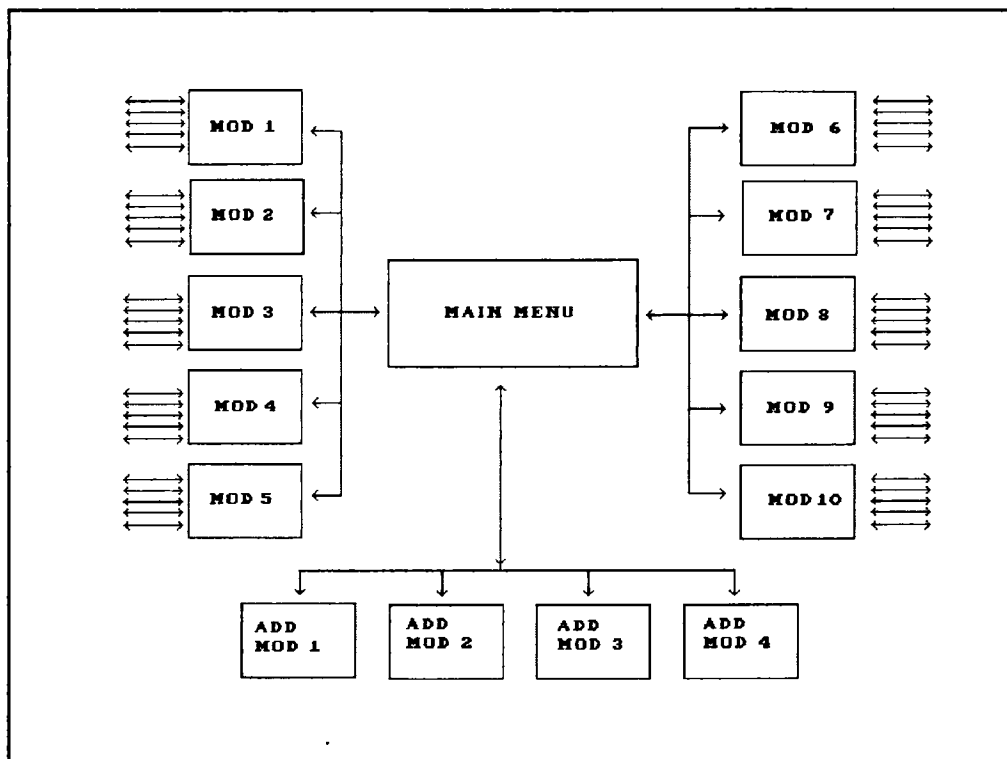


Figure 10 Overall Structure (Graphics Experiment)

The subjects carried out the task in different environments, therefore there was no consistency of environmental conditions. However, the shell recorded the subject's use of the material, making it straightforward to identify behaviour such as leaving the system switched on while the subject was occupied on other activities. Abnormal effects could thus be identified.

Robustness and Operation of Graphics Experiment

Since the Graphics experiment involved the posting of the software to the subjects and relied on the system operating correctly, it was tested on a small scale prior to the start of the experiment. Eight employees of the Employment Department's Learning Technology Unit were asked to test the system by carrying out the experiment. In addition to working through the system and completing the questionnaire, the subjects were asked to note any errors in the material.

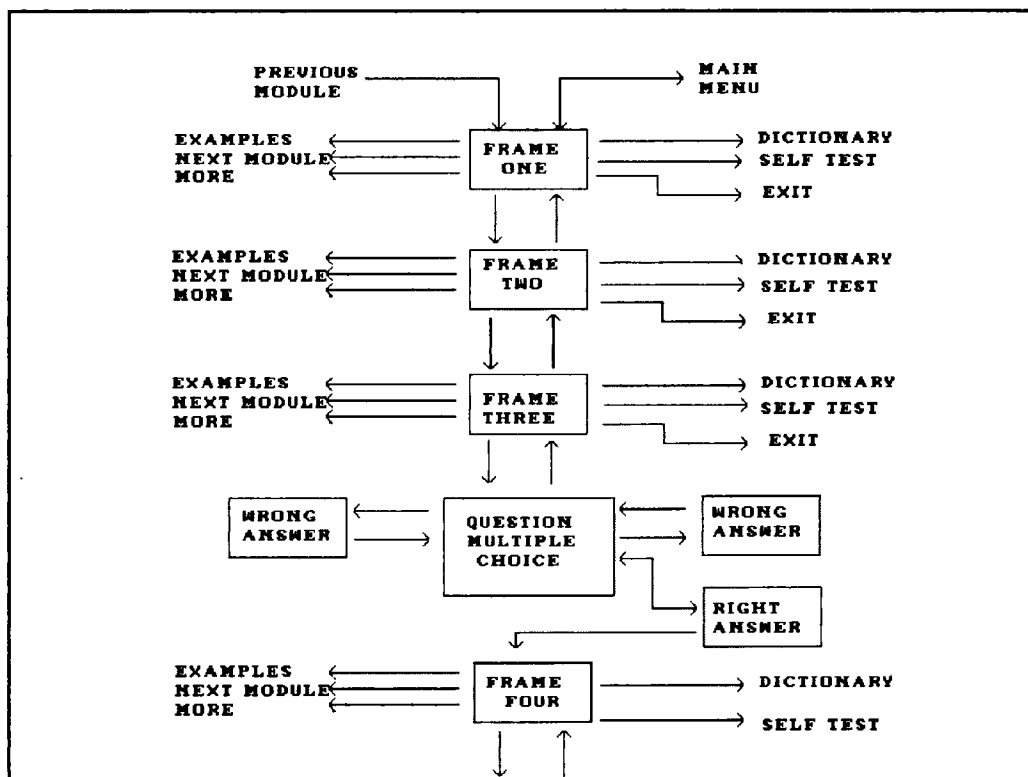


Figure 11 Tutorial Structure (Graphics Experiment)

Virzi (1992) reported that 80% of software usability problems were detected with four or five testers. His experimental results would predict that eight testers would detect over 90% of errors in the human-computer interface and that the most severe problems would be detected by the first 2 or 3 testers.

Examples of errors found by the testers included:

- (1) spelling mistakes;
- (2) incorrect instructions;
- (3) errors in the operation of the software; and
- (4) ambiguous statements.

The list of comments from the testers was used to amend the system and the instructions for its use. The testers were asked to verify the changes once they were made. The operation of the software during the experiment supported the pre-experiment tests with only one disk being returned by a subject who was unable to run the software.

3.4 Design of Colour Experimental Shell

The basic version of the tutorial was provided by version 0 which used three colours - black, white and grey. The display of information was consistently presented in the same positions on each frame with the title of each module and page number shown at the top of the screen in a rectangular box. Tutorial information was presented on the middle area of each frame with a justified left margin and double spaced text. The control menu was right justified and displayed at the bottom of the frame.

The addition of colour was redundant since the spatial positioning of text and the presentation of titles and menus in boxes were already providing the cues for the users. The use of a colour to highlight the title was thus an extra cue to draw the attention of a user to the title, tutorial information or control options. It was not providing new signposts for a user, merely reinforcing the ones already present. An example of a screen display from the Colour experiment (version 2) is shown in figure 12.



Figure 12 Menu Screen (Colour Experiment)

Measures

The Colour experiment to investigate the use of colour in computer-based learning materials used a variety of measurements. These were:

- (1) mean time spent on a tutorial frame;
- (2) mean time spent on a question frame;
- (3) mean time spent on a menu frame;
- (4) total number of frames used;
- (5) route taken through the computer-based learning module; and
- (6) a comprehension test.

3.5 Design of Graphics Experimental Shell

The Graphics experimental shell was designed to assist the investigation of the use of graphics in computer-based learning.

A range of factors were considered:

- (1) the use of graphics in computer-based learning;
- (2) the comparison of different types of graphic images in computer-based learning;
- (3) the use of different sizes of graphic images;
- (4) the complexity of the tutorial displays;
- (5) the effects of a learner's prior knowledge of the subject; and
- (6) a learner's use of additional modules not directly in the body of the main tutorial.

The major difference between the two experimental shells was that in the Colour experiment, the shell was used to produce six different versions of the computer-based learning material, whereas in the Graphics experiment, rather than produce a range of different versions, only one version was used. Within this single system, the different types of graphic image and other features were presented and subjects were free to access each component as they worked through the material. The comparison was thus between the same subjects using different elements of the material rather than different subjects using different versions of the material as in the Colour experimental shell. The outcomes of the Colour experiment were used to influence the design of the structure, software and tutorial of the Graphics experiment. Figure 13 provides an example of a screen display from the Graphics experiment.

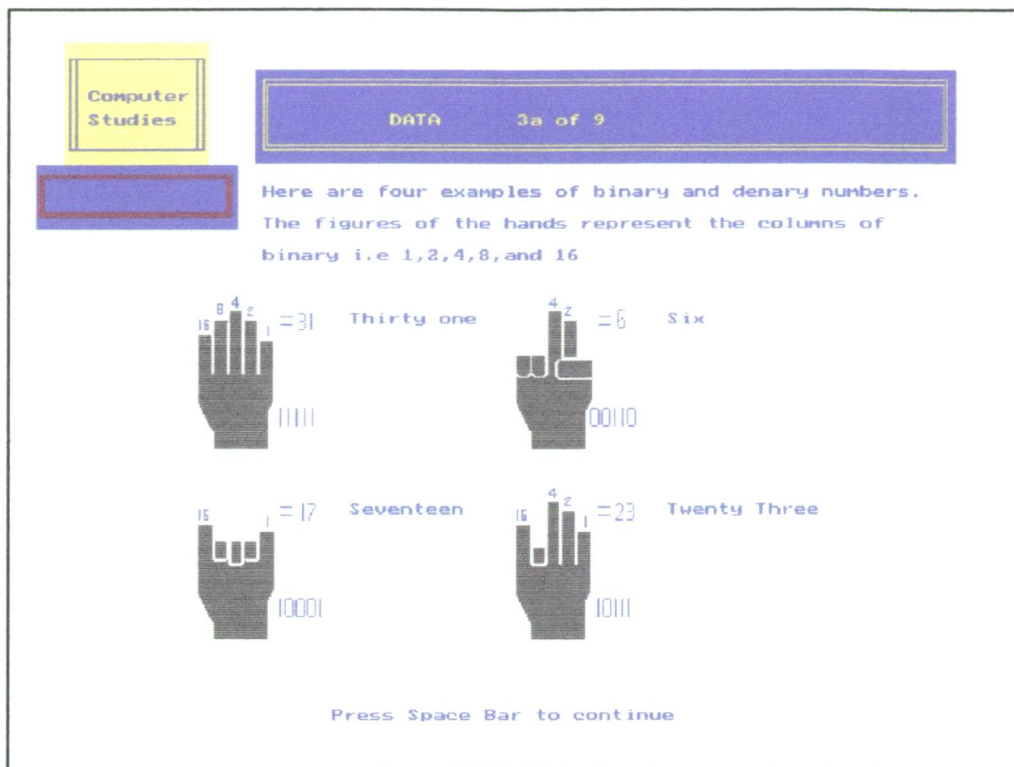


Figure 13 Tutorial Screen (Graphics Experiment)

Many computer-based learning experiments have been criticised on the grounds that they were confounded in not making realistic comparisons (Clark, 1983, 1985; Keane et al, 1991). The resulting changes identified in the experiments may be attributed to the uncontrolled effects of the different instructional methods, content or simply novelty (Clark, 1983, 1985). It is therefore critical to achieve a true comparison. In using computer-based learning materials, learners are free to choose the route they will follow and to interact with a range of screen displays, not simply with a single design. The survey of courseware undertaken during the research revealed that material did not use a single screen design but a range (for example, part and full-screen graphics, icons and menus and different types of image). The influence of a particular variable needs to be studied in the context of its interaction with the other screen design variables or the results obtained may well be flawed since they would not provide a realistic comparison. The Graphics experiment was designed to offer this realistic comparison with material used in the education and training worlds.

The Graphics experimental shell had twelve sections providing seven tutorial

elements (Computer System, Input, Output, Computer Development, Data Representation, Storage and Processing), a quiz, a dictionary of terms, examples of applications cross-referenced to the tutorial elements, information extending the tutorials and the exit system. The design of the Graphics shell was based on the evidence gathered during the use of the Colour system. All the elements were freely available to users who were required to choose their next move at each display. Table 28 shows the range of screen displays used in the Graphics experiment. Four additional elements were also available for the users. These are:

Frame 400 Dictionary which appeared as a pop-up text window. The window could be varied in size under the control of a user.

Frame 500 Example which was a series of relevant examples available as a pop-up window. A user could vary the size of the window.

Frame 600 More Information was a module which provided extra information on a given subject in a pop-up window. A user was able to control the size of the window.

Frame 700 Exit was a pop-up window which appeared once a user had chosen the exit option. It simply asked the user to confirm his or her decision to exit or return to the tutorial.

Table 28 Screen Designs Graphics Experiment

| Frame Number | Element | Feature |
|--|--|---|
| 2 3 to 5 6 7 to 8 | computer system computer system computer system computer system | part-structured text text part-structured text text |
| 10 to 12 13 14 15 to 23 | input input input input | text full-structured text part-structured text text |
| 25 to 31 | computer development | text |
| 33 to 34 35 36 to 40 | data representation data representation data representation | text full-structured text text |
| 42 to 46 47 | storage storage | text full-structured text |
| 50 to 56 | processing | text |
| 59 to 69 | output | text |
| 100 101 102 103 104 105 | computer system computer system computer system computer system computer system computer system | four part-screen representational graphics full-screen logical graphic three part-screen representational graphics full-screen logical graphic two part-screen representational graphics one part-screen representational graphic |

| | | |
|------------|----------------------|---|
| 106 to 113 | input | representational graphic in a pop-up window |
| 114 | input | two part-screen representational graphics |
| 115 to 116 | input | one part-screen representational graphic full-screen logical graphic |
| 117 | input | half-screen logical graphic |
| 118 | input | two part-screen representational graphics |
| 119 | input | two part-screen representational graphics |
| 120 | computer development | four part-screen representational graphics |
| 121 | computer development | one part-screen analogical graphic |
| 122 | computer development | three part-screen representational graphics |
| 123 | computer development | full-screen representational graphic |
| 124 to 126 | computer development | one part-screen representational graphic |
| 127 | data representation | four part-screen analogical graphics |
| 128 | data representation | one part-screen analogical graphic |
| 129 | data representation | six part-screen analogical graphics |
| 130 | data representation | four part-screen analogical graphics |
| 131 | data representation | two part-screen analogical graphics |
| 132 | data representation | four part-screen analogical graphics |

| | | |
|------------|------------|---|
| 133 | storage | two part-screen representational graphics |
| 134 | storage | two part-screen analogical graphics |
| 135 | storage | three part-screen representational graphics |
| 136 | storage | one part-screen representational graphic |
| 137 to 138 | storage | full-screen representational graphic |
| 139 | storage | four part-screen representational graphics |
| 140 | processing | one part-screen analogical graphic |
| 141 | processing | three part-screen analogical graphics |
| 142 | processing | two part-screen representational graphics |
| 143 to 144 | processing | one part-screen representational graphics |
| 145 | processing | full-screen representational graphic |
| 146 | processing | full-screen logical graphic |
| 147 to 148 | processing | one part-screen representational graphic |

| | | |
|------------|----------------------|---|
| 149 | output | full-screen logical graphic |
| 150 | output | full-screen representational graphic |
| 151 | output | three part-screen representational graphics |
| 152 | output | one part-screen representational graphic |
| 153 | output | two part-screen representational graphics |
| 154 | output | one part-screen representational graphic |
| 155 | output | full-screen logical graphic |
| 156 | output | one part-screen representational graphic |
| 157 | output | three part-screen representational graphics |
| 158 to 159 | output | two part-screen representational graphics |
| 200 | computer system | full-screen representational graphic |
| 201 | input | full-screen representational graphic |
| 202 | computer development | full-screen representational graphic |
| 203 | data representation | full-screen analogical graphic |
| 204 | storage | full-screen representational graphic |
| 205 | processing | full-screen representational graphic |
| 206 | output | full-screen representational graphic |

| | | |
|-----|------------|--|
| 207 | dictionary | full-screen representational graphic |
| 208 | quiz | full-screen representational graphic |

Measures

The experimental measures used in the Graphics experiment were:

- (1) Mean time spent on a text frame;
- (2) Mean time spent on a graphic frame;
- (3) Mean time spent on a structured text frame;
- (4) Mean time spent on a representative graphic frame;
- (5) Mean time spent on a full-screen representative graphic frame;
- (6) Mean time spent on a part-screen representative graphic frame;
- (7) Mean time spent on a one part-screen representative graphic frame;
- (8) Mean time spent on a two part-screen representative graphic frame;
- (9) Mean time spent on a three part-screen representative graphic frame;
- (10) Mean time spent on a four part-screen representative graphic frame;
- (11) Mean time spent on a full-screen logical graphic frame;
- (12) Mean time spent on a full-screen analogical graphic frame;
- (13) Mean time spent on a part-screen analogical graphic frame;
- (14) Total time spent on tutorial;
- (15) Total number of frames accessed;
- (16) Route taken through the computer-based learning modules;
- (17) Comprehension test; and
- (18) Complexity.

Complexity

In comparing different screen designs, it would be useful to have objective measures of what appears on the screen. Several researchers have reported on the importance of density and grouping of information displayed. Their views support the conclusion that grouping the information displayed and making it less complex will result in easier identification and better comprehension (Dodson and Shields, 1978; Tullis, 1981; 1983; Triesman, 1982; Haubner and Newmann, 1986). There is a limit to reducing complexity since a blank screen would have zero complexity but convey little information. The solution lies with some form of optimisation of the information displayed. The key to an objective measure of complexity is a means of quantifying what appears on the screen. There have been several published:

- (1) visual acuity;
- (2) information theory;
- (3) simplified information theory; and
- (4) groupings.

By combining these different pieces of work it is possible to produce a measure of complexity which can be used to compare different screen displays. There are three types of screen which form part of computer-based learning material:

- (1) text-only;
- (2) graphics-only; and
- (3) a combination of text and graphics.

Text-Only

Tullis (1983) measured screen complexity based on the work of Bonsiepe (1968), who proposed a method of measuring the complexity of typographically designed pages through the application of information theory. Table 29 shows a method of calculating complexity based on Tullis's work.

Table 29 Complexity Measure

1. Draw a rectangle around each field on the screen including captions, data and titles .
2. Count the number of fields and horizontal alignment points (the number of columns in which a field, inscribed by a rectangle, starts).
3. Count the number of fields and vertical alignment points (the number of rows in which a field, inscribed by a rectangle, starts)

Example

22 fields
6 horizontal points
20 vertical points

48 Complexity

The Tullis-based measure of complexity does, however, take no account of the size of each field. The capacity of the eye to resolve details is called visual acuity. It is the phenomenon that causes an object to become more distinct when individuals turn their eyes towards it and to lose distinction as they turn their eyes away from it. It has been calculated that when viewing a computer screen, the optimum size of object is 1.67 inches. This translates to a field of 15 characters wide and 7 rows deep (Bouma, 1970; Danchak, 1976) and is shown in figure 14. This is the size of an object where characters at the edge of the area have a 50/50 chance of being correctly identified. For a single line of text, the maximum optimum size of field is 15 characters. A field of text of up to 15 characters would thus have a complexity score of 1, a field of 16 to 30 characters would score 2 and a field of 31 to 45 characters would score 3.

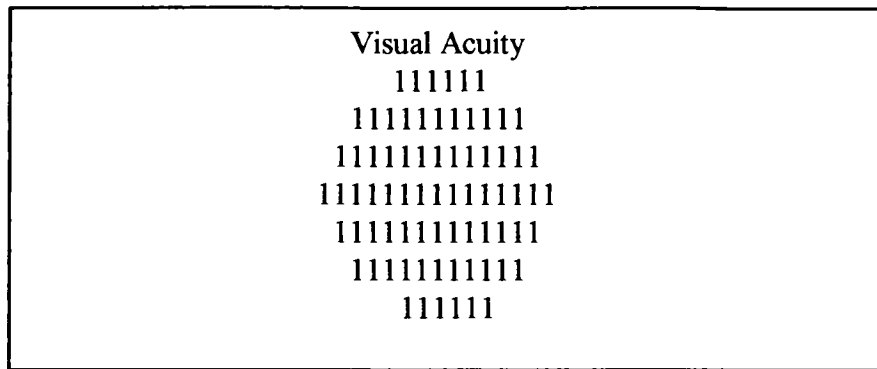


Figure 14 Visual Acuity

Graphics-only

The grouping of information has been shown to aid recall (Card, 1982). It can also add structure to a display and improve search times in identifying items of information (Dodson and Shields, 1978). An individual graphic could be considered as a single group of information. However, it is necessary to differentiate between full-screen and part-screen graphics. The key factors are the size and number of graphic objects. The measure of complexity chosen is based on these two factors. The optimum size of an image is taken to be that given by the value of visual acuity (1.67 inches). Combining this approach with that given for text-only, it is possible to measure complexity thus:

- (1) count the number of horizontal alignment points;
- (2) count the number of vertical alignment points;
- (3) count the number of graphical objects; and
- (4) for objectives less than the maximum optimum size (1.67 inches or 15 characters) count 1. For objects less than 30 characters (or 3.34 inches) but more than 15 characters (or 1.67 inches), count 2.

This approach will produce a total which is the measure of complexity of a graphics-only display.

Text and Graphics

To combine the factors in order to measure complexity for displays containing graphics and text requires the following steps:

- (1) draw a rectangle around each field on the screen including captions, data and titles;
- (2) count number of fields;
- (3) count number of graphic objects;
- (4) count number of horizontal alignment points;
- (5) count number of vertical alignment points; and
- (6) calculate size of fields and graphic objects.

This approach will produce a total which is a measure of complexity of a display containing both text and graphics.

3.6 Experimental Resources

The content of the two tutorials was designed to meet the standard of computer-based learning which is currently available to education and training organisations. The Colour experiment was based on a tutorial on study skills, while the Graphics experiment used a tutorial on computer studies.

Design of the Software Shell

The majority of research into screen design has been based on the use of single screen displays. Only a few experiments have used multiple screen displays and not all of these experiments have used a complete learning package. They have thus not considered the interaction of displays or learning material. The design of the software shell was based on multiple displays and contained learning materials.

The basic requirements of computer-based learning are given by Beevers et al (1991) and are:

- (1) displaying information;
- (2) controlling the route through the software;
- (3) controlling the rate of display;
- (4) obtaining input from a user;
- (5) sending messages and giving on-line help;
- (6) making use of the animation features of the microcomputer display; and
- (7) assessing data files to log student progress.

This list of elements does cover the basic requirements of computer-based learning materials but is based on the needs of mathematics materials and microcomputers in 1985. In the last few years, the facilities of microcomputers have grown extensively and in particular in the provision of colour and graphics displays.

The list now needs to be extended to:

- (8) making use of the graphics features of the microcomputer display; and
- (9) making use of the colour features of the microcomputer display.

The software shell for the research needs to include the provision of all these features. It must also be able to allow for the key variables of screen design being manipulated easily and in particular allow for the variables to be monitored and recorded. The shell must provide the features for several versions of the basic system to be created with the key variables changed to reflect the needs of the research.

Software Tools for Computer-Assisted Learning

One of the earliest decisions which had to be made was the choice of authoring software to be used in the development of the experimental shell. The basic choice was between a programming language and an authoring language or system. An

authoring facility would provide many of the tools required for computer-based learning such as multiple-choice questions, screen displays and different text types. However, authoring tools restrict the freedom of the author to some extent in order to provide these standard features. A programming language gives extensive flexibility but does not provide many of the standard features of an authoring language (Clarke, 1991). Bork (1987) states that the production of interactive learning materials requires the application of a programming language, as authoring languages and systems tend to restrict a designer. Dean (1993), in a review of authoring systems, makes the point that authors frequently need to use programming languages in order to achieve results that an authoring system is unable to supply. This research required flexibility, interaction, good structure, colour and graphics and therefore could not accept any restriction imposed by the authoring tools.

Turbo Pascal is a powerful and structured language which provides both colour and graphics facilities. It is widely used in both the academic (Leiblum, 1992) and commercial worlds and runs on IBM compatible computers which are the machines frequently used for computer-based learning. An IBM compatible computer was used throughout this research. Turbo Pascal has a range of third party toolboxes available which extends the language. The Technojock's Toolbox provides a number of routines which partially reduce the need to develop each part of the software from scratch. It thus reduces the gap between an authoring and a programming language without decreasing the flexibility of programming languages. During the research, the authoring software was upgraded several times as the producers released new versions of their packages. The following versions were used:

- (1) Turbo Pascal 4.0, 5.5, and 6.0 (Borland International, 1988; 1989; 1991);
- (2) TechnoJock's Turbo Toolbox 5.0 (TechnoJock Software, 1989);
- (3) TechnoJock's Object Toolbox (TechnoJock Software, 1991); and
- (4) Public Domain Graphics Functions (Public Domain Software Library, 1992a; 1992b).

Display of Text on the Screen

The screen of an IBM compatible computer operating in text mode is divided into 80 columns across the screen and 25 rows down. There is an alternative mode of 40 columns and 25 rows. However, the evidence provided by the review of research suggests that the 80 column and 25 row mode is the preferred learning mode. All text screens were thus provided in 80 column by 25 row mode.

The standard Turbo Pascal procedures of 'write and writeln' can be used to display text horizontally. Unfortunately, they do not allow for the precise positioning of text or for changing the colour of the text. In order to provide these, three additional procedures were used:

- (1) WriteAT(X,Y,Attr:byte;Str:string); positions text at co-ordinates X,Y in colour Attr;
- (2) WriteBetween(X1,X2,Y1,Attr:Byte;Str:string); centres text between the co-ordinates X1,Y1 and X2,Y1 in colour Attr; and
- (3) WriteCentre(Y,Attr:byte;Str:string); centres text on row Y in colourAttr.

The two standard and three additional procedures provided all the required control of the colour and positioning of text on the screen.

There was also the requirement to clear areas of the display in order to position new messages. Standard Turbo Pascal provides three procedures for erasing text. These are ClrScr, ClrEol and DelLine. ClrScr clears the screen or active window of all text, replacing it with the current background colour and moving the cursor to the top left hand corner. ClrEol clears the screen with the current background colour from the current cursor position to the end of the line without moving the cursor. DelLine clears the line containing the cursor, replacing it with the current background colour. Both ClrEol and DelLine operate within a defined window. In addition to these standard procedures were added three extra procedures to provide for greater control of erasing text.

These were:

- (1) `Clear(Attr:byte;Ch:char)`; this erases the entire screen and fills it with the character Ch;
- (2) `PartClear(X1,Y1,X2,Y2,Attr:byte; Ch:char)`; this is the same as `Clear` except only the rectangle X1, Y1 ,X2 ,Y2 is erased; and
- (3) `ClearText(X1,Y1,X2,Y2:byte)`; this clears all text from a rectangle X1, Y1, X2, Y2.

Colour

The experimental shell required good colour control for all aspects of the display.

The basic needs were:

- (1) menu and icon colour options;
- (2) title bar colour (foreground and background);
- (3) text area colour (foreground and background);
- (4) windows colour (background, foreground and border);
- (5) menu colours (foreground colour of highlighted pick and background colour of highlighted pick);
- (6) normal foreground colour;
- (7) normal background colour; and
- (8) foreground colour of box.

The approach adopted to control colours in the experimental shells was to give each colour element a variable name. This allowed the colour of the title text to be changed throughout the material by a single change to the value of the variable. The colours were allocated by the Initialise procedure which, in addition, randomly chose which version (Colour experiment) was presented to the user.

Menu

There are a range of approaches to user-control of computer packages. The main methods are menus, command languages, graphic interfaces (these are characterised by the use of a mouse and icons as links to other parts of the material) and combinations or variations of these main approaches. Menus are used extensively in computer-based learning since they are easy for inexperienced users and learning material is almost always used by students who are not familiar with the material. Menus present a clear list of labelled options for learners to choose from which even a computer novice can understand. The use of menus to provide control mechanisms for computer-based learning is the most prevalent (Schuerman and Peck, 1991a).

In recent years, graphic user interfaces have been used extensively in learning materials either on their own or combined with menus. The survey of thirty-four packages of computer-based learning material in use in the Employment Department's Learning Resource Centre discussed earlier confirmed that current learning material tends to use a main menu combined with a graphic interface. This was the control approach adopted in the Colour and Graphics experiments.

A main menu provides an overview of the material from where a student can choose to move to the major elements of the system. Within each element, a student uses the mouse or keyboard to point at icons which link to other components. Menus provide a structuring device for the learning material and are not simply a method of making choices. The order of the main menu, the links between menus and the division of options between menus is essentially the structure of the learning materials. The design of menus for computer-based learning has attracted only a few research studies and most of the available guidelines are based on intuition.

The main menu was developed as an open structure in which the number of choices could be easily increased or decreased using TechnoJock's Object Toolbox. The design of the menus used in the experiment was based on a consistent format with the choices left justified and organised into columns. This has been found to be the most efficient and effective by several researchers (Parkinson and Sisson, 1985; Bacs, Walrath and Hancock, 1987). The menus were presented in the same location on the screen and were distinguished from the remainder of the screen by presenting

them enclosed in a box.

The choices offered to the users were listed in a consistent manner which research evidence (Mitchell and Shneiderman, 1989) has shown users prefer and find quickest to use. When a user picked an option, the choice was highlighted with a bar of colour. Choices were clearly separated from each other by a number of techniques. These were:

- (1) numbering each item;
- (2) presenting choices in a consistent list;
- (3) giving the first letter of an option a different colour to the rest of the words; and
- (4) highlighting the choice in a different colour.

These devices allowed the menus in the experimental shell to follow the available research evidence and the current methods of commercial products. Figures 15 and 16 show the main and pop-up menus in the Graphics experiment.

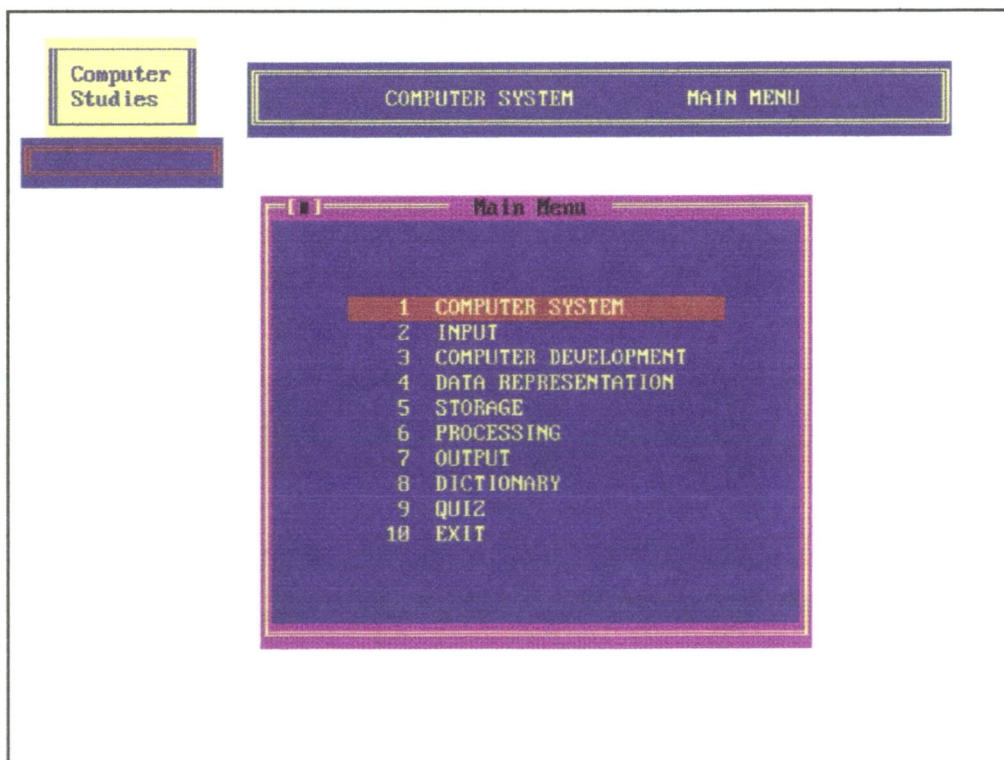


Figure 15 Main Menu (Graphics Experiment)

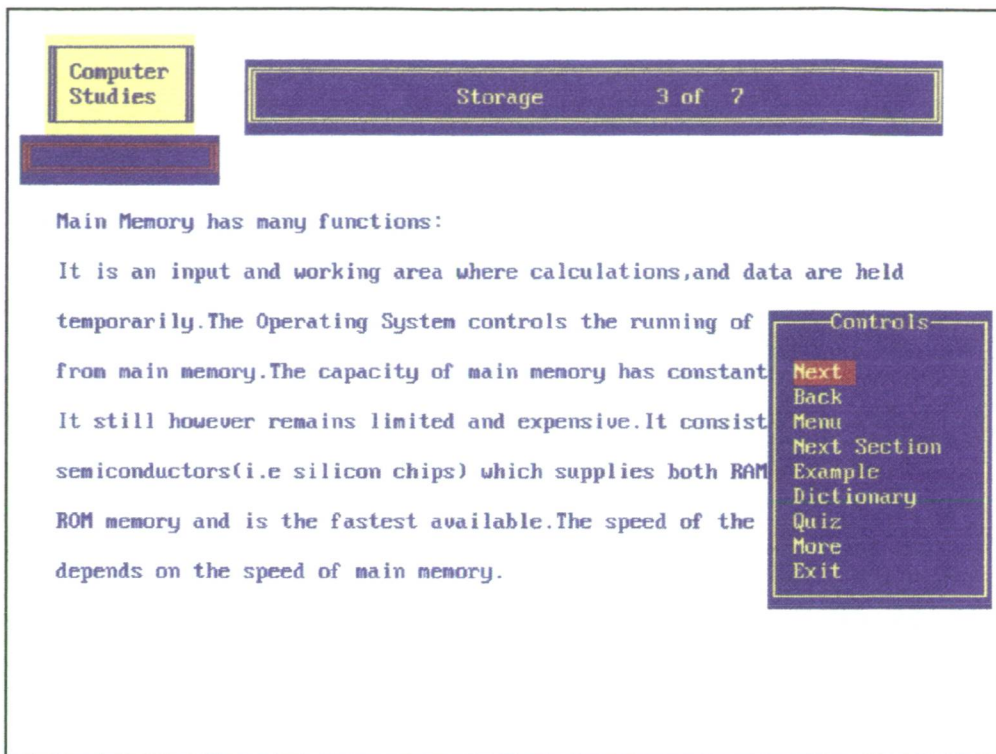


Figure 16 Pop-Up Menu (Graphics Experiment)

Graphics

IBM compatible computers can be equipped with a variety of video graphics adapter cards. Table 30 gives some examples of the IBM graphics standards (Jackson et al, 1994).

Each graphics adapter card supports a number of graphic resolution standards (for example VGA 320 x 200 and 640 x 480). For the purpose of these experiments, a VGA colour system was chosen using the 640 x 480 graphics standard which provides sixteen background and foreground colours. This provides high resolution with images which are clear and precise on a wide range of currently available computers. The trend over the last twenty years has been continuously to improve the resolution of screen displays.

Table 30 Examples of IBM Graphics Standards

| Type | Resolution | Colours | Total Palette |
|------|-------------|---------|---------------|
| CGA | 640 x 200 | 2 | 16 |
| | 320 x 200 | 4 | Pre-set |
| EGA | 640 x 350 | 16 | 64 |
| | 640 x 480 | 16 | 256 |
| | 880 x 600 | 16 | 256 |
| VGA | 640 x 480 | 16 | 262,144 |
| | 320 x 200 | 256 | 262,144 |
| SVGA | 1024 x 768 | 16/256 | 262,144 |
| | 1280 x 1024 | 16/256 | 262,144 |
| XGA | 1024 x 768 | 256 | 262,144 |
| | 640 x 480 | 65,536 | 262,144 |

The creation of graphical images requires both a considerable design and programming effort. Figure 17 shows the Turbo Pascal code for drawing a coloured image of a floppy disc on the screen.

Care has to be used in the choice and design of the graphical images in order to maximise the learning potential of the material and also to aid the experimental design. The productivity of the programming effort was improved by making a series of tools for the creation and storage of graphical images. The drawing tools produced images which could be stored as data files and read into the learning material when required. Two main tools were used, 'Iconnew' and 'Rmaster'. Some images produced using these tools are illustrated in figures 18 and 19. Both were based on public domain utilities which were enhanced to produce output compatible with Turbo Pascal. The time taken to produce a graphic image with these tools was far shorter than that taken to code an image in Turbo Pascal.


```

Procedure FloppyDisc; {Graphic of a Floppy Disc}
begin
  SetColor(1);
  Line(170,100,440,100);
  Line(440,100,470,130);
  Line(470,130,470,400);
  Line(470,400,170,400);
  Line(170,400,170,100);
  SetFillStyle(SolidFill,9);
  SetBkColor(9); {LightBlue}
  FloodFill(465,395,1);
  Rectangle(210,100,400,200);
  Rectangle(250,100,400,200);
  SetFillStyle(SolidFill,7);
  SetBkColor(7); {LightGrey}
  FloodFill(395,195,1);
  Rectangle(320,120,370,180);
  SetFillStyle(SolidFill,9);
  SetBkColor(9); {LightBlue}
  FloodFill(365,175,1);
  Rectangle(210,240,430,400);
  SetFillStyle(SolidFill,15);
  SetBkColor(15); {White}
  FloodFill(425,395,1);
  Rectangle(180,360,200,380);
  Rectangle(440,360,460,380);
  SetFillStyle(SolidFill,3);
  SetBkColor(3); {Cyan}
  FloodFill(185,365,1);
  FloodFill(445,365,1);
  Line(190,120,200,140);
  Line(200,140,195,140);
  Line(195,140,195,160);
  Line(195,160,185,160);
  Line(185,160,185,140);
  Line(185,140,180,140);
  Line(180,140,190,120);
  SetFillStyle(SolidFill,1);
  FloodFill(192,130,1);
end;

```

Figure 17 Turbo Pascal Floppy Disc Procedure

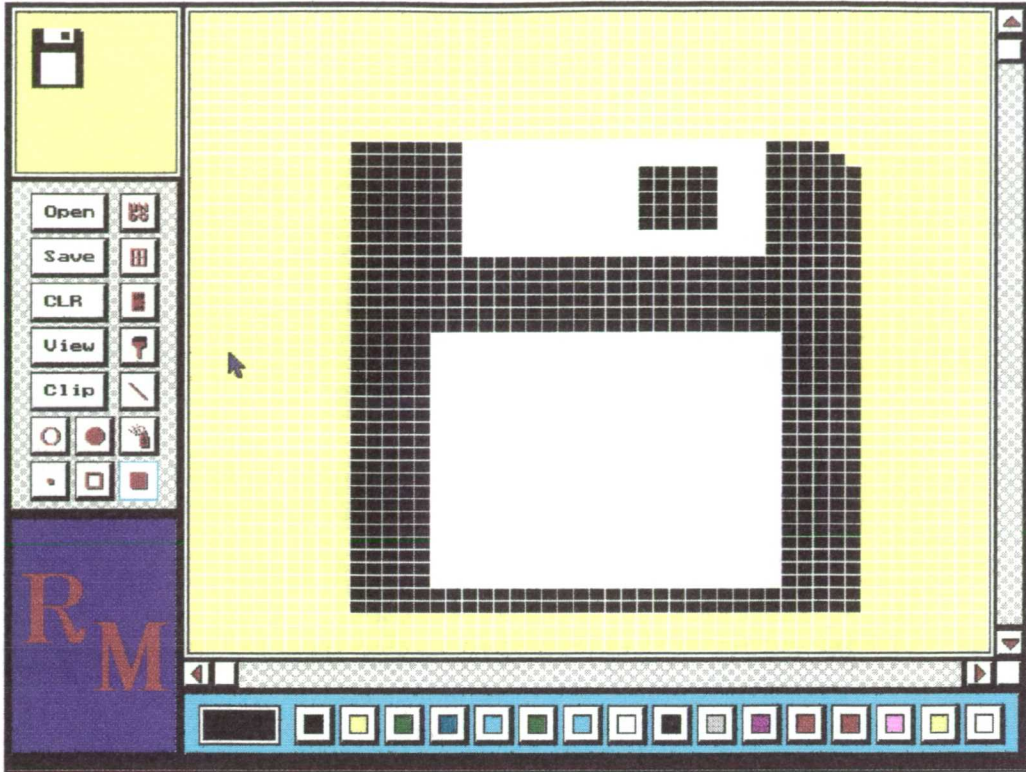


Figure 18 Rmaster Image of a Floppy Disk

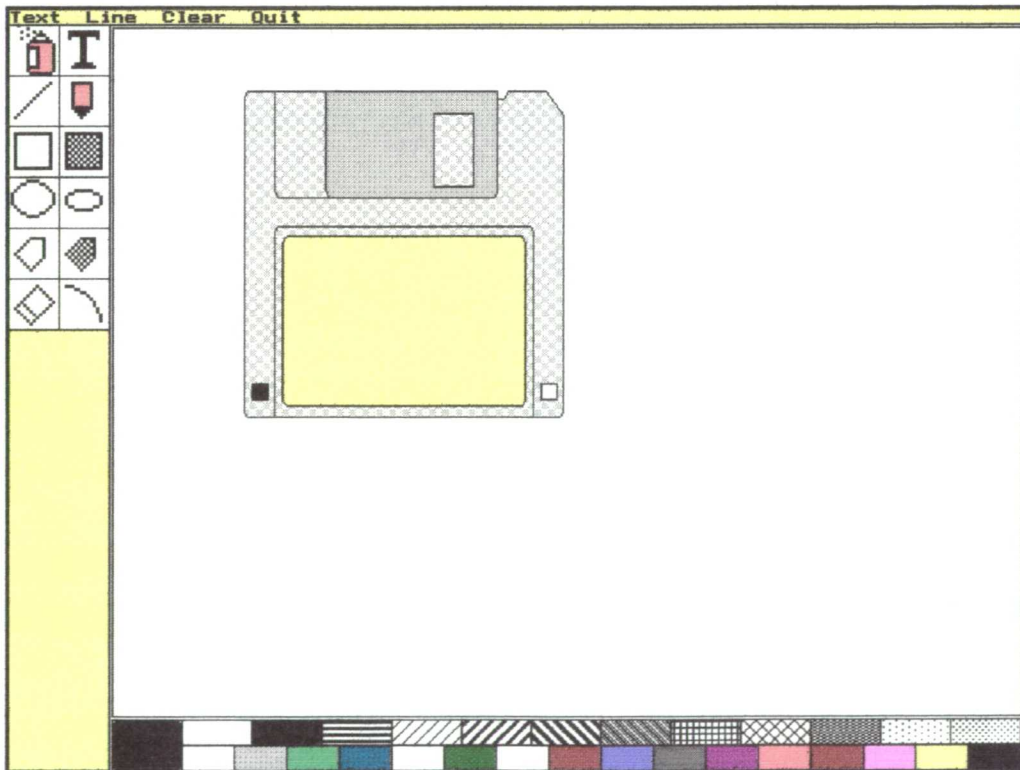


Figure 19 Iconnew Image of a Floppy Disk

Standard Turbo Pascal procedures provided to manipulate graphics do not allow for the precise positioning of images. In order to achieve this, three additional procedures were used:

- (1) `PixImage.Int('Chip.Img');`
- (2) `PixImage.Move(x,y);` and
- (3) `ReadXgf(Sales.xgf,x,y).`

These first two procedures were developed to handle the graphics file produced by Iconnew. The first procedure reads the graphics file into memory and the second positions the image anywhere on the screen. The third procedure works with image files produced by Rmaster; it reads the file into memory and then positions it on the screen. Figure 20 shows the `ReadXgf` procedure.

```
Procedure ReadXgf(var name:string;T,L:integer);
Var
  F   : File;
  Img : Pointer;
  Size : LongInt;

Begin
  Assign(F,name);{Graphic file must be in currentdirectory}
  Reset(F,1);
  Size:=FileSize(F);      {Get size of file}
  GetMem(Img,Size);      {allocate memory}
  BlockRead(F,Img^,Size); {Read file into memory}
  Close(F);
  PutImage(T,L,Img^,NormalPut); {display Image}
  FreeMem(Img,Size);     {Free Memory}
End;
```

Figure 20 ReadXgf (Graphics Procedure)

Initially, the Iconnew tool was used but problems related to file corruption led to the change to Rmaster. During this change, the efficiency of the shell was also improved (by replacing the two reading into memory and positioning procedures with a single procedure).

Student Files

Both Colour and Graphics experiments automatically tracked and recorded the subject's use of the learning material.

Colour Experimental Shell

In order to track and record the subjects' use of the experimental learning material, three procedures (figure 21) were written. The first, 'Collector', recorded a student's information on the route and time taken in an array. The second, 'Results', created a text file at the end of the experiment and recorded the contents of the array into the text file. The third served only to initialise the array.

This approach worked successfully during the Colour experiment but was inefficient in two ways. Firstly, information was held in memory throughout the experiment and thus could be lost entirely if a run-time error occurred. Secondly, an array uses a large amount of memory which could perhaps be better employed as part of the tutorial. The Graphics experimental shell was designed to remove these inefficiencies.

Graphics Experimental Shell

In order to track and record the subjects' use of the experimental learning material, two procedures were written. The first 'OpenFile' essentially opened a text file to allow the program to write information to the file (version of material, time taken and the number of frame).

```

procedure results (var Time:Total;i:Turns;Position:Map,Version:integer);
var
  TheFile:Text;
  Line:string[255];
  N,O:integer;
begin
  Assign(TheFile,'MYFILE.TXT');
  Rewrite(TheFile);
  Writeln(TheFile,'      Results of user'); Writeln(TheFile);
  Writeln(TheFile,'Version');
  Writeln(TheFile,Version); Writeln(TheFile);
  Writeln(TheFile,'No.   Time Taken   Repeats   ');
  Writeln(TheFile);
  For N:=1 to 300 do
  begin
    Writeln(TheFile,N,'   ',Time[N]:10:2,'   ',i[N] );
    Writeln(TheFile);
  end;
  Writeln(TheFile,'No.   Route Taken');Writeln(TheFile);
  For O:=1 to 300 do
  begin
    Writeln(TheFile,O,'   ',Position[O]:10);
  end;
  Writeln(TheFile); Writeln(TheFile,'***** End of File *****');
  Close(TheFile);
  MainQuestion;
  Halt;
end;      {*****End of results*****}
procedure Collector(var Hour, Minute, Second, Sec:Clock; j, page:integer;
  Temporary:real);
begin
  GetTime(Hour[j],Minute[j],Second[j],Sec[j]); Temporary:=0;
  Temporary:=(Hour[j]*3600+Minute[j]*60+Second[j]+Sec[j]/100)-
  (Hour[j-1]*3600+Minute[j-1]*60+Second[j-1]+Sec[j-1]/100);
  Time[page]:=Time[page]+Temporary;
end;
procedure Initiate(var Hour, Minute,Second,Sec: Clock; j,page: integer;
  Temporary:real);
begin
  i[page]:=i[page]+1; GetTime(Hour[j-1],Minute[j-1],Second[j-1],Sec[j-1]);
end;

```

Figure 21 Turbo Pascal Procedures Results and Collector

The second procedure 'AddToFile' was included in those sections of the programme from where experimental information needed to be recorded. This simple use of two procedures allowed considerable flexibility and ease of recording the experimental information. Student information was recorded onto the disc and thus could not be lost if a run-time error was encountered. No information was held in memory thus providing an efficient use of memory.

Questions/Tests

In all CAL software one of the most important elements is some form of test which allows the students the chance to assess their understanding of the subject and to reinforce the theory being presented (Beever et al, 1991). All forms of computer-based learning need to provide learners with a means of testing their understanding of the tutorial material.

Testing was provided in the Graphics experimental software in two forms:

- (1) a block of multiple-choice questions which the learners can choose to access whenever they wish; and
- (2) individual multi-choice questions which form part of the tutorial.

The structure of the questions was identical, with a choice of three different answers being offered to the learners. Each choice led to a feedback window which explained the reasons why an answer was right or wrong. The feedback thus provided a substantial component of the Graphics tutorials. In the original Colour experiment, test questions were only provided in a block but feedback for each question was provided in an identical way.

Windows

A window is an area of the screen (usually rectangular in shape) which is enclosed by a border and contains a particular message. It can be moved, sized and rendered

independent on the screen. Windows can be small, containing a short message or very large covering the entire screen. A display may contain one, two or more windows within its boundaries.

Windowing can best be regarded as a means of allowing a learner access to several pieces of information from different sources on the same screen. In commercial software products, windows are already extensively used. It is therefore essential that the experimental shell also used windows for the same purposes to allow the design of the display to be investigated. The experimental shell used windows for:

- (1) control menu window;
- (2) Exit window;
- (3) Dictionary window;
- (4) overlay windows in the questions;
- (5) Examples windows;
- (6) More Information windows; and
- (7) tutorial windows.

The dangers of using windows were expressed by Benest and Dukic (1989). They stated that a computer display could become chaotic by providing too many windows with too many options to resize or move. Users would spend a great deal of time choosing different options or searching for information on one of the windows being displayed. To avoid this possibility, the number and properties of the windows in the experimental shell were restricted. All the windows were restricted so that only one of each type could be overlaid on to the main screen at any one time. The windows automatically closed once users indicated they wished to move on. Only a limited range of the windows offered extended properties, such as the ability to change size and have text scrolling within them. These were the windows giving access to additional features, that is, the Dictionary, Examples and More Information modules.

Users were provided with an initial window in which to choose between options. The options were: to enter the item they wished to find such as a term in the

dictionary or a numbered example; to see the entire contents presented as continuous text through which a user could scroll; or to return to the main system. If a user opted for the entire contents list, an additional choice was offered - to view it in a small window or a window occupying the whole screen.

Exit

The Colour tutorial did not allow users accidentally to choose to exit the tutorial and return. Users were committed to leave as soon as they had made the choice to exit. The Graphics experiment presented users with the option to return. This brings the experiment into line with the normal practice of commercial learning materials. It also provides the opportunity to check if the screen display contributes to the accidental choice of the Exit option.

The Exit option was provided consistently throughout the shell as the last choice of the main and control menus. Once a learner had made the choice, an overlay window appeared in the centre of the display (figure 22 illustrates the Exit window). Learners were then given the opportunity to confirm their choice of either leaving or returning to the tutorial.

Structure of Software

The overall structure of the software is shown in figures 8 and 10. Unfortunately the overviews presented in these figures tend to hide the complexity of the experiments. In view of this, figures 23 and 24 reveal the underlying structure of two components of the experiments by showing the structures of the Question and Input modules, respectively.



Figure 22 Exit Window

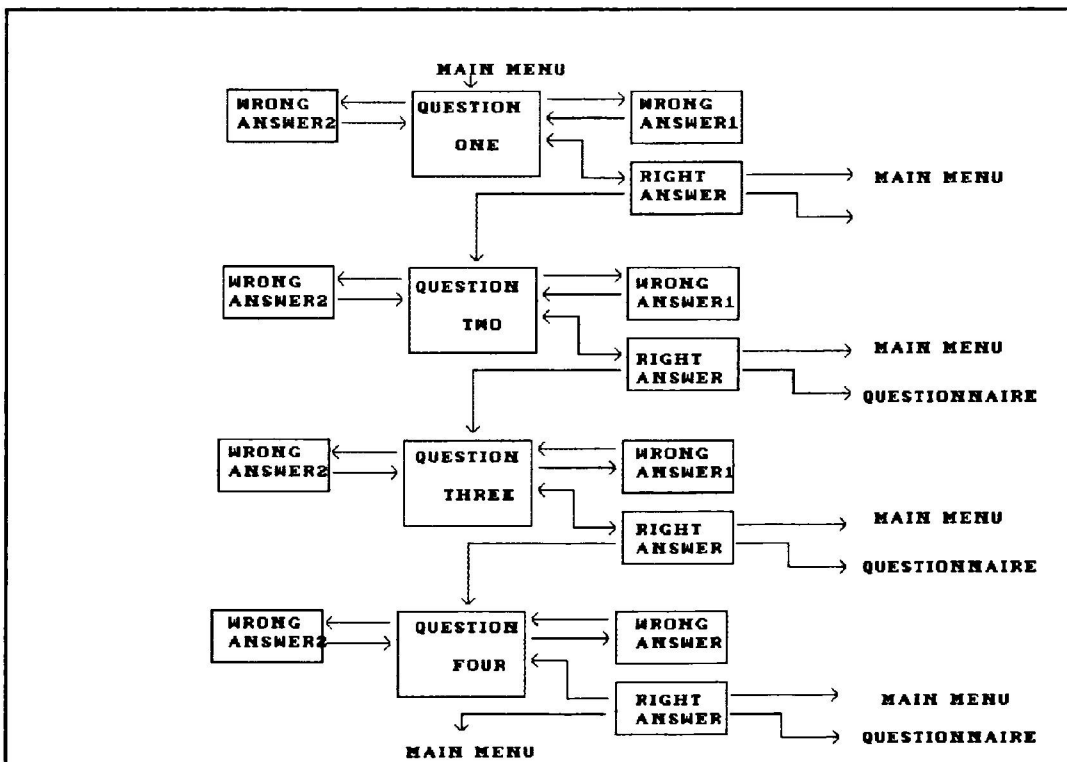


Figure 23 Structure of Question Module (Colour Experiment)

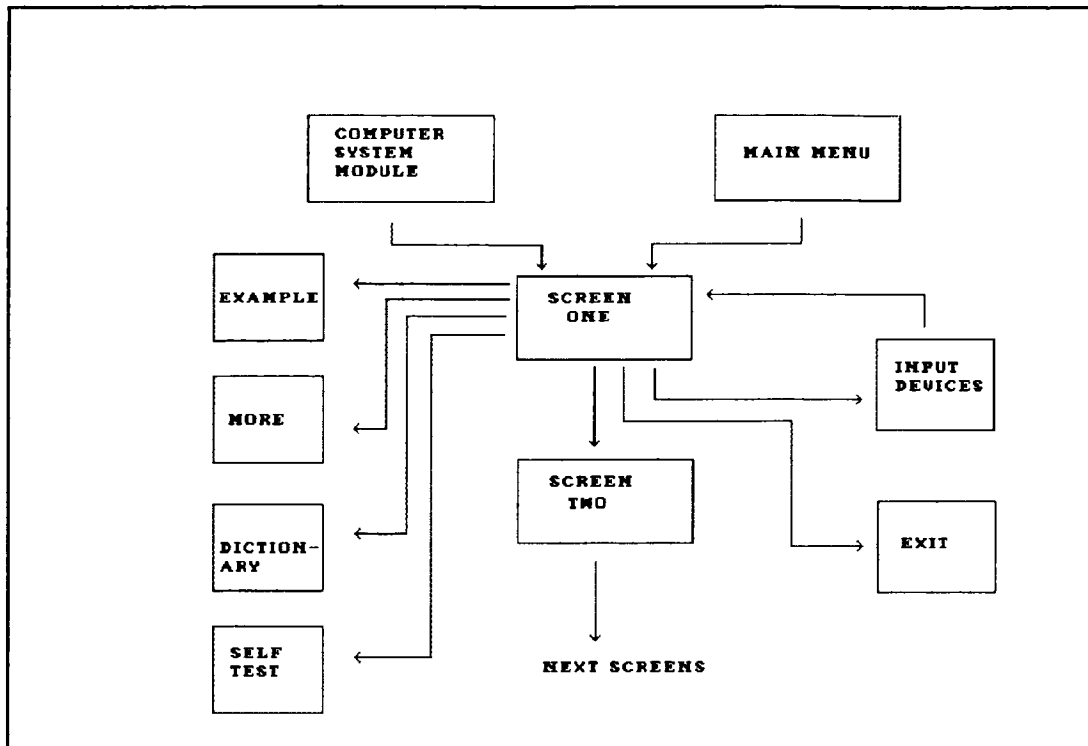


Figure 24 Structure of Input Module (Graphics Experiment)

Units

Units are the basis of modular programming in Turbo Pascal. They are used to create libraries of procedures which can be included in programs or to divide large programs into logically related modules. Units served both these purposes within the experimental shell program. Library units were provided in the standard units supplied with Turbo Pascal and the TechnoJock's Toolboxes. The experimental shell code was divided into unit modules to simplify the large program.

The Colour experimental shell was constructed using Turbo Pascal 5.5 and TechnoJock's Turbo Toolbox. The shell was compiled into Improve.exe which was divided into five units - Include, Text, Control, Prequest and Temp. A range of standard Turbo Pascal and TechnoJock's Turbo Toolbox were used to produce the shell and units. The basic structure of the shell consists of standard Turbo Pascal 5.5 units, TechnoJock's units and experimental units. Table 31 provides a list of units used in the Colour experiment.

Table 31 Units of Colour Experiment

| |
|--|
| <p>Standard Turbo Pascal 5.5 Units</p> <p>Improve.exe - 110368 bytes</p> <ol style="list-style-type: none">1. Crt2. Dos3. Graph34. Turbo3 <p>Standard TechnoJock's Turbo Toolbox 5.0 Units</p> <ol style="list-style-type: none">1. KeyTTT52. FastTTT53. WinTTT54. MiscTTT55. MenuTTT56. PullTTT57. ListTTT58. DirTTT59. ReadTTT510. IOTTT511. StrnTTT5 <p>Experiment Units</p> <ol style="list-style-type: none">1. Include (Include.tpu 7232 bytes)2. Text (Text.tpu 57824 bytes)3. Controlt (Controlt.tpu 7232 bytes)4. Prequest (Prequest.tpu 15319 bytes)5. Temp (Temp.tpu 13664 bytes) |
|--|

The technical limitations of Turbo Pascal 5.5 restricted the size of the tutorial. Units could not exceed 64000 bytes and this placed a limit on the contents of each unit. Turbo Pascal 5.5 was the most recent version of the language available during the construction of the shell.

The Graphics experimental shell was constructed using Turbo Pascal 6.0 and TechnoJock's Object Toolbox. The shell was compiled into Shell.exe which was divided into eight units Shelltext, ShellGraphics, ShellGWin, ShellFrames, ShellDictionary, ShellQuestion, ShellMore and ShellExamples. A range of standard Turbo Pascal 6.0 and TechnoJock's Object Toolbox units were used to produce the shell and units. Table 32 provides a list of units in the Graphics experiment.

Table 32 Units in Graphics Experiment

| |
|--|
| <p>Standard Turbo Pascal 6.0 Units</p> <ol style="list-style-type: none">1. Dos2. Crt3. Graph <p>Standard TechnoJock's Object Toolbox Units</p> <ol style="list-style-type: none">1. totFAST2. totMENU3. totMSG4. totIO15. totLOOK6. totStr <p>Experiment Units</p> <p>Shell.exe - 153696 bytes Shell.ovr - 166672 bytes (Overlay)</p> <ol style="list-style-type: none">1. Shelltext (Shelltext.tpu 82704 bytes)2. ShellGraphics (ShellGraphics.tpu 76592 bytes)3. ShellGWin (ShellGWin.tpu 85680 bytes)4. ShellFrames (ShellFrames.tpu 36176 bytes)5. ShellDictionary (ShellDictionary.tpu 7040 bytes)6. ShellQuestion (ShellQuestion.tpu 87840 bytes)7. ShellMore (ShellMore.tpu 9072 bytes)8. ShellExamples (ShellExamples.tpu 10040 bytes) |
|--|

Text Files

The experimental shell used three text files to include large elements of text in the program. These provided a means of including a range of examples in the tutorial, more information to extend the tutorial and a dictionary of terms. This approach provided a simple means for changing the contents of the shell merely by changing the text files. Each text file could be prepared as an ASCII file using a wordprocessor. The three files were:

- (1) Example.doc - 13241 bytes;
- (2) More.doc - 10610 bytes; and
- (3) Dictionary.doc - 19090 bytes.

Graphic Files

There were 126 graphic files each of which held a single graphic image. The size of the Graphics experimental shell grew significantly so that it exceeded the available run-time memory of the computer. In order to reduce the run-time memory, overlays were used. Overlays are parts of a program that share a common memory area. Only the parts of the program that are required for a given function reside in memory at the same time; they can overwrite each other during execution. Overlays allow programs to be executed that are far larger than the available memory. Turbo Pascal generates an overlay file (Shell.ovr) which contains all elements of the program that will be swapped in and out of memory during the execution of the program. The executable file (Shell.exe) contains the elements of the program which are static and not swapped in and out of memory. The overlays were transparent to users.

Design of Learning Material

In order to provide a consistent design framework for the experiments, a number of screen design guidelines were followed.

These were:

- (1) the length of a line of text should be short (between 8 to 10 words);
- (2) each line of text should be broken at a syntactic point;
- (3) the text should be displayed in meaningful units (a single concept at a time);
- (4) text should be presented left justified with a ragged right edge;
- (5) text should use a mixture of upper and lower case letters;
- (6) vertical scrolling in smooth mode should be used as an aid to browsing or skimming, not as the main presentation method;
- (7) the contrast between foreground and background should be maximised to aid readability; and
- (8) presentation of information should be consistent.

These are a combination of published guidelines (Galitz, 1985; 1993; 1994; Spannaus, 1985; Clarke, 1989; 1992; Rivlin et al, 1990; Harrison, 1991; Sherr, 1993). They are all text guidelines and research evidence has shown they are reliable and valid.

In addition, the results of the Colour experiment were used to provide a consistent approach to the use of colour in the Graphics experiment. This also served to further validate the guidelines identified in the Colour experiment on the use of colour within the Graphics experiment.

3.7 Conclusions

Two experiments were developed. The Colour experiment investigated the use of colour and experimental methods while the Graphics experiment considered the use of graphics. Both experimental designs were based on a post-test-only design. The outcomes of the Colour experiment contributed to the design of the Graphics experiment.

Both Colour and Graphics experiments were computer-based, used subjects drawn from the working population, asked users to undertake a learning task and employed a variety of different screen displays. The review of experimental design and survey of computer-based learning materials discussed in chapter 2 provided the basis for the development of both approach and presentation of learning material. Some features however, were intentionally varied to allow their effects to be considered as part of the research. In particular, a higher proportion of graphic images were used than was typically found in the survey of materials.

The Colour experiment consisted of six versions of a tutorial on study skills. Each version was identical except that the number of colours varied from 3 in version 0 to 11 in version 5. In contrast, the Graphics experiment consisted of a single system containing different types and sizes of graphics and other features. Subjects were free to access each part of the system as they interacted with the learning material.

The experimental systems were developed using Turbo Pascal, TechnoJock's Turbo Toolbox, TechnoJock's Object Toolbox and some public domain graphics utilities. In order to improve productivity, a number of procedures were developed to position text, control colour changes, clear areas of the screen, manipulate and create graphic images.

Chapter Four

EXPERIMENTAL RESULTS

4.0 Introduction

The aim of the Colour and Graphics experiments was to investigate the effects of some of the factors which influence the design of computer-based learning materials. A number of factors and hypotheses were examined.

The Colour experiment concentrated on use of colour in computer-based learning and research methods to be employed.

The Graphics experiment was focused on the use of graphics and the complexity of the screen displays in computer-based learning. Section 1.10 lists the hypotheses which were investigated as part of the Colour and Graphics experiments. This chapter will cover the results obtained from both experiments, including experimental subjects and statistical methods. The results from the Colour experiment will concentrate on the subjects' use and opinion of the computer-based learning material. That is:

- (1) time spent on different types of display;
- (2) route taken by subjects through the material;
- (3) subjects' responses to post-test questionnaire; and
- (4) design of the colour experiment.

The results from the Graphics experiment will also focus on the subjects' use and opinion of the computer-based learning material. That is:

- (1) time spent on different types of display;
- (2) complexity of the display;
- (3) subjects' responses to post-test questionnaire;
- (4) route taken by subjects through the material; and
- (5) subjects' mode of operation of the tutorial.

4.1 Characteristics of Subjects

This section considers the characteristics of the subjects of the Colour and Graphics experiments.

Colour Experiment Subjects

Sixty users took part in the Colour experiment. They were drawn from two groups; postgraduate students at the University of Sheffield and employees of the Employment Department in Sheffield.

Tables 33, 34 and 35 provide an analysis of the users in terms of age, gender and academic achievement.

Table 33 Characteristics of Subjects (Colour Experiment)

(N = 60)

| Age range | Male | Currently studying | Female | Currently studying | Total |
|-----------|------|--------------------|--------|--------------------|-------|
| Under 20 | 1 | 1 | 1 | 1 | 2 |
| 20 to 30 | 10 | 9 | 18 | 12 | 28 |
| 30 to 40 | 7 | 3 | 9 | 4 | 16 |
| 40 to 50 | 7 | 2 | 4 | 1 | 11 |
| 50 to 60 | 0 | 0 | 2 | 0 | 2 |
| over 60 | 1 | 0 | 0 | 0 | 1 |
| Total | 26 | 15 | 34 | 18 | 60 |

The tutorial was aimed at developing the learners' study skills. The subjects were chosen as having this particular need or interest and thus would be motivated to learn. Of the experimental subjects, 33 of the 60 were undertaking courses of study during the experiment.

Table 34 Gender and Qualifications of Subjects (Colour Experiment)**(N = 60)**

| Qualifications | Male | Female | Total |
|-----------------------|-------------|---------------|--------------|
| None | 2 | 1 | 3 |
| Below GCSE O-Level | 0 | 3 | 3 |
| GCSE O-level | 3 | 6 | 9 |
| GCSE A-level | 5 | 6 | 11 |
| First degree | 12 | 16 | 28 |
| Postgraduate | 4 | 2 | 6 |
| Total | 26 | 34 | 60 |

Table 35 Highest Qualification (Colour Experiment) (N = 60)

| Age | Above A-Level GCSE or equivalent | A-Level GCSE or equivalent | O-level GCSE or equivalent | None |
|------------|---|---|---|--------------|
| 16-19 | 0% | 1.7% | 0% | 1.7% |
| 20-24 | 11.7% | 3.3% | 5.0% | 5.0% |
| 25+ | 45.0% | 15.0% | 8.3% | 3.3% |
| All | 56.7% | 20.0% | 13.3% | 10.0% |

A comparison between the Colour experiment subjects and the working population of Great Britain as shown in table 36 (Labour Force Survey, 1988) in terms of age and highest qualification reveals a number of issues:

- (1) the experimental sample contained a higher proportion of subjects with qualifications above A-level GCSE or equivalent;
- (2) the experimental sample contained a lower proportion of subjects with O-level GCSE or equivalent or no qualifications; and

- (3) the experimental sample contained a similar proportion of subjects with A-level or equivalent qualifications.

Table 36 Highest Qualification in Great Britain

| Age | Above A-Level GCSE or equivalent | A-Level GCSE or equivalent | O-level GCSE or equivalent | None |
|------------|---|-----------------------------------|-----------------------------------|-------------|
| 16-19 | 0.05% | 14.8% | 54.1% | 28.0% |
| 20-24 | 9.8% | 29.3% | 38.3% | 19.5% |
| 25+ | 15.5% | 21.8% | 16.3% | 37.3% |
| All | 13.3% | 22.1% | 22.9% | 34.0% |

Graphics Experiment Subjects

Seventy-three users took part in the Graphics experiment. They were drawn from several groups of students at the University of Sheffield, employees of the Employment Department in Sheffield, computer professionals and volunteers attending various education and training conferences. None of the subjects involved in this study took part in the Colour experiment. All subjects were regular users of computers and were of working age. Forty males and thirty-three females took part in the Graphics experiment.

The characteristics of the experimental subjects are shown in tables 37, 38, 39, 40 and 41. Table 37 provides an analysis of all subjects by gender, age and qualifications. Tables 38, 39 and 40 provide an analysis of the expert, intermediate and beginners groups, respectively. Table 41 shows the highest qualification held by the experimental subjects by age.

Table 37 Characteristics of Subjects (Graphics Experiment)

(N = 73, M = male and F = female)

| Age | Above A-Level GCSE or equivalent | | A-Level GCSE or equivalent | | O-Level GCSE or equivalent | | None | | Total | |
|-------|----------------------------------|----|----------------------------|---|----------------------------|---|------|----|-------|----|
| | M | F | M | F | M | F | M | F | M | F |
| 16-19 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 2 | 1 |
| 20-24 | 1 | 2 | 1 | 1 | 2 | 0 | 2 | 4 | 6 | 7 |
| 25-34 | 1 | 4 | 5 | 4 | 1 | 2 | 1 | 3 | 8 | 13 |
| 35-49 | 6 | 3 | 4 | 2 | 3 | 0 | 5 | 3 | 18 | 8 |
| 50-64 | 1 | 1 | 1 | 1 | 2 | 0 | 2 | 2 | 6 | 4 |
| Total | 9 | 10 | 11 | 8 | 9 | 3 | 11 | 12 | 40 | 33 |

Table 38 Characteristics of Expert Subjects (Graphics Experiment)

(N = 24, M = male and F = female)

| Age | Above A-Level GCSE or equivalent | | A-Level GCSE or equivalent | | O-Level GCSE or equivalent | | None | | Total | |
|-------|----------------------------------|---|----------------------------|---|----------------------------|---|------|---|-------|---|
| | M | F | M | F | M | F | M | F | M | F |
| 16-19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 20-24 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 3 | 1 |
| 25-34 | 1 | 1 | 3 | 2 | 1 | 0 | 0 | 2 | 5 | 5 |
| 35-49 | 3 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 6 | 3 |
| 50-64 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| Total | 5 | 2 | 4 | 4 | 3 | 0 | 3 | 3 | 15 | 9 |

Table 39 Characteristics of Intermediate Subjects (Graphics Experiment)**(N = 25, M = male and F = female)**

| Age | Above A-Level GCSE or equivalent | | A-Level GCSE or equivalent | | O-Level GCSE or equivalent | | None | | Total | |
|-------|----------------------------------|---|----------------------------|---|----------------------------|---|------|---|-------|----|
| | M | F | M | F | M | F | M | F | M | F |
| 16-19 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| 20-24 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 2 | 2 |
| 25-34 | 0 | 1 | 2 | 1 | 0 | 2 | 1 | 1 | 3 | 5 |
| 35-49 | 1 | 1 | 0 | 0 | 2 | 0 | 2 | 1 | 5 | 2 |
| 50-64 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 3 | 2 |
| Total | 2 | 4 | 3 | 2 | 4 | 2 | 5 | 3 | 14 | 11 |

Table 40 Characteristics of Beginner Subjects (Graphics Experiment)**(N = 24, M= male and F = female)**

| Age | Above A-Level GCSE or equivalent | | A-Level GCSE or equivalent | | O-Level GCSE or equivalent | | None | | Total | |
|-------|----------------------------------|---|----------------------------|---|----------------------------|---|------|---|-------|----|
| | M | F | M | F | M | F | M | F | M | F |
| 16-19 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| 20-24 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 3 | 1 | 4 |
| 25-34 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 3 |
| 35-49 | 2 | 1 | 3 | 1 | 0 | 0 | 2 | 1 | 7 | 3 |
| 50-64 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 2 | 2 | 2 |
| Total | 2 | 4 | 4 | 2 | 2 | 1 | 3 | 6 | 11 | 13 |

Table 41 Highest Qualification (Graphics Experiment)

| Age | Above A-Level GCSE or equivalent | A-Level GCSE or equivalent | O-Level GCSE or equivalent | None |
|------------|---|-----------------------------------|-----------------------------------|-------------|
| 16-19 | 0% | 0% | 2.7% | 1.3% |
| 20-24 | 4.1% | 2.7% | 2.7% | 8.2% |
| 25+ | 21.9% | 23.3% | 10.9% | 21.9% |
| All | 26.0% | 26.0% | 16.3% | 31.4% |

A comparison between the Graphics experimental subjects (table 41) and the working age population of Britain in terms of age and highest qualification (table 36) reveals a number of issues:

- (1) the experimental sample contained a higher proportion of subjects with qualifications above A-Level GCSE or equivalent;
- (2) the experimental sample contained a lower proportion of subjects with O-Level GCSE or equivalent or no qualifications in the 16-19 and 20-24 age groups; and
- (3) the experimental group contained broadly similar proportions of 25+ year olds as does the working population of Britain.

An important issue is whether the results of the research can be transferred to the wider population. The experimental subjects were generally more highly qualified and fewer had no qualifications, compared to the working population as a whole. However, the trend is for the proportion of the working population who hold qualifications to increase. In 1984, 63.6% of working population held qualifications. By 1989 this had risen to 67.7% (Training Agency, 1990).

Size of Sample

In all experiments there is a need to draw a balance between the three interrelated factors of significance, power and cost of the experiment; that is, the degree of significance, the reduction of the probability of failing to detect an effect against the increase in the sample size and the greater the size of the sample, the higher the cost of the experiment in terms of both money and other resources.

A simple analysis would seem to suggest that the sample size should always be as large as possible to maximise the detection of a false hypothesis. However, an excessively large sample may produce an extra sensitive hypothesis test that detects a very small effect lacking any practical importance. A sample which is too small may produce an insensitive test with a large standard error that will fail to detect even a large important effect (Witte, 1989).

To minimise these type of effect, the sample was monitored during the experiments by carrying out the required statistical tests as the individual results became available. This continuous testing allowed control to be maintained. In addition, the effects of different sample sizes on the standard error were considered, once the complete results of the sample were taken. Table 42 provides an overview of the two experimental samples while table 43 shows the relationship between standard error and sample size. It allowed the relationship between increasing the size of the experiment to be compared to the cost.

Table 42 Colour and Graphics Experimental Samples

| Measurement | Text Screens | Graphic Screens |
|--------------------|---------------------|------------------------|
| Sample size | 60 | 73 |
| Mean | 22.36 | 16.33 |
| Standard Deviation | 11.75 | 5.390 |

Table 43 Standard Error and Sample Size

| Sample | Text Standard Error | Graphics Standard Error |
|---------------|----------------------------|--------------------------------|
| 10 | 3.72 | 1.71 |
| 20 | 2.63 | 1.21 |
| 30 | 2.14 | 0.98 |
| 40 | 1.86 | 0.85 |
| 50 | 1.66 | 0.76 |
| 60 | 1.52 | 0.70 |
| 73 | 1.37 | 0.63 |
| 100 | 1.18 | 0.54 |
| 200 | 0.83 | 0.38 |
| 1000 | 0.37 | 0.17 |

4.2 Colour Experiment

In this experiment, different subjects were randomly allocated to the experimental groups. The statistical tests employed were:

- (1) Mann-Whitney Test;
- (2) Kruskal-Wallis one-way analysis of variance; and
- (3) Jonckheere trend test (one-tailed test).

Statistical tests aim to discover the percentage probability that the experimental results are randomly distributed rather than a significant effect due to the independent variable (number of colours used). All three tests required a statistical value to be calculated (observed value) and compared with the value (critical value) obtained from standard statistical tables. The critical values were provided for different percentage levels of significance and the number of experimental subjects. A symbol

is given for each test's statistical value. These are:

- (1) U' for Mann-Whitney;
- (2) H for Kruskal-Wallis; and
- (3) S for Jonckheere.

Tables 44, 45 and 46 provide an overall basic analysis of the results of the Colour experiment. Table 44 shows the mean time spent by the subjects on individual frames and standard deviation for each version of the experiment. Table 45 shows the mean time spent by the subjects and standard deviations for each type of display (menu, question and tutorial page) for each version of the experiment. Table 46 shows the total number of frames accessed and the total mean time (seconds) taken to complete the tutorial for each experimental version.

Table 44 Means and Standard Deviations for Six Colour Experimental Versions

| Version of Colour Experiment | Mean (seconds) | Standard Deviation |
|-------------------------------------|-----------------------|---------------------------|
| 0 | 7.28 | 3.29 |
| 1 | 10.23 | 4.69 |
| 2 | 7.98 | 3.29 |
| 3 | 9.77 | 5.11 |
| 4 | 9.34 | 5.14 |
| 5 | 10.99 | 4.14 |

Table 45 Colour Experiment Means and Standard Deviations for each Type of Display (SD = Standard Deviation)

| Version | Overall Mean | Overall SD | Menu Mean | Menu SD | Question Mean | Question SD | Page Mean | Page SD |
|---------|--------------|------------|-----------|---------|---------------|-------------|-----------|---------|
| 0 | 7.28 | 3.29 | 6.34 | 3.39 | 9.40 | 3.84 | 6.27 | 1.38 |
| 1 | 10.23 | 4.69 | 9.27 | 2.26 | 11.27 | 4.06 | 10.2 | 6.77 |
| 2 | 7.98 | 3.29 | 7.60 | 3.59 | 9.97 | 3.14 | 6.38 | 2.08 |
| 3 | 9.77 | 5.11 | 9.16 | 4.57 | 11.93 | 5.06 | 8.37 | 5.34 |
| 4 | 9.34 | 5.14 | 11.18 | 7.89 | 9.38 | 1.98 | 7.47 | 3.30 |
| 5 | 10.99 | 4.14 | 12.16 | 5.27 | 11.58 | 2.56 | 9.24 | 4.18 |

Table 46 Frames Accessed and Mean Time (Colour Experiment)

| Version | Colours | Mean Time | Mean Frames |
|-----------|---------|-----------|-------------|
| Version 0 | 2 | 642 | 103 |
| Version 1 | 5 | 687 | 77 |
| Version 2 | 6 | 653 | 96 |
| Version 3 | 7 | 748 | 91 |
| Version 4 | 9 | 585 | 74 |
| Version 5 | 11 | 558 | 58 |

Mann-Whitney Test

The Mann-Whitney test analysed the difference between the performance of all the different subjects in two experimental conditions. Thus the analysis was conducted between pairs of conditions, that is:

- (1) version 0 - version 1 (5 colours);
- (2) version 0 - version 2 (6 colours);

- (3) version 0 - version 3 (7 colours);
- (4) version 0 - version 4 (9 colours); and
- (5) version 0 - version 5 (11 colours).

The test compared the monochrome version of the tutorial with the coloured versions of the computer-based learning material. The null hypothesis is that there is no differences between the versions, so there should be roughly equal scores and therefore approximately equal ranks in the pairs of versions. If there are many low or high ranks in one version than another, this is likely not to be due to chance but to the effects of the independent variable (colour). If observed value is lower than critical value, the null hypothesis can be rejected. The results of statistical analysis use a one-tailed test (colour increases time spent on each frame) of mean time spent on each frame of tutorial.

Version 0 to Version 1 (monochrome to 5 colours)

Observed value $U' = 30$

Critical value $U' = 31$ (0.005% level of significance)

Critical value $U' = 47$ (0.05% level of significance)

Observed value is smaller than critical value at 0.005% and 0.05% significance level.

The null hypothesis can be rejected.

Version 0 to Version 2 (monochrome to 6 colours)

Observed value $U' = 70$

Critical value $U' = 31$ (0.005% level of significance)

Critical value $U' = 47$ (0.05% level of significance)

Observed value is larger than critical value at both 0.005% and 0.05% significance

levels. The null hypothesis cannot be rejected.

Version 0 to Version 3 (monochrome to 7 colours)

Observed value $U' = 50$

Critical value $U' = 38$ (0.005% level of significance)

Critical value $U' = 55$ (0.05% level of significance)

Observed value is smaller than critical value at 0.05% significance level. The null hypothesis can be rejected.

Version 0 to Version 4 (monochrome to 9 colours)

Observed value $U' = 25$

Critical value $U' = 10$ (0.005% level of significance)

Critical value $U' = 19$ (0.05% level of significance)

Observed value is larger than critical value at both 0.005% and 0.05% significance levels. The null hypothesis cannot be rejected.

Version 0 to Version 5 (monochrome to 11 colours)

Observed value $U' = 16$

Critical value $U' = 13$ (0.005% level of significance)

Critical value $U' = 20$ (0.05% level of significance)

Observed value is smaller than critical value at 0.05% significance level. The null hypothesis can be rejected.

Table 47 gives a summary of the statistical results of the Mann-Whitney test comparing the different versions of the Colour experiment.

Table 47 Summary of Mann-Whitney Tests for Colour Experiment

(n/s - not significant)

| Comparison | Observed Value | Critical Value 0.005% | Critical Value 0.05% | Level of Significance |
|-------------------|-----------------------|------------------------------|-----------------------------|------------------------------|
| 0 to 1 | 30 | 31 | 47 | 0.005% |
| 0 to 2 | 70 | 31 | 47 | n/s |
| 0 to 3 | 50 | 38 | 55 | 0.05% |
| 0 to 4 | 25 | 10 | 19 | n/s |
| 0 to 5 | 16 | 13 | 20 | 0.05% |

The results show that the increase in use of colour is significant at the 0.05% level of significance for the comparison between:

- (1) version 0 and version 1;
- (2) version 0 and version 3; and
- (3) version 0 and version 5.

The results also show that the increased use of colour is significant at the 0.005% level of significance for the comparison between version 0 and version 1. This indicates that the use of colour significantly affects the time spent on each frame of the tutorial in versions 1, 3 and 5.

Kruskal-Wallis One-Way Analysis of Variance

This test is essentially an extension of the Mann-Whitney test when there is a need to analyse data from more than two conditions. The subjects are again randomly assigned to the different experimental conditions. It is used to test the overall differences between the scores but it does not provide any analysis of a trend in the differences between the conditions. The critical value H , the statistic by which the differences are measured, represents the minimum value that a difference in ranks

must reach in order to have a chance probability of less than the required significance level; in this case the null hypothesis can be rejected. The observed value of H must be equal to or larger than the critical value, that is, the larger the difference between conditions the more likely it is significant. The analysis was conducted of:

- (1) version 0 - version 1 - version 3;
- (2) version 0 - version 3 - version 5; and
- (3) version 1 - version 3 - version 5.

This last comparison was between three versions using colour and excludes the base monochrome version. The other two comparisons were between the monochrome version and the two coloured versions of the tutorial. These combinations were chosen to adequately cover the range of colours used in the experiment. The results of the earlier Mann-Whitney analysis were also used to identify the versions.

Version 0 - Version 1 - Version 3 (monochrome - 5 colours - 7 colours)

Observed value $H = 7.23$ Degrees of freedom 2

Critical value $H = 5.99$ (0.05% level of significance)

Critical value $H = 9.21$ (0.01% level of significance)

The observed value is larger than the critical value at the 0.05% significance level and smaller than the 0.01% critical value. The null hypothesis can be rejected at 0.05% level of significance.

Version 0 - Version 3 - Version 5 (monochrome - 7 colours - 11 colours)

Observed value $H = 7.29$ Degrees of freedom 2

Critical value $H = 5.99$ (0.05% level of significance)

Critical value $H = 9.21$ (0.01% level of significance)

The observed value is larger than the critical value at the 0.05% significance level and smaller than the 0.01% critical value. The null hypothesis can be rejected at 0.05% level of significance.

Version 1 - Version 3 - Version 5 (5 colours - 7 colours - 11 colours)

Observed value $H = 1.33$ Degrees of freedom 2

Critical value $H = 5.99$ (0.05% level of significance)

Critical value $H = 9.21$ (0.01% level of significance)

The results show that the increased use of colour was not significant at any level of significance for the comparison between three different versions of the tutorial, all of which include colour. This comparison did not include the control version 0 which was monochrome. This result suggests that the essential difference is between monochrome and colour, not between the degree of colour used.

Jonckheere Trend Test (One-Tailed Test)

This test is essentially an extension of the Kruskal-Wallis one-way analysis of variance for dealing with three or more experimental conditions. Subjects are randomly assigned to different experimental conditions. The Jonckheere test evaluates the predicted trend across the scores from different experimental conditions. In this experiment, the trend that is predicted is that the mean time spent on each display will be greater as the use of colour increases. The observed value of S must be greater than the critical value of S for the results to be significant. The analysis was conducted of:

- (1) version 0 - version 1 - version 3;
- (2) version 0 - version 3 - version 5;
- (3) version 0 - version 1 - version 3 - version 5; and
- (4) version 0 - version 1 - version 2 - version 3 - version 4 - version 5.

These combinations were chosen to adequately reflect the use of colour in the different experimental versions, that is, the increasing use of colour across the versions of the experiment and the results of the earlier analysis using the Mann-Whitney and Kruskal-Wallis tests.

Version 0 - Version 1 - Version 3 (monochrome - 5 colours - 7 colours)

Observed value $S = 98$

Critical value $S = 88$ (0.05% significance level)

Critical value $S = 124$ (0.01% significance level)

The observed value is larger than the critical value at the 0.05% significance level and smaller than the 0.01% critical value. The null hypothesis can be rejected at the 0.05% level of significance.

Version 0 - Version 3 - Version 5 (monochrome - 7 colours - 11 colours)

Observed value $S = 59$

Critical value $S = 53$ (0.05% significance level)

Critical value $S = 74$ (0.01% significance level)

The observed value is larger than the critical value at the 0.05% significance level and smaller than the 0.01% critical value. The null hypothesis can be rejected at the 0.05% level of significance.

Version 0 - Version 1 - Version 3 - Version 5

(monochrome - 5 colours - 7 colours - 11 colours)

Observed value $S = 90$

Critical value $S = 82$ (0.05% significance level)

Critical value $S = 115$ (0.01% significance level)

The observed value is larger than the critical value at the 0.05% significance level and smaller than the 0.01% critical value. The null hypothesis can be rejected at the 0.05% level of significance.

Version 0 - Version 1 - Version 2 - Version 3 - Version 4 - Version 5
(monochrome - 5 colours - 6 colours - 7 colours - 9 colours - 11 colours)

Observed value $S = 58$

Critical value $S = 121$ (0.05% significance level)

Critical value $S = 170$ (0.01% significance level)

The observed value is smaller than the critical value at the 0.05% and the 0.01% significance level. The null hypothesis cannot be rejected.

The overall results show a significant trend with increased use of colour, the mean time spent on each frame of the tutorial increased. The exception was the analysis of all six versions to identify a significant trend.

Analysis of Individual Types of Display

The analysis of the individual types of display (menu, questions and tutorial frames) produced similar results to the analysis of the whole tutorial.

Mann-Whitney Test

The Mann-Whitney test analyses the difference between the performance of all the different subjects in two experimental conditions. The analysis was therefore conducted between pairs of conditions for the menu, question and tutorial pages, that is, comparing:

- (1) version 0 - version 1;
- (2) version 0 - version 3; and
- (3) version 0 - version 5.

These comparisons were chosen on the basis of the analysis of all frames. All the statistical analysis was based on a one-tail test using the same null hypothesis as that used for the analysis of the whole tutorial.

Version 0 - Version 1 (monochrome - 5 Colours)

Menu pages

Observed value $U' = 34$

Critical value $U' = 35$ (0.01% level of significance)

The observed value is less than the critical value at the 0.01% significance level. The null hypothesis can be rejected at the 0.01% level of significance.

Question pages

Observed value $U' = 44$

Critical value $U' = 38$ (0.05% level of significance)

The observed value is greater than the critical value at the 0.05% significance level. The null hypothesis cannot be rejected at the 0.05% level of significance.

Tutorial pages

Observed value $U' = 33$

Critical value $U' = 35$ (0.01% level of significance)

The observed value is less than the critical value at the 0.01% significance level. The null hypothesis can be rejected at the 0.01% level of significance.

Version 0 - Version 3 (monochrome - 7 colours)

Menu pages

Observed value $U' = 33$

Critical value $U' = 34$ (0.005% level of significance)

The observed value is less than the critical value at the 0.005% significance level. The null hypothesis can be rejected at the 0.005% level of significance.

Question pages

Observed value $U' = 29$

Critical value $U' = 31$ (0.005% level of significance)

The observed value is less than the critical value at the 0.005% significance level.

The null hypothesis can be rejected at the 0.005% level of significance.

Tutorial pages

Observed value $U' = 45.5$

Critical value $U' = 51$ (0.05% level of significance)

The observed value is less than the critical value at the 0.05% significance level. The null hypothesis can be rejected at the 0.05% level of significance.

Version 0 - Version 5 (monochrome - 11 colours)

Menu pages

Observed value $U' = 3$

Critical value $U' = 13$ (0.005% level of significance)

The observed value is less than the critical value at the 0.005% significance level.

The null hypothesis can be rejected at the 0.005% level of significance.

Question pages

Observed value $U' = 9$

Critical value $U' = 12$ (0.005% level of significance)

The observed value is less than the critical value at the 0.005% significance level.

The null hypothesis can be rejected at the 0.005% level of significance.

Tutorial pages

Observed value $U' = 14$

Critical value $U' = 16$ (0.01% level of significance)

The observed value is less than the critical value at the 0.01% significance level. The null hypothesis can be rejected at the 0.01% level of significance.

This analysis shows the effects of colour across a variety of displays serving different purposes. The overall results indicate a significant trend as the use of colour increases so does the mean time spent on each frame of the tutorial. This trend was apparent on all types of display (menu, question and tutorial pages).

Mean Times and Frames Accessed

The mean time to complete the tutorial showed an interesting trend, in that it increased from the base level in the monochrome version up to and including version 3 (7 colours). In the last two versions 4 and 5 mean time fell below that taken for the monochrome version.

The analysis of the total time taken to use the tutorial using the Mann-Whitney test was not found to be significant in any of the comparisons which were based on the results of the earlier analysis, that is comparing:

- (1) version 0 and version 1;
- (2) version 0 and version 3; and
- (3) version 0 and version 5.

A comparison between the total time taken to use the tutorials and the total number of frames accessed shows an interesting contrast. The analysis of the number of frames accessed using the Mann-Whitney test produced the following results:

Version 5 - Version 0 (11 colours - monochrome)

Observed value $U' = 23.5$

Critical value $U' = 24$ (0.05% level of significance)

The null hypothesis can be rejected, that is, subjects accessed significantly fewer frames in version 5 than in version 0. Version 5 used 11 colours while version 0 was monochrome.

Version 5 - Version 3 (11 colours - 7 colours)

Observed value $U' = 18$

Critical value $U' = 21$ (0.05% level of significance)

The null hypothesis can be rejected, that is, subjects accessed significantly fewer frames in version 5 employing 11 colours than version 3 using 7 colours.

Version 3 - Version 0 (7 colours - monochrome)

Observed value $U' = 74$

Critical value $U' = 47$ (0.05% level of significance)

The null hypothesis cannot be rejected, that is, there is no significant difference between the number of frames accessed in version 3 (using 7 colours) compared to the monochrome version 0.

The results suggest that the number of frames accessed by the users is not significantly different in versions of the tutorial using seven or fewer colours. However, with versions using more than seven, the number of frames used is significantly smaller. This is an important outcome in that it indicates the optimum number of colours to be employed.

Test Scores from On-Line Test

The on-line test results (table 48) indicated that users of all six versions of the learning material succeeded in answering the built-in questions. An analysis of the subjects' scores showed no significant difference between the different experimental versions. This could be the result of two different factors. Firstly, all the experimental versions were designed to provide effective learning so that this lack of significantly different results could be confirmation that the material has achieved this

purpose. The lowest mean score was in version 3 (7 colours) but even here, a mean score of 69.2% was achieved. The highest score was 80% in the monochrome version of the tutorial. Secondly, the test questions built into the tutorial were not sensitive enough to reveal the differences between the versions.

Table 48 Colour Experiment Test Scores from On-Line Test

| Version | Score | Percentage |
|----------------|--------------|-------------------|
| Version 0 | 4.00 | 80.00 |
| Version 1 | 3.73 | 74.60 |
| Version 2 | 3.92 | 78.40 |
| Version 3 | 3.46 | 69.20 |
| Version 4 | 3.60 | 72.00 |
| Version 5 | 3.57 | 71.40 |

In order to improve the sensitivity of the internal tutorial questions in the Graphics experiment, the tests were expanded to provide a more detailed analysis of comprehension. The detailed changes were to allow a user access to the self-test from any frame of the tutorial and to embed questions into the tutorial as well as the bank of self-test items. This is the normal way questions are used in commercial computer-based learning materials.

Post-test Questionnaire (Colour Experiment)

Users were each presented with a questionnaire on the screen when they had finished working on the tutorial. They were asked to provide some basic information about themselves and answer ten questions concerning their views of the tutorial version they had used.

Basic Information

Each user was asked to state their age, gender and highest qualification.

Questions

Each user was asked to answer the following questions:

1. How easy was the system to use?
2. How interesting was the system to use?
3. How tiring was the system to use?
4. Too much colour or too little colour?
5. Information displayed was hard to read or easy to read?
6. Information displayed was confusing or clear?
7. Sequence of screens was confusing or clear?
8. Learning to operate the system was difficult or easy?
9. Moving around the system was difficult or easy?
10. Position of messages on the screen was inconsistent or consistent?

The subjects were asked to indicate by their choice of scores from 0 to 6, the strength of their views to the extreme conditions shown by definitions of scores 0 and 6.

Table 49 provides a summary of the attitudes to the different experimental versions, ranging from a monochrome version to one employing 11 colours. The answers were analysed using the Mann-Whitney test and no significant differences between the versions were identified - with the exception of the responses to the colour question. This is not surprising since the versions only differed in the use of colour.

However, even considering the responses to the colour question, there was no significant difference between the versions except for:

Version 0 - Version 1 (monochrome - 5 colours)

Observed value $U' = 3$

Critical value $U' = 11$ (0.01% level of significance)

Critical Value $U' = 17$ (0.05% level of significance)

Since the observed value is less than the critical value at both the 0.01% and the 0.05% levels there is a significant difference between the attitude of users to version 0 and version 1. That is, between the monochrome and the 5 colour version.

Table 49 Results of Attitude Questionnaire (Colour Experiment)

| Questions | 0 | 1 | 2 | 3 | 4 | 5 |
|-----------|-----|-----|-----|-----|-----|-----|
| Use | 4.8 | 5.2 | 4.9 | 5.2 | 4.5 | 4.4 |
| Interest | 3.5 | 2.4 | 3.7 | 2.7 | 3.0 | 3.7 |
| Tiring | 4.2 | 4.1 | 4.3 | 4.6 | 4.0 | 4.3 |
| Colour | 4.4 | 3.6 | 3.3 | 2.9 | 2.8 | 1.9 |
| Reading | 4.5 | 4.9 | 5.1 | 5.0 | 5.2 | 3.4 |
| Confusing | 4.3 | 4.8 | 4.1 | 4.7 | 4.0 | 4.3 |
| Sequence | 2.8 | 3.1 | 3.3 | 3.1 | 3.2 | 3.4 |
| Learn | 4.7 | 5.1 | 5.4 | 5.6 | 4.3 | 4.7 |
| Move | 4.1 | 4.8 | 4.2 | 4.9 | 5.0 | 4.4 |
| Position | 4.1 | 3.8 | 4.0 | 4.6 | 5.2 | 4.4 |

The comparison between version 0 and version 5 just failed the significance test at the 0.05% level.

Version 0 - Version 5 (monochrome - 11 colours)

Observed value $U' = 12$

Critical value $U' = 6$ (0.01% level of significance)

Critical value $U' = 11$ (0.05% level of significance)

Although there is limited evidence of significant difference, there are clearly differences in attitude to the use of colour in the tutorial:

- (1) 2 colours - attitude score 4.4;
- (2) 5 colours - attitude score 3.6;
- (3) 6 colours - attitude score 3.3;
- (4) 7 colours - attitude score 2.9;
- (5) 9 colours - attitude score 2.8; and
- (6) 11 colours - attitude score 1.9.

The attitude score ranged from 6 (too little colour) to 0 (too much colour). The trend is clearly to regard the use of less than 7 colours as too little colour and the application of more than 7 colours as too much. This simple analysis shows a distinct change of attitude between six and nine colours. The subjects who had used versions 0 to 2 indicated a modest request for more colour, while versions 4 and 5 showed a modest request for less colour. The optimum number of colours appears from the results of the attitude questions to be seven.

Routes taken through the Colour Tutorial

The tutorial was designed to offer the subjects the choice of which sections to study and in what order to consider them. The subjects were free to leave at any time and to complete all or part of the material. The tutorial was essentially menu driven, with an initial choice of reading an introduction to the tutorial or opting to begin with one of the main tutorials on notetaking in lectures or from text books. All the options were identically presented. Table 50 shows the percentage of learners who accessed the introduction or another module first. The pattern of the users' choices has been called the route and is a map of their use of the whole tutorial. The route taken through each tutorial was logged by the shell without users being aware of the process while they were using the tutorial.

Table 50 Routes (Colour Experiment)

| Version | Introduction Frames | Others |
|----------------|----------------------------|---------------|
| 0 | 84.6% | 15.4% |
| 1 | 75.0% | 25.0% |
| 2 | 75.0% | 25.0% |
| 3 | 66.6% | 33.4% |
| 4 | 50.0% | 50.0% |
| 5 | 57.1% | 42.9% |

The trend was that as the use of colour increases, so did the users' choice of an option other than the introduction. The tutorial modules were accessed from a main menu which linked to two sub-menus. The main menu options were introduction, notes from books, notes from lectures, self-test and exit. The sub-menus provided links to modules on reading strategy, lecture notes, types of notes, equipment and records, faults to avoid, revision aids and notes from books.

The analysis of the subjects' choice of route through the learning material showed a distinct linear approach. The key influence of where learners chose to go was the order the modules appear on the menus. Table 51 shows the order (1st to 8th) in which users accessed individual modules in each of the experimental versions (0 to 5). The analysis shown in the table indicates a tendency to follow a linear sequence.

Table 51 Choice of Modules of Tutorial (Colour Experiment)

| Options | Order | 0 | 1 | 2 | 3 | 4 | 5 |
|-----------------------|-------|-----|-----|-----|-----|-----|-----|
| Introduction | 1st | 85% | 69% | 75% | 64% | 50% | 50% |
| Notes from Books | 2nd | 85% | 77% | 75% | 64% | 50% | 50% |
| Self-Test | 3rd | 54% | 31% | 58% | 57% | 33% | 33% |
| Reading Strategy | 4th | 77% | 62% | 75% | 57% | 50% | 50% |
| Types of Notes | 5th | 23% | 0% | 8% | 7% | 0% | 16% |
| Equipment and Records | 6th | 31% | 0% | 8% | 0% | 0% | 0% |
| Faults to avoid | 7th | 15% | 8% | 8% | 0% | 0% | 16% |
| Revision Aids | 8th | 15% | 0% | 8% | 0% | 0% | 0% |

Table 52 shows the percentage of subjects who accessed each module of the Colour tutorial in the six experimental versions. This shows that all modules were used by the subjects although there were large differences between the number of subjects who accessed particular modules.

Table 52 Access to Tutorial Modules (Colour Experiment)

| Options | 0 | 1 | 2 | 3 | 4 | 5 |
|-----------------------|------|------|------|-----|------|------|
| Introduction | 92% | 92% | 75% | 86% | 66% | 66% |
| Notes from Books | 92% | 100% | 92% | 93% | 100% | 100% |
| Notes from Lectures | 85% | 100% | 83% | 86% | 100% | 66% |
| Self-Test | 92% | 92% | 100% | 93% | 83% | 100% |
| Reading Strategy | 100% | 85% | 100% | 93% | 83% | 83% |
| Lecture Notes | 54% | 77% | 75% | 50% | 50% | 50% |
| Types of Notes | 62% | 31% | 42% | 57% | 66% | 66% |
| Equipment and Records | 62% | 38% | 42% | 43% | 33% | 17% |
| Faults to avoid | 54% | 46% | 50% | 64% | 50% | 33% |
| Revision Aids | 62% | 31% | 42% | 36% | 66% | 50% |

Analysis of Methods Employed in Colour Experiment

The Colour experiment was also designed to test the methods to be employed during the research. This was used to influence the design of the Graphics experiment. The key factors to be investigated were:

- (1) post-test-only design;
- (2) design of tutorial;
- (3) automatic recording of subjects' responses;
- (4) technical issues (for example, choice of software tools); and
- (5) robustness of software.

The post-test-only model was found to be successful during the Colour experiment and was employed in the Graphics experiment. The post-test was encountered immediately a user exited the learning material. The Colour tutorial did not allow a user accidentally to exit the tutorial and return. Once subjects had made the choice, they were committed to leave. This did leave open the possibility that the experimental subjects had accidentally exited the material and were therefore being asked to complete the post-test before they were ready.

To eliminate this possibility, the Graphics experiment was changed so as to ask the subjects to confirm that they wished to exit before allowing them to do so. 29% of learners entered exit and chose to return to tutorial during the Graphics experiment, providing evidence in support of this design change.

The structure of the Colour tutorial is shown in figure 9 and reveals a menu driven approach. The main menu gave users the choice of moving to the two sub-menus representing the tutorial on notetaking from books and lectures, or to an introduction to whole subject, or to a self-test. A user was free to choose but was always required to return to the main menu in order to make a new choice unless the choice was to exit the tutorial.

The detailed structure of the Colour experiment's learning material is shown in figures 8 and 9. The structure of the Question module is illustrated in figure 23.

The overall design of the tutorial was aimed to reflect the current practice in

the design of computer-based learning material. The structure is reasonably typical but the use of the Colour tutorial shell served to highlight the changes required in the Graphics shell to bring it closer to this goal. The results of the Colour experiment showed that the tutorial shell needed to be amended to enhance the Graphics experiment.

The Colour tutorial restricted the users' freedom by limiting them to return to the main menu or sub-menu before they could move to a new section.

In the Colour experiment, only a limited number of questions were used in the self-test. To improve the sensitivity of the test, the number of questions was increased and combined with a considerable increase in the freedom to choose when to take a test. Access to the self-test was provided from any frame of the tutorial and questions were embedded into the tutorial as well as the bank of self-test items.

The random allocation of users to versions of the learning material worked well but did not reflect the use of computer-based learning materials. In schools and commercial training departments, learning materials contain a range of options within each package, for example, different ranges and numbers of colours, different types of graphics and text-only displays. Learners choose which options to take and how long to spend on each element of the material. Their choice may well be influenced by the displays they have already encountered.

The design of the software shell to allow the users' responses to the tutorial to be automatically recorded as a text file on the disk worked well. The system recorded all the experimental results including:

- (1) time spent on individual frames;
- (2) route taken through the material;
- (3) number of frames accessed; and
- (4) post-test questionnaire results.

With the exception of some files which were corrupted during the experiment, the approach was a success and was adopted for the Graphics experiment.

The only changes were:

- (1) additional testing of the material prior to the experiment to eliminate the problem of corrupted files; and
- (2) improvement of the structure of the Turbo Pascal code to improve software efficiency.

In an experimental design in which the material is presented to subjects on a floppy disk and which automatically records their reactions, robust software is essential. Each copy of the material was tested by the author prior to use and this did result in the vast majority of individual results being acceptable. However, a small number of files were found to be corrupted or to contain software bugs.

A total of 67 disks were issued and 7 contained errors or were found to be corrupted after use. In order to eliminate these problems in the Graphics experiment, a small trial to check the material was conducted using subjects other than the author. This resulted in a number of errors being identified and corrected before the experiment started. During the Graphics experiment, no disks were found to be corrupted.

The design of the Graphics experimental shell was adjusted in a number of ways in response to the analysis of the Colour experiment.

1. User Freedom

Direct access from any individual screen to many of the elements of the learning material (for example, Quiz, Dictionary, More Information and Examples modules).

2. Accidentally Exiting the Tutorial

Before users can exit, they must confirm their decision.

3. Overcoming Technical Limitations of Turbo Pascal

The Graphics experimental shell was coded in Turbo Pascal 6.0.

TechnoJock's Object Toolbox provided additional software tools.

Object oriented programming techniques were employed to improve the structure of program.

4. Self-Test

More questions were provided in self-test to improve sensitivity.

Access to self-test was improved to give learners more choice.

Questions were placed within the tutorial in addition to self-test to reflect current commercial practice.

5. Learners' Choice of Options

Learners could choose which options to take.

Learners could choose how long to spend on each section.

6. Use of Colour

Guidelines derived from the Colour experiment were applied to the Graphics experiment tutorial (seven colours optimum).

4.3 Graphics Experiment

During the Graphics experiment, subjects were asked to work through a tutorial on computer studies and complete a questionnaire once they had finished the tutorial. The questionnaire asked the subjects to provide basic details of themselves (age, gender, highest qualification, how frequently they used computers and their own assessment of their level of understanding of computer studies).

This information was used to place the subjects into three groups:

- (1) Expert - Subjects who rated themselves as having 'a great deal of computer knowledge';

- (2) Intermediate - Subjects who rated themselves as having 'a reasonable amount of computer knowledge'; and
- (3) Beginner - Subjects who rated themselves as having 'some knowledge of computers'.

A fourth group who rated themselves as having 'no knowledge of computers' was not included in the statistical analysis. The subjects were not asked to identify themselves but were sent copies of the tutorial and asked to complete the experiment and return the disk and questionnaire. After eliminating subjects who had not completed the questionnaire, the sample contained seventy-three subjects. These were:

- (1) 24 expert subjects;
- (2) 25 intermediate subjects; and
- (3) 24 beginner subjects.

All the subjects were regular users of computers and rated themselves as competent. The difference between the groups was the degree of self-assessed knowledge of computers. The experimental materials were sent to subjects during the period September 1992 to February 1993.

The range of use can be demonstrated by considering the range of time spent on the tutorial by the subjects. Table 53 illustrates this point by showing the minimum and maximum time spent on the tutorial.

Table 53 Range of Use (Graphics Experiment)

| Characteristic | Minimum Use | Maximum Use |
|---------------------------------|--------------------|--------------------|
| Total Time (seconds) | 196 | 4301 |
| Mean Time per Frame (seconds) | 10.31 | 52.46 |
| Total Number of Frames Accessed | 19 | 82 |

The software system automatically recorded the use of material by a subject. This information was stored on the disk and later printed out for analysis.

Statistical Tests

Two main statistical tests were employed to analyse the experimental results during the Graphics experiment. These were the Mann-Whitney and t tests. The Mann-Whitney test was discussed earlier.

The t test is used for designs having two conditions, when the same subjects are carrying out both conditions. The t test is a parametric test which provides a means of discovering the probability of the results of the experiment occurring randomly by chance. A non-parametric test such as the Wilcoxon is also able to fulfil this function. The difference between the tests is that the parametric takes into account not only the ranking of the results, but also differences in size of the results. The t test is thus a more 'powerful' test in that it is more likely to detect a significant difference in the experimental results (Greene and D'Oliveira, 1982).

The experiment involved measuring the behaviour of the same subjects carrying out a task. The task was to use a tutorial on Computer Studies. The tutorial contained a range of screen designs such as text-only, structured text and graphic frames.

Subjects were free to move through the material in any way they chose. This mirrors the use of computer-based learning in education and training departments. The subjects could choose to leave areas of the tutorial untouched which meant that during the analysis, different areas of the tutorial needed to be matched. To compare the behaviour of a subject using a text frame with an analogical graphic design required that both were used. This was not necessarily the case with all the subjects who undertook the experiment.

Complexity

The common features of all screens were not included in the total measure of complexity. These were the control icon and the title bar. These appeared on all text

screens and respectively produced measures of complexity of 5 and 16. The total complexity of each frame of the material including common features would therefore increase by 21. For text and graphics screens, only the title bar was in common and thus the complexity measure would increase by 16.

Tables 54 and 55 show the individual components of the complexity score for individual frames. Table 54 relates to text-only displays and table 55 relates to frames containing both text and graphics.

Table 54 Complexity Measurements (Text-Only Displays)

| Frame No. | Fields | Horizontal | Vertical | Graphics | Total |
|------------------|---------------|-------------------|-----------------|-----------------|--------------|
| 2 | 21 | 6 | 11 | 0 | 38 |
| 3 | 21 | 3 | 8 | 0 | 32 |
| 4 | 28 | 3 | 9 | 0 | 40 |
| 5 | 14 | 4 | 7 | 0 | 25 |
| 6 | 19 | 3 | 6 | 0 | 28 |
| 7 | 24 | 2 | 7 | 0 | 33 |
| 8 | 32 | 2 | 9 | 0 | 43 |
| 10 | 16 | 2 | 6 | 0 | 24 |
| 25 | 28 | 2 | 9 | 0 | 39 |
| 26 | 21 | 2 | 9 | 0 | 32 |
| 27 | 11 | 3 | 7 | 0 | 21 |
| 28 | 22 | 2 | 9 | 0 | 33 |
| 29 | 22 | 2 | 9 | 0 | 33 |
| 30 | 16 | 2 | 7 | 0 | 25 |
| 31 | 20 | 2 | 8 | 0 | 30 |
| 33 | 17 | 6 | 8 | 0 | 31 |
| 34 | 17 | 3 | 8 | 0 | 28 |
| 35 | 18 | 6 | 9 | 0 | 33 |
| 36 | 22 | 4 | 11 | 0 | 37 |
| 37 | 17 | 4 | 9 | 0 | 30 |
| 38 | 26 | 3 | 9 | 0 | 38 |
| 39 | 19 | 3 | 7 | 0 | 29 |
| 42 | 18 | 4 | 9 | 0 | 31 |
| 50 | 24 | 2 | 9 | 0 | 35 |
| 58 | 16 | 4 | 7 | 0 | 27 |

Table 55 Complexity Measurements (Text and Graphics Displays)

| Frame No. | Fields | Horizontal | Vertical | Graphics | Total |
|------------------|---------------|-------------------|-----------------|-----------------|--------------|
| 100 | 13 | 10 | 13 | 4 | 40 |
| 101 | 11 | 15 | 25 | 6 | 57 |
| 102 | 24 | 3 | 13 | 3 | 43 |
| 103 | 25 | 30 | 26 | 12 | 93 |
| 104 | 17 | 4 | 9 | 2 | 32 |
| 105 | 19 | 3 | 12 | 3 | 37 |
| 106 | 17 | 4 | 10 | 1 | 32 |
| 107 | 17 | 4 | 11 | 1 | 33 |
| 108 | 18 | 4 | 11 | 1 | 34 |
| 109 | 15 | 4 | 11 | 1 | 31 |
| 110 | 15 | 4 | 11 | 1 | 31 |
| 120 | 16 | 7 | 11 | 4 | 38 |
| 121 | 29 | 3 | 12 | 1 | 45 |
| 122 | 7 | 9 | 7 | 4 | 27 |
| 123 | 10 | 3 | 7 | 4 | 24 |
| 124 | 16 | 2 | 8 | 1 | 27 |
| 125 | 14 | 2 | 7 | 1 | 24 |
| 126 | 16 | 3 | 8 | 1 | 28 |
| 127 | 18 | 6 | 8 | 4 | 36 |
| 128 | 11 | 3 | 7 | 1 | 22 |
| 129 | 22 | 6 | 10 | 6 | 44 |
| 131 | 15 | 4 | 9 | 2 | 30 |
| 132 | 5 | 4 | 10 | 8 | 27 |
| 200 | 8 | 7 | 7 | 10 | 32 |
| 201 | 8 | 7 | 7 | 10 | 32 |
| 202 | 8 | 8 | 6 | 5 | 27 |

| | | | | | |
|-----|----|---|----|---|----|
| 203 | 16 | 7 | 9 | 4 | 36 |
| 207 | 5 | 6 | 7 | 5 | 23 |
| 208 | 15 | 7 | 12 | 8 | 42 |

Mean Complexity by Display

Table 56 shows the mean complexity scores of the different types of display. These were calculated using all screens, and not simply limited to those which were most frequently used.

Table 56 Mean Complexity by Screen Type

| Types of Screen | Mean Complexity |
|---|-----------------|
| Text | 39.35 |
| Structured Text | 45.50 |
| Graphics | 35.51 |
| Representative Graphic | 33.12 |
| Full-Screen Representative Graphic | 33.31 |
| Part-Screen Representative Graphic | 33.05 |
| One Part-Screen Representative Graphic | 36.05 |
| Two Part-Screen Representative Graphics | 30.11 |
| Three Part-Screen Representative Graphics | 35.75 |
| Four Part-Screen Representative Graphics | 41.00 |
| Logical Graphic | 56.00 |
| Analogue Graphic | 36.55 |
| Full-Screen Analogue Graphic | 36.00 |
| Part-Screen Analogue Graphic | 31.30 |

In order to investigate the relationship between the time spent by the experimental subjects on each type of display and the measure of complexity for that display, a

correlation analysis was undertaken. In each case, the display frames chosen to be included in the analysis were the ones accessed most frequently by the subjects and therefore likely to produce a more representative value of the mean time of access.

The Spearman rank correlation coefficient is designed to measure the degree of relationship between two variables, that is, the time spent by subjects and the complexity score of a display. It uses data which can be put into a rank order. The amount of correlation is given by statistic, r_s which ranges from -1 to +1. A perfect +1 score would indicate that there is a positive one to one correlation between the two variables. A perfect -1 score would indicate there is a negative one to one correlation between the variables. A score of 0 would indicate a purely random relationship between the variables. The test also provides each measure of correlation with an appropriate level of significance.

Text Screens (Coefficient of Correlation)

The Spearman rank correlation coefficient was used to measure the degree of relationship between actual time spent on a text frame and the complexity measure.

N (Number of Screens) = 25.00

Observed Value of Coefficient $r_s = 0.432$

Critical Value of Coefficient $r_s = 0.409$ (0.025% level of significance one-tail test)

The observed value of the coefficient was greater than the critical value. There was a positive correlation between the time spent on each text screen and the measure of complexity. Regression analysis of the text-only data provided a regression line equation of:

$$y = 10.22 + 0.49x \quad \begin{array}{l} y = \text{time} \\ x = \text{complexity} \end{array}$$

Text and Graphics screens

The Spearman rank correlation coefficient was used to measure the degree of relationship between actual time spent on a graphics and text frame and the complexity measure.

N (Number of Screens) = 29.00

Observed Value of Coefficient $r_s = 0.49$

Critical Value of Coefficient $r_s = 0.448$ (0.01% level of significance one-tail test)

The observed value of the coefficient was greater than the critical value therefore there is a positive correlation between the time spent on each graphics and text screen and the measure of complexity.

Regression analysis of the combined graphics and text frames data provided a regression line equation of:

$$y = 9.25 + 0.25x \quad \begin{array}{l} y = \text{time} \\ x = \text{complexity} \end{array}$$

Text and Graphics and Text-only screens

In order to ensure that the data was compatible, the complexity scores for the text-only screens was increased by five. This was owing to the common features on all text screens being eliminated in the text-only correlation measures. However, text and graphics screens do not share the same identical features with text-only screens and so the complexity of the non-common elements had to be added to the text-only screens. These are screens containing both graphic images and text and text-only screens.

N (Number of Screens) = 30.00

Observed Value of the Coefficient $r_S = 0.38$

Critical Value of Coefficient $r_S = 0.364$ (0.025% level of significance one-tail test)

The observed value of the coefficient was greater than the critical value. There was a positive correlation between the time spent on each graphics and text and text-only screens and the measure of complexity. Combining the data for both text-only and text and graphics frames gives a regression line of:

$$y = 19.19 + 0.18x \quad \begin{array}{l} y = \text{time} \\ x = \text{complexity} \end{array}$$

Comparison of Screen Types

Tables 57 shows a comparison between the different types of display against their complexity score, mean time spent on that type of display, percentage complexity compared to mean complexity of a text display and percentage of mean time compared to mean time of a text screen. Table 58 continues the comparison begun in table 57 by showing the actual mean times spent on each type of display against the mean time which the complexity score would predict, based on the norm of the text screen mean complexity. Table 59 compares the types of display against the predicted mean time, based on the text complexity score for each group of learners.

Table 57 Complexity of Screens compared to Text Screens

| Screen Type | Complexity | Mean Time (seconds) | Percentage Complexity | Percentage Time |
|---|-------------------|--------------------------------|----------------------------------|----------------------------|
| Text-only | 39.35 | 22.36 | 100% | 100% |
| Structured Text | 45.50 | 29.00 | 116% | 130% |
| Graphics | 35.51 | 16.33 | 90% | 73% |
| Representative Graphics | 33.12 | 16.56 | 84% | 75% |
| Full-Screen Representative Graphics | 33.31 | 16.65 | 85% | 75% |
| Part-Screen Representative Graphics | 33.05 | 17.44 | 84% | 78% |
| One Part-Screen Representative Graphics | 36.05 | 17.44 | 92% | 78% |
| Two Part-Screen Representative Graphics | 30.11 | 20.75 | 77% | 93% |
| Three Part-Screen Representative Graphics | 35.75 | 21.95 | 91% | 98% |
| Four Part-Screen Representative Graphics | 41.00 | 19.37 | 104% | 87% |
| Logical Graphics | 56.00 | 19.28 | 142% | 86% |
| Analogue Graphics | 36.55 | 16.68 | 93% | 75% |
| Full-Screen Analogue Graphics | 36.00 | 14.13 | 92% | 63% |
| Part-Screen Analogue graphics | 31.30 | 17.04 | 80% | 76% |

Table 58 Predictions of Mean Time

| Screen Type | Complexity | Mean Time Actual (seconds) | Predicted Mean Time (seconds) |
|---|-------------------|-----------------------------------|--------------------------------------|
| Text-only | 39.35 | 22.36 | 22.36 |
| Structured Text | 45.50 | 29.00 | 25.57 |
| Graphics | 35.51 | 16.33 | 20.18 |
| Representative Graphics | 33.12 | 16.56 | 18.82 |
| Full-Screen Representative Graphics | 33.31 | 16.65 | 18.93 |
| Part-Screen Representative Graphics | 33.05 | 17.44 | 18.78 |
| One Part-Screen Representative Graphics | 36.05 | 17.44 | 20.49 |
| Two Part-Screen Representative Graphics | 30.11 | 20.75 | 17.11 |
| Three Part-Screen Representative Graphics | 35.75 | 21.95 | 20.31 |
| Four Part-Screen Representative Graphics | 41.00 | 19.37 | 23.3 |
| Logical Graphics | 56.00 | 19.28 | 31.82 |
| Analogue Graphics | 36.55 | 16.68 | 20.77 |
| Full-Screen Analogue Graphics | 36.00 | 14.13 | 20.46 |
| Part-Screen Analogue Graphics | 31.30 | 17.04 | 17.79 |

Table 59 A Comparison of the Actual Mean Time by Learner Group compared to Complexity Prediction

| Screen Type | Mean Time Beginner | Mean Time Intermediate | Mean Time Expert | Predicted Mean Time |
|---|---------------------------|-------------------------------|-------------------------|----------------------------|
| Text-only | 20.76 | 26.17 | 19.32 | 19.32 |
| Structured Text | 31.76 | 33.56 | 21.34 | 25.57 |
| Graphics | 15.66 | 17.77 | 15.24 | 20.18 |
| Representative Graphics | 15.85 | 18.59 | 15.08 | 18.82 |
| Full-Screen Representative Graphics | 19.27 | 17.92 | 12.82 | 18.93 |
| Part-Screen Representative Graphics | 15.99 | 19.22 | 16.36 | 18.78 |
| One Part-Screen Representative Graphics | 13.28 | 18.07 | 14.79 | 20.49 |
| Two Part-Screen Representative Graphics | 17.65 | 25.29 | 17.05 | 17.11 |
| Three Part-Screen Representative Graphics | 27.00 | 23.31 | 16.41 | 20.31 |
| Four Part-Screen Representative Graphics | 18.83 | 21.66 | 17.05 | 23.30 |
| Logical Graphics | 15.26 | 22.97 | 14.61 | 31.82 |
| Analogue Graphics | 16.31 | 16.84 | 16.68 | 20.77 |
| Full-Screen Analogue Graphics | 13.65 | 16.62 | 11.76 | 20.46 |
| Part-Screen Analogue Graphics | 16.83 | 17.06 | 17.00 | 17.79 |

By applying the Mann-Whitney test to the complexity predictions and the actual times of the three groups of subjects (beginner, intermediate and expert), it was possible to determine if differences between them shown in table 59 are significant.

Complexity Predicted - Overall Mean Time

There was no significant difference between the predicted mean time and the overall mean time subjects spent on accessing frames, that is, the mean of all three groups of users is not significantly different from the mean time predicted by the measure of complexity. This is not surprising considering the positive correlation between the time spent on each graphics and text and text-only screens and the measure of complexity. It serves to confirm the use of complexity measures to predict the time a learner will spend on a given display.

Complexity Predicted - Beginner Group Mean Time

Observed value $U' = 34$

Critical value $U' = 39$ (two-tail test at the 0.02% level of significance)

The observed value was smaller than the critical value at the 0.02% level of significance. There was a significant difference between the actual mean time of the beginner group on each screen and the time predicted by the complexity measure. Beginners spent generally less time on a display than the complexity measure would predict. Although there are some exceptions with beginners spending more time on frames containing structured text, full-screen representative graphics, two part-screen representative graphics and three part-screen representative graphics.

Complexity Predicted - Intermediate Group Mean Time

There was no significant difference between the predicted mean time and the mean time intermediate subjects spent accessing frames.

Complexity Predicted - Expert Group Mean Time

Observed value $U' = 7$

Critical value $U' = 34$ (two-tail test at the 1% level of significance)

The observed value was smaller than the critical value at the one percent level of significance. There was a significant difference between the actual mean time of the expert group on each screen and the time predicted by the complexity measure. The experts spent less time on each type of screen than the complexity measure would predict.

Beginner Group Mean Time - Expert Group Mean Time

There was no significant difference between the beginner group's mean time and the expert group's mean time spent, although on most types of display the beginners spent more time than the experts. This is reasonable given the experts' greater knowledge and understanding of the tutorial contents.

Beginner Group Mean Time - Intermediate Group Mean Time

Observed value $U' = 53$

Critical value $U' = 55$ (two-tail test at the 5% level of significance)

The observed value was less than the critical value at the five percent level of significance. There was a significant difference between the two groups in the mean time spent on each screen. The difference was mainly in that the beginner group of subjects spent less time than the intermediate group on each screen, with two exceptions - full-screen representative graphics and three part-screen representative graphics, where beginners spent more time than the intermediate group.

Expert Group Mean Time - Intermediate Group Mean Time

Observed value $U' = 23$

Critical value $U' = 42$ (two-tail test at the 1% level of significance)

The observed value was less than the critical value at the one percent level of significance. There was a significant difference between the two groups in that the expert group spent less time on each screen than the intermediate group.

The evidence of this analysis supports the view that the prior knowledge of learners will affect their behaviour with particular screen designs. In this case, the main difference between the groups is the knowledge of the subject being studied. Learners will spend less time studying the tutorial display as their knowledge of the subject increases.

Questionnaire (Graphics Experiment)

The subjects were asked to complete a questionnaire after they had completed the tutorial, broadly similar to the questionnaire used in the Colour experiment. The main difference was that instead of employing a scale of 0 to 6 (Colour experiment), subjects were asked to respond using a 1 to 5 scale representing their judgement of how much they agreed with the extreme answers to the questions. The scale was narrowed to obtain a more specific response from the subjects. The subjects were asked for some basic information about themselves. This consisted of their age, gender, highest qualification, use of computers and computing expertise. The specific questions related to the tutorial were:

- (1) How easy was the system to use?
- (2) How interesting was the system to use?
- (3) How tiring was the system to use?
- (4) How much colour was used in the system?

- (5) The information displayed was difficult or easy to understand?
- (6) There were too many or too few illustrations?
- (7) The information displayed was confusing or clear?
- (8) The sequence of screens was confusing or clear?
- (9) Learning to operate the system was difficult or easy?
- (10) Moving around the system was difficult or easy?
- (11) The position of messages on the screen was inconsistent or consistent?
- (12) What size illustrations do you prefer?
- (13) What type of illustrations do you prefer?

Table 60 provides a summary of the results of the questionnaire. The mean scores for each question are given for the three groups of subjects (Beginner, Intermediate and Expert). Tables 61 and 62 show the subjects preferences for the size and type of graphic illustration.

Table 60 Results of Attitude Questionnaire (Graphics Experiment)

| Question | Expert | Intermediate | Beginner | All |
|-----------------|--------|--------------|----------|------|
| 1. Ease | 3.6 | 3.46 | 3.18 | 3.62 |
| 2. Interest | 3.1 | 3.36 | 3.09 | 3.18 |
| 3. Tiring | 3.0 | 3.0 | 3.09 | 3.09 |
| 4. Colour | 2.8 | 3.09 | 3.18 | 3.03 |
| 5. Understand | 3.7 | 3.36 | 3.27 | 3.44 |
| 6. Illustration | 3.2 | 3.18 | 3.36 | 3.25 |
| 7. Confusing | 3.2 | 3.46 | 3.46 | 3.37 |
| 8. Sequence | 2.6 | 3.72 | 3.18 | 3.17 |
| 9. Learning | 3.0 | 3.72 | 3.81 | 3.51 |
| 10. Moving | 3.0 | 3.36 | 3.63 | 3.33 |
| 11. Position | 3.3 | 3.56 | 3.81 | 3.56 |

Table 61 Size of Illustrations - Preferences

| Size | Expert | Intermediate | Beginner | All |
|-----------------------|---------------|---------------------|-----------------|------------|
| Whole-Screen | 5.0 | 3.0 | 4.0 | 3.0 |
| Half-Screen | 2.0 | 2.0 | 2.0 | 2.0 |
| Quarter-Screen | 1.0 | 1.0 | 2.0 | 1.0 |
| Stamp-Size | 6.0 | 4.0 | 6.0 | 6.0 |
| Multiple Small Images | 3.5 | 5.0 | 5.0 | 5.0 |
| Do not care | 3.5 | 6.0 | 2.0 | 4.0 |

Table 62 Types of Graphics - Preferences

| Type | Expert | Intermediate | Beginner | All |
|-------------------|---------------|---------------------|-----------------|------------|
| Realistic | 2.0 | 1.0 | 1.0 | 1.0 |
| Cartoon Type | 1.0 | 3.0 | 4.0 | 2.0 |
| Line Drawings | 4.0 | 2.0 | 2.0 | 3.0 |
| Charts and Tables | 5.0 | 5.0 | 5.0 | 5.0 |
| Diagrams | 3.0 | 4.0 | 3.0 | 4.0 |
| Analogies | 6.0 | 6.0 | 6.0 | 6.0 |

Overall

The initial analysis considered whether there was a significant difference between the three groups in their overall answers to the questions. The responses of the three groups of subjects to the questions was compared using the Mann-Whitney test. The null hypothesis was that there was no difference between the groups of learners responses.

Expert - Intermediate Groups

Observed value $U' = 28$

Critical value $U' = 30$ (two-tailed at the 5% level of significance)

The observed value is smaller than the critical value at the five percent significance level. The null hypothesis can be rejected. There is a significant difference between the expert and intermediate groups of learners.

Expert - Beginner Group

Observed value $U' = 30$

Critical value $U' = 34$ (two-tailed at the 10% level of significance)

The observed value is smaller than the critical value at the ten percent significance level. The null hypothesis can be rejected. There is a significant difference between the expert and beginner groups of learners.

Intermediate - Beginner Group

Observed value $U' = 60$

Critical value $U' = 34$ (two-tailed at the 10% level of significance)

The observed value was greater than the critical value at the 10 percent significance level. The null hypothesis can be accepted. There is no significant difference between the two groups of learners.

The results show that there was a significant difference between the responses of the beginner and expert groups to the post-tutorial questionnaire.

Attitudes to the Use of Colour compared to the Results of Colour Experiment

The use of colour in the Graphics tutorial was based on the outcomes of the Colour experiment which centred on considering the effects of colour on subjects. The key result was that there was a critical change in the subjects' behaviour once more than

seven colours were employed. In the Graphics experiment, this guideline was used throughout the tutorial and table 63 provides some examples of the use of colour in the Graphics tutorial.

Table 63 Examples of Types of Screen

| Example Screens | Number of Colours |
|-------------------------------------|--------------------------|
| Full-screen graphic frames | 4 |
| Text-only frames | 7 |
| Text and part-screen graphic frames | 6 or 7 |
| Test frames | 7 |

This use of the Colour experimental guidelines provided an opportunity to confirm the outcomes of the experiment, at least in terms of the subjects' attitudes to the use of colour.

The attitude questionnaire in the Graphics experiment contained an identical colour question to that used in the Colour experiment. The results from the Graphics experiment questionnaire used a scale of 1 (far too much colour) to 5 (far too little colour). The results were:

- (1) expert users 2.8;
- (2) intermediate users 3.09;
- (3) beginner users 3.18; and
- (4) all users 3.02.

These results show that the users of the Graphics experiment were content (score of 3) with the use of colour, with only a modest difference between the user groups, that is, a difference of 0.38 between the expert and beginner groups with the experts indicating a tendency towards too many colours and beginners towards too few

colours. However, the difference is small. This is important evidence supporting the outcomes of the Colour experiment and adding the additional evidence from the analysis of all types of subjects reacting to the use of colour in a similar way.

Individual Questions (Graphics Experiment)

A review of the learners' answers to the post experimental questionnaire revealed an interesting pattern of answers and differences between the three groups. All three groups showed a positive response to the design of the tutorial (table 60). The intermediate and beginner groups answered all eleven questions positively, except to indicate a marginal request for increased use of colour and illustrations. The expert group's responses were again generally positive with marginal requests for more colour and illustrations. In addition they found the sequence of screens marginally confusing.

Using these responses as a measure of approval, it was clear that the design of the tutorial was valid for these three groups (table 60). This provides evidence to support the underlying screen design guidelines used.

The Graphics experiment also adopted the outcomes of the Colour experiment with regard to colour, that is, the optimum number of colours is seven. The positive reaction of learners to the use of colour in the Graphics tutorial indicates support for these original conclusions.

Expert - Beginner Groups

The comparison of the answers given by the two groups showed a significant difference, although both groups were generally positive to the design of the materials. The beginners' responses compared to the experts' show that:

- (1) they found the system easier to use;
- (2) they found the system more relaxing to use;
- (3) they would have preferred the system to use slightly more colour;
- (4) they found the information displayed was more difficult to understand;

- (5) they would have preferred the tutorials to have used slightly more illustrations;
- (6) they found the information displayed clearer;
- (7) they found the sequence of screens clearer;
- (8) they found the system easier to learn;
- (9) they found moving round the system easier; and
- (10) they found the position of messages on the screen was more consistent.

The general trend of this analysis was that the beginners had a more positive attitude to the tutorial than the experts with the exception of colour, illustrations and the display of information. Beginners found the information displayed more difficult to understand than the experts. This is not surprising since an expert is likely to be more familiar with the information on the subject than a beginner. What is less easy to explain is the beginners' desire for more colour and illustrations than the experts. The difference is not large but could be interpreted as the beginners' desire for easier explanations of the tutorial information which they equate to the increased use of colour and illustrations. The comparison with the beginners' answers showed that experts were more critical of the tutorial perhaps due to their greater knowledge of the subject except for colour, illustrations and understanding the information displayed.

Expert - Intermediate Groups

The comparison of the answers to the post-tutorial questionnaire given by the two groups showed a significant difference, although both groups were generally positive to the design of the material. The intermediates' responses showed that:

- (1) they found the system harder to use;
- (2) they found the system more interesting to use;
- (3) they found the system more relaxing to use;
- (4) they would have preferred the system to use more colour;
- (5) they found the information displayed was more difficult to understand;
- (6) they found the information displayed was clearer;

- (7) they found the sequence of screens clearer;
- (8) they found the system easier to learn;
- (9) they found moving round the system easier; and
- (10) they found the position of messages on the screen was more consistent.

This analysis is more complex than that of the expert to beginner comparison. The intermediates tended to be more positive to the design of the tutorial than the experts with a number of exceptions. The intermediate group was less positive about using the system, colour and understanding the information displayed. A similar path of reasoning to the expert-beginner comparison would indicate that the experts' greater knowledge explains the difference of attitude to the understanding of information shown. This does not explain why the intermediate group found the system more difficult to use except that the difference was small and both groups did not find using the system was very difficult. The difference in knowledge between the two groups should be less than the expert-beginner comparison. Thus, the result that the intermediate group's attitude to illustrations being broadly the same as the experts is consistent with the analysis that the different responses to the questionnaire are due to different levels of knowledge.

Beginner - Intermediate Groups

The comparison of the answers to the post-tutorial questionnaire given by the two groups showed no significant difference. The beginners' responses showed that:

- (1) they found the system easier to use;
- (2) they found the system less interesting to use;
- (3) they would prefer more illustrations to have been used;
- (4) they found the sequence of screens less clear;
- (5) they found moving round the system easier; and
- (6) they found the position of messages on the screen was more consistent.

It is certainly not clear whether the answers from the two groups reflect the degree of difference in their knowledge of the tutorial subject. Five of the questions did not

produce clear differences and the remaining six showed that beginners were more positive to the design in three answers and more negative in three. One explanation is that the difference in knowledge between two groups was not great. The two groups identified their own level of knowledge and may have not always accurately identified their status.

Illustrations - Size and Type

A review of the learners' answers shown in tables 61 and 62 to the post-experimental questions concerned with the size and type of illustrations showed that there was no significant difference between the groups in response to the two questions:

- (1) what size illustrations do you prefer? and
- (2) what type of illustrations do you prefer?

In fact the three groups responded to the questions in a similar manner.

The three groups of learners clearly (table 61) showed a preference for illustrations which occupy a quarter or half of the screen. The beginners, however, gave equal value to 'Do not care' which suggests that the size of illustration was not important to this group. The expert group ranked 'Do not care' as its third equal choice so again, size preferences were not necessarily strong. Only the intermediate group marked 'Do not care' as its sixth choice.

A review of current computer-based learning material shows that the typical use of graphics was in the displaying of a single illustration occupying less than half of the screen. Thus, learners' preference is in line with current practice.

The three groups of learners' answers to the post-tutorial questionnaire clearly reveal that they did not like analogical illustrations. This may, however, be due to learners' failing to understand what is an analogical image. This was partially countered by the learners completing the questions after working through the tutorial containing examples of this type of illustration. In addition to helping them understand what an analogical image is, the question also gave extra guidance in the form of an example of an analogy.

This lack of interest in analogies was reinforced by the preference for realistic images expressed by all three groups. The overall preferences were:

- First Choice: Realistic;
- Second Choice: Cartoon Type;
- Third Choice: Line Drawings;
- Fourth Choice: Diagrams;
- Fifth Choice: Charts and Tables; and
- Sixth Choice: Analogies.

The current practice, as shown by the review of computer-based learning material, is in line with these preferences, that is, representative (realistic) graphics are the most used type of illustration.

Learners' Choice of Route through the Tutorial (Graphics Experiment)

The tutorial modules provided seven tutorials and five additional modules. The tutorial modules contained learning material which was central for the learners to study in order to understand the subject. In addition to the tutorial modules, learners were provided with extra features called 'additional modules' to assist their studies.

Table 64 lists both the tutorial and additional modules.

Table 64 Tutorial and Additional Modules

| Tutorial Modules | Additional Modules |
|-------------------------|---------------------------|
| 1. Computer System | 1. Dictionary |
| 2. Input | 2. Quiz |
| 3. Computer Development | 3. Examples |
| 4. Data Representation | 4. More Information |
| 5. Storage | 5. Exit |
| 6. Processing | Not Applicable |
| 7. Output | Not applicable |

It shows the mean number of modules accessed, the range of modules accessed for each group and the percentage of subjects who accessed all modules. All three groups of learners contained few individuals who accessed all the tutorial modules. The intermediate group contained the highest number, 24% of learners, who accessed all the modules. Table 65 provides an overview of the tutorial modules which each of the three groups of learners accessed.

Table 65 Graphics Experiment Modules Accessed

| Group | Mean (Number of modules accessed) | Range of modules | Accessed all modules |
|--------------|--|-----------------------------|---------------------------------|
| Beginner | 4.0 | 1 to 7 | 8% |
| Intermediate | 4.1 | 1 to 7 | 24% |
| Expert | 4.0 | 2 to 6 | 0% |
| Overall | 4.0 | 1 to 7 | 12% |

Table 66 shows the order in which the tutorial modules were used in relation to the order in which they appeared in the main menu. The percentage of each group of learners (beginner, intermediate and expert) who accessed the tutorial module in that particular order is shown. Table 67 shows the total percentage of each group who accessed the tutorial modules in any order. Thus module one was accessed by 90% of the subjects as their first choice and by 95% of the subjects during their whole use of the Graphics experiment.

Table 66 Choice of Modules of Tutorial (Graphics Experiment)

| Module | Order | Beginner | Intermediate | Expert | Overall |
|---------------|--------------|-----------------|---------------------|---------------|----------------|
| One | 1st | 71% | 100% | 100% | 90% |
| Two | 2nd | 50% | 60% | 42% | 51% |
| Three | 3rd | 58% | 52% | 21% | 40% |
| Four | 4th | 50% | 46% | 8% | 30% |
| Five | 5th | 21% | 25% | 8% | 21% |
| Six | 6th | 8% | 25% | 0% | 12% |
| Seven | 7th | 8% | 25% | 0% | 12% |

Table 67 Tutorial Module Access (Graphics Experiment)

| Module | Beginner | Intermediate | Expert | Overall |
|---------------|-----------------|---------------------|---------------|----------------|
| One | 79% | 100% | 100% | 95% |
| Two | 71% | 60% | 42% | 56% |
| Three | 79% | 68% | 71% | 73% |
| Four | 79% | 60% | 58% | 66% |
| Five | 50% | 44% | 50% | 47% |
| Six | 29% | 44% | 42% | 38% |
| Seven | 8% | 44% | 29% | 27% |

The tutorial modules, main and pop-up menus contained a range of references to the five additional modules. These were intended to encourage the subjects to access these additional modules when they needed them. The tutorial contained 13 cross-references to the Examples additional module spread across all the tutorial modules. Tables 68, 69, 70, 71, 72 and 73 show different aspects of the subjects' use of the additional modules.

Examples Module

Table 68 shows the effect of the number of cross-references encountered in the tutorial in relation to accessing the Examples additional module for each group of subjects. It reveals the pattern of use of the Examples module. It is limited to the first three encounters with a cross-reference since subsequent encounters did not significantly influence subjects to access the Examples module.

Table 68 Cross-References (Examples Module)

| References | None | First | Second | Three or more | No access |
|-------------------|-------------|--------------|---------------|----------------------|------------------|
| Beginner | 0% | 10% | 10% | 10% | 70% |
| Intermediate | 0% | 17% | 8% | 33% | 42% |
| Expert | 10% | 20% | 10% | 20% | 40% |
| Overall | 3% | 16% | 9% | 22% | 50% |

A large proportion of all three groups of subjects did not access the Examples module. This is an interesting result, considering the number of encounters to cross-references to Examples in the tutorial modules these subjects came across. The mean number of cross-references encountered were:

- (1) beginners encountered 5.66 references;
- (2) intermediates encountered 4.25 references; and
- (3) experts encountered 6.75 references.

Dictionary Module

A dictionary of terms was available from both the main and pop-up menus. It was an option which could be picked from every frame of the tutorial. There were, however, no references to the dictionary in the tutorial text. The only prompts that the subjects

had were the choices provided on the menu displays.

Tables 69 and 70 provide a comparison of the use of the Dictionary module by all three groups of subjects. Table 69 shows the percentage of each group of subjects who did not access the Dictionary module, those who used it once and more than once. Table 70 shows a comparison of which route (main or pop-up menu) was taken to access the Dictionary module for each of the three groups of subjects.

Table 69 Access to the Dictionary

| Subjects | No assess | At least once | More than once |
|-----------------|------------------|----------------------|-----------------------|
| Beginner | 50% | 50% | 10% |
| Intermediate | 38% | 62% | 8% |
| Expert | 40% | 60% | 10% |
| Overall | 42% | 58% | 9% |

Table 70 Access to the Dictionary from Main and Pop-Up Menus

| Subjects | Main Menu | Pop-up Menu |
|-----------------|------------------|--------------------|
| Beginner | 50% | 0% |
| Intermediate | 38% | 24% |
| Expert | 50% | 10% |
| Overall | 46% | 12% |

More Information Module

The More Information additional module was available from each frame of the tutorial as an option on the pop-up menu. It was provided to assist subjects who wished to explore a tutorial area in more depth. The module contained extra information on the tutorial topics. There were no references to the section in the

tutorial text. The only prompts that the subjects had were the choices provided on the pop-menu display.

Table 71 shows the percentage of each group of subjects who did not access the More Information module, those who used it once and more than once. The majority of all three groups of subjects did not access the More Information module.

Table 71 Access to More Information Module

| Subjects | No assess | At least once | More than once |
|-----------------|------------------|----------------------|-----------------------|
| Beginner | 92% | 8% | 0% |
| Intermediate | 76% | 24% | 8% |
| Expert | 73% | 27% | 9% |
| Overall | 79% | 21% | 7% |

Quiz Module

The Quiz additional module was available from both the main and pop-up menus. It was an option which could be picked from every frame of the tutorial and was provided to assist subjects who wished to test themselves against the content of the tutorial. There were no references to the section in the tutorial text. However, a test question in the same format as those available in the Quiz module was included at the end of each module of the tutorial.

Tables 72 and 73 provide a comparison of the use of the Quiz module by all three groups of subjects. Table 72 shows the percentage of each group of subjects who did not access the Quiz module, those who used it once and more than once. The majority of the subjects used the Quiz module, with all the experts accessing the module at least once and 58% of them more than once. In contrast, none of the beginners and intermediates used the Quiz more than once. Table 73 shows a comparison of which route (main or pop-up menu) was taken to access the Quiz module for each of the three groups of subjects.

Table 72 Access to Quiz Module

| Subjects | No access | At Least Once | More than Once |
|-----------------|------------------|----------------------|-----------------------|
| Beginner | 21% | 79% | 0% |
| Intermediate | 24% | 76% | 0% |
| Expert | 0% | 100% | 58% |
| Overall | 15% | 85% | 19% |

Table 73 Access to Quiz Module from Main and Pop-Up Menus

| Subjects | Main menu | Pop-Up Menu |
|-----------------|------------------|--------------------|
| Beginner | 34% | 45% |
| Intermediate | 28% | 48% |
| Expert | 78% | 22% |
| Overall | 44% | 41% |

Exit Module

The Exit option was available from both the main and pop-up menus. It was thus an option from each frame of the tutorial. The module had the straightforward purpose of allowing subjects to leave the tutorial. Once the subjects had picked the Exit module, they were asked to confirm their choice and therefore could return to the tutorial if they had made a mistake. 71% of subjects entered Exit once and left the tutorial but 29% of subjects entered Exit and then chose to return to the tutorial. 7% of learners entered Exit more than once and then chose to return. These results indicate that many subjects had picked Exit by mistake.

Subjects entered Exit from a wide range of frames and modules. However, the four commonest choices of Exit location were tutorial frame (41%), main menu (25%), Quiz module (22%) and other additional modules (12%).

Mode of Operation

Learners' operation of the tutorial was classified into three types:

- (1) browsing: dipping in and out of modules only considering one or two frames of the tutorial in a module before moving to a new module;
- (2) completing the majority of a module before moving to a new module; and
- (3) mixture of browsing and completing a module.

Table 74 shows the different modes of operation by the three groups of subjects. It does indicate considerable differences between the groups of subjects. Only 8% of the beginner group operated in a browsing mode compared to 42% of the experts.

Table 74 Learners Modes of Operating (Graphics Experiment)

| Subjects | Browse | Complete | Mixture |
|-----------------|---------------|-----------------|----------------|
| Beginner | 8% | 42% | 50% |
| Intermediate | 16% | 44% | 40% |
| Expert | 42% | 29% | 29% |
| Overall | 22% | 38% | 40% |

Analysis of Different Screen Designs

Tables 76, 77, 78 and 79 show a series of comparisons of the different types of display against a text-only display. Each table provides the mean time users spent on each type of screen display along with the standard deviation, number of users (N), value of statistic t from a two-tail t test and the degree of significance compared with a text-only display.

Table 75 gives the mean times users spent on text-only displays. Table 76

shows the results of all the subjects while tables 77, 78 and 79 indicate the results of the beginner, intermediate and expert groups, respectively. In all cases, the comparison is made between the screen display and a text-only display. It is a one-for-one comparison in that individual users must have accessed both the text-only screen display and the type of display being compared. The null hypothesis in all cases was that there was no difference between the time spent on text-only and other displays.

Table 75 Users Mean Times

| Subjects | Mean Times (seconds) |
|-----------------|-----------------------------|
| All Subjects | 22.36 |
| Beginner | 20.76 |
| Intermediate | 26.17 |
| Expert | 19.32 |

Table 76 Comparison of Various Screen Designs with Text-Only Screens

| Screen Type | Mean Time | Standard Deviation | N | t | Significance |
|--|------------------|---------------------------|----------|----------|---------------------|
| Structured Text | 29.00 | 20.09 | 71 | -2.76 | 0.01% |
| Analogue Graphics | 16.68 | 6.72 | 47 | 2.05 | 0.05% |
| Part-Screen Analogue Graphics | 17.04 | 7.18 | 47 | 1.90 | 0.01% |
| Full-Screen Analogue Graphics | 14.13 | 4.98 | 29 | 2.16 | 0.05% |
| Graphics | 16.33 | 5.39 | 73 | 4.01 | 0.001% |
| Logical Graphics | 19.28 | 10.53 | 47 | 2.33 | 0.05% |
| Representative Graphics | 16.65 | 6.71 | 73 | 4.70 | 0.001% |
| Full-Screen Representative Graphics | 16.69 | 11.42 | 73 | 3.04 | 0.01% |
| Part-Screen Representative Graphics | 17.44 | 6.36 | 71 | 3.75 | 0.001% |
| One Part-Screen Representative Graphics | 17.44 | 7.81 | 68 | 4.78 | 0.001% |
| Two Part-Screen Representative Graphics | 20.75 | 9.61 | 45 | 1.48 | 0.20% |
| Four Part-Screen Representative Graphics | 19.37 | 8.10 | 69 | 2.29 | 0.05% |
| Window Text | 64.26 | 112.14 | 58 | -1.98 | 0.10% |
| Dictionary | 45.07 | 30.49 | 30 | -3.46 | 0.01% |
| Examples | 93.92 | 152.76 | 36 | -1.98 | 0.10% |
| Exit | 19.07 | 13.09 | 21 | -3.94 | 0.001% |
| Questions | 10.87 | 3.84 | 62 | 5.70 | 0.001% |

**Table 77 Comparison of Various Screen Designs with Text-Only Screens
(Beginner)**

| Screen Type | Mean Time | Standard Deviation | N | t | Significance |
|---|------------------|---------------------------|----------|----------|---------------------|
| Structured Text | 31.76 | 15.28 | 22 | -1.87 | 0.1% |
| Analogue Graphics | 16.31 | 6.60 | 22 | 1.70 | 0.2% |
| Full-Screen Analogue Graphics | 13.65 | 8.15 | 14 | 2.70 | 0.02% |
| Graphics | 15.66 | 5.61 | 24 | 2.97 | 0.01% |
| Representative Graphics | 15.85 | 5.07 | 24 | 2.87 | 0.01% |
| Part-Screen Representative Graphics | 15.99 | 5.03 | 22 | 3.29 | 0.01% |
| One Part-Screen Representative Graphics | 13.28 | 3.88 | 19 | 8.25 | 0.001% |
| Two Part-Screen Representative Graphics | 17.65 | 6.62 | 14 | 1.69 | 0.2% |
| Questions | 9.81 | 3.06 | 19 | 8.40 | 0.001% |

**Table 78 Comparison of Various Screen Designs with Text-Only Screens
(Intermediate)**

| Screen Type | Mean Time | Standard Deviation | N | t | Significance |
|--|------------------|---------------------------|----------|----------|---------------------|
| Structured Text | 33.6 | 26.31 | 25 | -1.93 | 0.1% |
| Analogue Graphics | 16.8 | 5.03 | 15 | 1.58 | 0.2% |
| Part-Screen Analogue Graphics | 17.1 | 5.21 | 15 | 1.56 | 0.2% |
| Full-Screen Analogue Graphics | 16.6 | 2.79 | 8 | 1.40 | 0.2% |
| Graphics | 17.8 | 6.17 | 25 | 2.32 | 0.05% |
| Logical Graphics | 23.0 | 12.54 | 19 | 1.40 | 0.2% |
| Representative Graphics | 18.6 | 8.90 | 25 | 2.73 | 0.02% |
| Full-Screen Representative Graphics | 18.0 | 11.89 | 25 | 3.12 | 0.01% |
| Part-Screen Representative Graphics | 19.2 | 8.34 | 25 | 2.17 | 0.05% |
| One Part-Screen Representative Graphics | 18.1 | 8.96 | 25 | 2.35 | 0.05% |
| Four Part-Screen Representative Graphics | 21.7 | 10.86 | 25 | 1.49 | 0.2% |
| Window Text | 61.1 | 48.14 | 21 | -2.27 | 0.05% |
| Dictionary | 64.0 | 34.98 | 8 | -2.90 | 0.05% |
| Examples | 103.0 | 124.21 | 15 | -1.79 | 0.01% |
| Exit | 20.6 | 17.51 | 9 | 2.50 | 0.05% |
| Questions | 10.5 | 3.28 | 19 | 3.27 | 0.01% |

**Table 79 Comparison of Various Screen Designs with Text-Only Screens
(Expert)**

| Screen Type | Mean Time | Standard Deviation | N | t | Significance |
|--|------------------|---------------------------|----------|----------|---------------------|
| Full-Screen Analogue Graphics | 11.8 | 3.48 | 7 | 1.98 | 0.1% |
| Graphics | 15.2 | 3.47 | 24 | 4.51 | 0.001% |
| Logical Graphics | 14.6 | 2.92 | 24 | 1.70 | 0.2% |
| Representative Graphics | 15.1 | 3.81 | 24 | 4.38 | 0.001% |
| Full-Screen Representative Graphics | 12.8 | 4.26 | 24 | 4.80 | 0.001% |
| Part-Screen Representative Graphics | 16.4 | 3.98 | 24 | 3.56 | 0.01% |
| One Part-Screen Representative Graphics | 14.8 | 7.01 | 24 | 2.50 | 0.02% |
| Two Part-Screen Representative Graphics | 17.1 | 5.55 | 13 | 1.93 | 0.1% |
| Four Part-Screen Representative Graphics | 17.1 | 3.29 | 24 | 1.79 | 0.1% |
| Window Text | 26.3 | 7.90 | 24 | -3.10 | 0.01% |
| Dictionary | 29.5 | 14.92 | 17 | -2.60 | 0.02% |
| Examples | 27.1 | 10.52 | 14 | -1.70 | 0.2% |
| Exit | 9.82 | 2.89 | 8 | 2.89 | 0.05% |
| Questions | 13.1 | 4.28 | 24 | 3.41 | 0.01% |

Tables 80, 81, 82 and 83 show a series of comparisons of the different types of graphic displays. Each table provides the number of users (N), value of statistic t from a two-tail t test and the degree of significance.

Table 80 shows the results of all the subjects while tables 81, 82 and 83 reveal the results of the beginner, intermediate and expert groups, respectively. It is a one-for-one comparison in that individual users must have accessed both types of graphics screen displays. In cases where there was no significant difference or insufficient users, a 'No' has been entered under all headings in the table. The null hypothesis in all cases was that there was no difference between the time spent on the different graphic displays.

Table 80 Comparison between different Graphic Screen Displays (All Subjects)

| Comparison | N | t | Significance |
|--|----------|----------|---------------------|
| Representative compared with Analogue Graphics | No | No | No |
| Representative compared with Logical Graphics | 47 | -1.93 | 0.1% |
| Analogue compared with Logical Graphics | 37 | -1.70 | 0.1% |
| Part-Screen Representative compared with Analogue Graphics | No | No | No |
| One Part-Screen Representative compared with Analogue Graphics | No | No | No |
| Two Part-Screen Representative compared with Analogue Graphics | 37 | 1.85 | 0.1% |
| Three Part-Screen Representative compared with Analogue Graphics | No | No | No |
| Four Part-Screen Representative compared with Analogue Graphics | No | No | No |
| Full-Screen Representative compared with Analogue Graphics | No | No | No |
| Full-Screen Representative compared with Logical Graphics | 47 | -3.86 | 0.001% |
| Part-Screen Representative compared with Logical Graphics | No | No | No |
| One Part-Screen Representative compared with Logical Graphics | No | No | No |
| Two Part-Screen Representative compared with Logical Graphics | No | No | No |
| Three Part-Screen Representative compared with Logical Graphics | No | No | No |
| Four Part-Screen Representative compared with Logical Graphics | No | No | No |

**Table 81 Comparison between different Graphic Screen Displays
(Beginner)**

| Comparison | N | t | Significance |
|--|----------|----------|---------------------|
| Representative compared with Analogue Graphics | No | No | No |
| Representative compared with Logical Graphics | No | No | No |
| Analogue compared with Logical Graphics | No | No | No |
| Part-Screen Representative compared with Analogue Graphics | No | No | No |
| One Part-Screen Representative compared with Analogue Graphics | No | No | No |
| Two Part-Screen Representative compared with Analogue Graphics | No | No | No |
| Three Part-Screen Representative compared with Analogue Graphics | No | No | No |
| Four Part-Screen Representative compared with Analogue Graphics | No | No | No |
| Full-Screen Representative compared with Analogue Graphics | No | No | No |
| Full-Screen Representative compared with Logical Graphics | No | No | No |
| Part-Screen Representative compared with Logical Graphics | No | No | No |
| One Part-Screen Representative compared with Logical Graphics | No | No | No |
| Two Part-Screen Representative compared with Logical Graphics | No | No | No |
| Three Part-Screen Representative compared with Logical Graphics | No | No | No |
| Four Part-Screen Representative compared with Logical Graphics | No | No | No |

**Table 82 Comparison between different Graphic Screen Displays
(Intermediate)**

| Comparison | N | t | Significance |
|--|----------|----------|---------------------|
| Representative compared with Analogue Graphics | 15 | 2.94 | 0.02% |
| Representative compared with Logical Graphics | 19 | -2.20 | 0.05% |
| Analogue compared with Logical Graphics | 15 | -2.00 | 0.1% |
| Part-Screen Representative compared with Analogue Graphics | 15 | 1.40 | 0.2% |
| One Part-Screen Representative compared with Analogue Graphics | 17 | 1.53 | 0.2% |
| Two Part-Screen Representative compared with Analogue Graphics | 13 | 1.95 | 0.1% |
| Three Part-Screen Representative compared with Analogue Graphics | No | No | No |
| Four Part-Screen Representative compared with Analogue Graphics | 15 | 1.55 | 0.2% |
| Full-Screen Representative compared with Analogue Graphics | No | No | No |
| Full-Screen Representative compared with Logical Graphics | 19 | -3.10 | 0.01% |
| Part-Screen Representative compared with Logical Graphics | No | No | No |
| One Part-Screen Representative compared with Logical Graphics | No | No | No |
| Two Part-Screen Representative compared with Logical Graphics | No | No | No |
| Three Part-Screen Representative compared with Logical Graphics | No | No | No |
| Four Part-Screen Representative compared with Logical Graphics | No | No | No |

**Table 83 Comparison between different Graphic Screen Displays
(Expert)**

| Comparison | N | t | Significance |
|--|----------|----------|---------------------|
| Representative compared with Analogue Graphics | No | No | No |
| Representative compared with Logical Graphics | No | No | No |
| Analogue compared with Logical Graphics | No | No | No |
| Part-Screen Representative compared with Analogue Graphics | No | No | No |
| One Part-Screen Representative compared with Analogue Graphics | No | No | No |
| Two Part-Screen Representative compared with Analogue Graphics | 9 | 1.49 | 0.2% |
| Three Part-Screen Representative compared with Analogue Graphics | No | No | No |
| Four Part-Screen Representative compared with Analogue Graphics | No | No | No |
| Full-Screen Representative compared with Analogue Graphics | 11 | -1.80 | 0.2% |
| Full-Screen Representative compared with Logical Graphics | 13 | -2.20 | 0.05% |
| Part-Screen Representative compared with Logical Graphics | No | No | No |
| One Part-Screen Representative compared with Logical Graphics | No | No | No |
| Two Part-Screen Representative compared with Logical Graphics | No | No | No |
| Three Part-Screen Representative compared with Logical Graphics | No | No | No |
| Four Part-Screen Representative compared with Logical Graphics | No | No | No |

Test Questions (Graphics Experiment)

The majority of learners (85%) chose to answer the Quiz questions. However, only the expert group (58%) used the test more than once. The other main difference between the groups was that all the experts used the Quiz, while both beginners (21%) and intermediates (24%) contained a minority who did not attempt the test.

The difference in behaviour of the three groups does seem to indicate that a greater knowledge and understanding of a subject provides the confidence to initially use the Quiz and later to return to it. For designers of computer-based learning, the evidence suggests that a self-test will be used extensively by all types of learner. The knowledgeable users are, however, likely to make full use of it and not restrict their access to a single visit.

The differences between the expert and the other two groups continue when considering where the learners chose to access the test questions from. 78% of experts selected the Quiz from the main menu, while only 34% of beginners and 28% of intermediates made the same decision. The users who chose the self-test from the tutorials could be considered to be influenced by the content and design of the learning material. In comparison, selecting the Quiz from the main menu is more likely to be independent of a particular tutorial module.

All three groups achieved a high degree of success with the test questions, which perhaps confirms that the learning material provided an effective tutorial of the subject. Tables 84, 85 and 86 provide a breakdown of both the percentage of users who answered each question and the percentage who answered the questions correctly at their first attempt. The application of the Mann-Whitney test on the results of the three groups shows that there was no significant difference between the beginner and expert groups and the beginner and intermediate groups. There was, however, a significant difference between the expert and intermediate groups (Observed value $U' = 46$, Critical value $U' = 51$ with two-tail test at the 0.10% level of significance). The experts answered significantly more questions right first time than the intermediates which, given their greater knowledge, is not unexpected.

Table 84 Beginners' Answers (Quiz)

| Question | Percentage of group | Right first time |
|-----------------|----------------------------|-------------------------|
| Question 1 | 71% | 88% |
| Question 2 | 71% | 88% |
| Question 3 | 71% | 100% |
| Question 4 | 71% | 29% |
| Question 5 | 71% | 59% |
| Question 6 | 71% | 41% |
| Question 7 | 71% | 100% |
| Question 8 | 29% | 100% |
| Question 9 | 29% | 100% |
| Question 10 | 42% | 50% |
| Question 11 | 21% | 60% |
| Question 12 | 29% | 100% |
| Question 13 | 29% | 71% |

Table 85 Intermediates' Answers (Quiz)

| Question | Percentage of group | Right first time |
|-----------------|----------------------------|-------------------------|
| Question 1 | 84% | 38% |
| Question 2 | 84% | 62% |
| Question 3 | 84% | 71% |
| Question 4 | 76% | 42% |
| Question 5 | 76% | 53% |
| Question 6 | 76% | 42% |
| Question 7 | 68% | 65% |
| Question 8 | 60% | 100% |
| Question 9 | 32% | 50% |
| Question 10 | 40% | 60% |
| Question 11 | 8% | 0% |
| Question 12 | 16% | 100% |
| Question 13 | 24% | 100% |

Table 86 Experts' Answers (Quiz)

| Question | Percentage of group | Right first time |
|-----------------|----------------------------|-------------------------|
| Question 1 | 92% | 55% |
| Question 2 | 83% | 90% |
| Question 3 | 83% | 90% |
| Question 4 | 83% | 45% |
| Question 5 | 83% | 100% |
| Question 6 | 83% | 55% |
| Question 7 | 83% | 90% |
| Question 8 | 63% | 73% |
| Question 9 | 8% | 100% |
| Question 10 | 29% | 71% |
| Question 11 | 29% | 100% |
| Question 12 | 8% | 100% |
| Question 13 | 8% | 100% |

The evidence from all three groups shows that the computer-based learning material is effective. Thus the underpinning screen design guidelines are likely to be sound.

4.4 Discussion of Results

The subjects for both the Colour and Graphics experiments were drawn from the working population of Great Britain. However, they tended to be more highly qualified and there were fewer subjects with no qualifications than in the working population as a whole.

The results of the Colour and Graphics experiments were analysed using a variety of statistical tests. The Colour experiment results were analysed using Mann-Whitney, Kruskal-Wallis and Jonckheere statistical tests, while the Graphic results were analysed using Mann-Whitney and t tests.

The analysis of the Colour results showed a significant trend in that as the number of colours increased, so did the mean time spent on each type of frame of the tutorial. In contrast to this trend, the number of frames accessed was significantly reduced in versions of the experiment using more than seven colours. This suggests that the motivational effects of colour are positive when up to seven colours are used but negative when more than seven colours are employed.

The Colour experiment was also used to test the research methods and the outcomes contributed to the design of the Graphics experiment. The main influences were to increase learners' freedom of choice, to employ a more sensitive self-test and to provide the guidelines for the use of colour in the Graphics experiment.

The subjects of the Graphics experiment were divided into three groups on the basis of their knowledge of the tutorial content. The three groups were beginner, intermediate and expert. This was in order to consider the effects of the users' prior knowledge.

A key measure was the time users spent on each type of screen design compared to the complexity of the display. Complexity was found to have a positive correlation with the time spent on the display. Regression analysis of complexity measurements of text-only and text and graphics displays produced regression equations. Complexity measurement allowed the time spent on particular displays to be predicted and compared with the actual results from the three groups of users. This showed that the learners' prior knowledge of the tutorial content affects their behaviour to different screen designs.

All subjects of the Graphics experiment completed a post-tutorial questionnaire. An analysis of the answers given by the three groups of users showed that there were significant differences between their attitudes to the tutorial, particularly in the degree of their reaction to individual questions. However, all three groups had a positive attitude to the learning material and the use of colour in the material. This provides some confirmation of the outcomes of the Colour experiment. In a similar way, Colour experiment subjects also regarded the tutorial as effective learning material. The subjects showed a preference for realistic (representational) graphics occupying a quarter to a half of a display.

Only a minority of learners using the Graphics experiment accessed all seven

modules of the tutorial. The pattern of access was sequential and followed the linear presentation of the modules in the menu. In addition to the seven tutorial modules, there were five additional modules, some of which were cross-referenced in the main tutorial. Access was improved by using cross-references. However, the cross-references did not guarantee the use of these additional modules by the majority of the users. A more influential effect appeared to be the users' prior knowledge of the additional module.

The learners' mode of operating the tutorial was classified into three types: browsing; completing whole modules; and mixed. There were considerable differences between the modes of operation of the three groups of learners. Beginners tended not to browse whereas experts did.

A series of comparisons were made using statistical tests between different types of display against a text-only display. Comparisons were made between the overall results from all subjects and from each group of users. In a similar way, comparisons were also made between different types of graphic display for all subjects and each group of users.

4.5 Conclusions

The experimental subjects were more representative of the working population than typically involved in screen design research identified in chapter 2. Subjects were satisfied with the learning material in both Colour and Graphics experiments.

The results of the Colour experiment, which were confirmed by the outcomes of the Graphics experiment, indicate that seven colours is the optimum number of colours for computer-based learning materials. In addition, the Colour experiment confirmed the overall design of the experiment with a few minor changes to improve the design of the Graphics experiment.

The results of the Graphics experiment showed that the learners' prior knowledge of the tutorial subject influenced their behaviour and attitude to the learning material. This was related to the time spent on different types of graphic display, the route taken through the learning material and the mode of operation.

However, there was no significant difference between learners' preference for the type and size of illustration, that is, they all preferred representative images occupying a quarter to a half of the screen.

Chapter Five

DISCUSSION OF RESULTS

5.0 Introduction

This chapter will consider the results of the research presented earlier in this thesis. The outcomes of both the Colour and Graphics experiments will be discussed. In particular, the chapter considers:

- (1) experimental design;
- (2) effects of increasing the numbers of colours used;
- (3) prior knowledge of the tutorial subject;
- (4) different types and sizes of graphic image; and
- (5) accessing additional modules.

5.1 Experimental Design

There is an extensive body of research evidence related to screen design. Unfortunately, it is limited in a number of ways. The balance of material is skewed with only text having an extensive well-supported body of research findings. Research into the use of colour and graphics is often in relation to media other than computer screens (for example, illustrations in books, colour television and aircraft displays). The research results are frequently based on experiments which did not have a learning focus, that is, the content of the displays was not designed to support learning and the subjects used were not learners. A key question is how much of the research can be transferred to the screen design of computer-based learning materials.

Computer-based learning consists of multiple screens of tutorial with features such as glossaries of terms, tests and additional information or help systems. It is thus an integrated environment of different types of display through which the learner moves and interacts. The survey of 41 research projects discussed earlier showed

that only 6 were computer-based, learning-focused, employed an integrated display and used a broad range of learners. This tends to reinforce the doubt expressed above about the transferability of the research literature to computer-based learning.

The research evidence presented earlier in this thesis provided a number of screen design guidelines relating to the use of text which were used in both Colour and Graphics experiments. The learners' success as measured by the comprehension test and their positive attitude to both experimental designs, revealed by the post-experimental questionnaire, does provide support for these guidelines. This is an indication that the research evidence for text is transferable to computer-based learning.

The basic approach to the experimental research undertaken in the current project was the creation of a real learning environment for the subjects. Learners using computer-based learning materials interact with a number of displays, not a single design. Screen designers need to take into account that a learner's reaction to a single display does not emulate the real world situation. Learners will normally encounter a range of displays such as graphics, text, monochrome and colour. Their behaviour is likely to be influenced by the screen design employed in this environment. Both the Colour and Graphics experiments were designed to measure subject responses to screen designs in an environment of multiple displays.

The Colour Experiment: Multiple displays of text and colour.

The Graphics Experiment: Multiple displays of text, colour and graphics.

During the two experiments, a number of hypotheses were investigated. These concerned the use of colour and graphics. All versions of the experimental tutorials were designed to be effective learning material. The overall structure of the material was not altered, except in the design of the display. In all versions of the Colour experiment, the users successfully learned the content of the tutorial as shown by the internal test. The role of colour throughout the material was to augment the organisation of the information and controls already provided by text and spatial cues. The consistently high scores on the comprehension test in both experiments

verifies the overall design of the material and supports the underpinning screen design guidelines used in the tutorial design. The learners' answers to the post-experimental questionnaire in both experiments also indicate an overall positive attitude to the design of the tutorial.

The outcomes of the Colour experiment were used to improve the structure of the Graphics experimental shell by:

- (1) providing the screen design guidelines for the use of colour within the Graphics experiment; and
- (2) improving the structure of the Graphics experiment by providing greater choice for the subjects and overcoming technical problems.

By utilising the results of the Colour experiment as the guidelines for the use of colour in the Graphics experiment, it was possible to provide a further test of the results. Learners in the Graphics experiment expressed satisfaction with the use of colour, thus providing further confirmation for the results of the Colour experiment.

The survey of 41 research projects discussed earlier in chapter two showed that experimental subjects were frequently not representative of the whole population of learners. The subjects for this research were drawn from a wider group than the normal focus on undergraduate students but they still had a number of differences from the whole population. They were, however, more representative than the typical subjects of the surveyed projects and tended to be broadly similar to those individuals who used open learning materials.

The experiments investigated the screen design of computer-based learning materials. The experimental materials were designed to be broadly similar to available learning materials. The survey of existing computer-based learning materials revealed that all three types of graphical image were employed, although analogical graphics only occasionally. Both full and part-screen images were used but in general, graphical images did not appear to be employed to their full potential.

In designing the Graphics experiment, all three types of graphics were used, employing a higher proportion than usually found in computer-based learning

materials, in order to make full use of their potential. Analogical images appeared more frequently than the current practice revealed in the survey, so as to provide a more meaningful test of their effects. The users' request for even more graphics on the post-experimental questionnaire does suggest that the use of graphics within the experimental learning design was valid, even though not in line with current practice. Perhaps current practice is failing to meet the learners' expectations.

The survey also showed that the typical approach to controlling the material was a combination of menus and icons. These methods were adopted in both experiments.

The use of colour in text-only computer-based learning materials was limited, with up to only four colours being employed. The Colour experiment provided six versions of the learning material using 3, 5, 6, 7, 9 and 11 colours, respectively. This allowed the effects of increasing the number of colours on learners' behaviour to be investigated. The consistently high scores on the comprehension test supports the general conclusion that the design of the learning material was sound.

The Colour experiment focused on two factors:

- (1) the use of colour in computer-based learning; and
- (2) the research methods to be employed.

The approach adopted during the Colour experiment was to compare learners' use of multiple versions of the same tutorial, differing only in the use of colour in each version.

The Graphics experiment was focused on a range of factors:

- (1) the use of graphics in computer-based learning;
- (2) a comparison of different types of graphic image in computer-based learning;
- (3) the use of different sizes of graphic image;

- (4) the complexity of the tutorial displays;
- (5) the effects of the learners' prior knowledge of the subject; and
- (6) the learners' use of additional modules not directly in the body of the main tutorial.

The major difference between the two experimental shells was that rather than produce a range of different versions, the Graphics experiment only used a single system. Within this single system, the different types of graphic image and other features were present and subjects were free to access each component as they worked through the material. The comparison was thus between the same subjects using different elements of the material, rather than different subjects using different versions of the material, as was the case in the Colour experiment.

The two experiments were inter-related in a number of ways:

- (1) the Colour experiment tested the underlying approach to the research (structure, software and design);
- (2) both experiments used a number of standard screen guidelines; and
- (3) the outcomes of the Colour experiment were applied to the Graphics experiment.

5.2 The Colour Experiment

The Colour experiment showed that the use of colour increased the mean time spent on each frame. Compared to the monochrome version, the increase was significantly different in versions 1, 3 and 5 using 5, 7 and 11 colours, respectively. Versions 2 and 4 did not show a significant difference, although there was an increase in mean time spent on each frame. This effect was apparent for all types of frame, that is, menu, question and tutorial frames. This effect started from version 1 which used

only limited amounts of colour. Further to version 0, which used black, white, and grey, version 1 offered users a red title bar and blue text. This indicates the power of colour to motivate learners to interact with computer-based learning materials.

The number of frames accessed by the users did not change significantly in versions of the tutorial using seven or fewer colours. However, in versions of the tutorial using more than seven colours, the number of frames used reduced significantly. The post-test questionnaire asked users' opinion of the amount of colour in the version they had used. For versions using seven or less colours, they responded positively to the view that 'Too little colour' was used, while for versions using more than seven colours, their reaction was that 'Too much colour' was used. Version 3 with seven colours scored 2.9 or a 'just right' score. The results show two main trends:

- (1) a general increase in the mean time spent on each frame of tutorial; and
- (2) for tutorials using more than seven colours, a negative reaction from users (too much colour), and a decrease in the number of frames accessed.

There are several possible explanations for these results. One is that colour adds interest to the tutorial which increases with the use of colour. The Jonckheere Trend test showed a significant trend across versions 0, 1, and 3 and versions 0, 3, and 5 for the mean time spent on each frame. An alternative explanation is that colour makes the reading or understanding more difficult, so increasing the time spent on each frame. The users' post-test questionnaire results do not support this alternative view. The questionnaire showed no major change between monochrome and coloured versions in the users' assessment of the tutorial. Users did not indicate that colour made the material more difficult to use, read, learn from, or find their way around. In fact, coloured versions scored higher than the monochrome in many cases.

The versions which used colour extensively (more than seven colours) resulted in the users' accessing significantly less frames. This seems to indicate that too much colour may start to have a negative effect on the users. The mean time to

complete the tutorial showed an interesting trend, in that it increased from the base level in the monochrome version up to and including version 3 (7 colours). In the last two versions (4 and 5) it fell below the time taken for the monochrome version. Kelley (1988) reported that users found a tutorial using fourteen colours in an indiscriminate way tiring and confusing to use. However, in the tutorial using three and five colours users found colours helpful in indicating different types of information. Kelley (1988) did not examine tutorials using more than five colours except for fourteen colours. He also found that users took less time to complete the fourteen coloured version than the other versions. This agrees with the results reported in this thesis. Version 5 (eleven colours) used colour in a systematic way and users reported no particular feelings of tiredness or confusion. This is probably due to the systematic use of colour to aid information separation, rather than an indiscriminate use.

The test built into the tutorials did not produce results significantly different comparing each version. This agrees with Kelley (1988) who reported no significant differences in comprehension test results from both a built-in test and a post-tutorial paper and pencil test.

The results can be explained by combining the alternative explanations, that colour increases interest and motivation so that users are encouraged to spend more time studying individual frames. However, as the number of colours is increased beyond seven, users find it more distracting and so need to spend more time on each frame to understand it. This results in fewer frames being accessed as users try to quickly finish the tutorial. In both cases, however, the users are successfully learning the content of the tutorial as shown by the built-in test scores. This is not surprising as the monochrome (version 0) tutorial was designed as an effective learning module. Colour was used throughout the material to add to the organisation of the information and controls already provided by text and spatial cues. The consistently high scores on the comprehension test perhaps verifies the overall design of the material and supports the screen design guidelines used.

5.3 Changes to the Design of Experimental Shell

The Colour tutorial was essentially a menu-driven tutorial. The main menu gave users four choices - from moving to the two sub-menus representing the tutorials on notetaking from books and lectures, to an introduction to whole subject or to the self-test. A user was free to choose but was always required to return to the main menu in order to make a new choice, unless the choice was to exit the tutorial. The Exit led users to the questionnaire on their opinions of the learning materials they had just used.

The overall aim of the tutorial was to emulate the design of learning material that is currently in use. The review of learning materials undertaken as part of the research provided the basic model. The structure was reasonably typical but the use of the Colour tutorial served to highlight the changes required in the Graphics tutorial to bring it closer to this goal.

Figures 10, 11 and 24 show the structure of the Graphics experiments learning material and serve to illustrate the changes brought about by the analysis of the Colour tutorial (increased user freedom, richer learning environment and technical changes).

5.4 The Graphics Experiment

The Graphics experiment was aimed at investigating the use of graphic images in computer-based learning. It compared different types and sizes of image as well as text-only displays. This section will consider:

- (1) complexity;
- (2) the prior knowledge of the subject;
- (3) structured text;
- (4) graphic screen displays;

- (5) the routes taken through the learning material;
- (6) the additional modules; and
- (7) the operation of the tutorial.

Complexity

The positive correlation between the time spent on each graphics and text and text-only screen offers a useful method of comparing different displays. Many publications offer the simple advice to designers to keep displays uncluttered, clear and straightforward. This is far from easy to achieve, not least because the standard suggested is subjective. A measure of complexity has the potential to provide greater objectivity for a designer.

The experimental evidence shows that there is no significant difference between the predicted mean time and the overall mean time, that is, the mean of all three groups is not significantly different from the mean time predicted by the complexity measure. This overview does however disguise a more complex situation between the groups of learners.

There were significant differences between the predicted and actual mean times for the beginner and expert groups, while the intermediate group showed none. Both the beginners and experts spent less time on the displays than predicted. The principle difference between the groups was their understanding of the tutorial subject. Beginners may have spent less time on the display because they did not have sufficient understanding of the subject to appreciate the full content of each display, while the experts, having far greater understanding of the subject, may not have needed to spend time studying the screen since they rapidly realised that they already understood its contents.

This analysis is supported by the experimental results which show that the experts tended to spend less time on the displays than the beginners.

Knowledge of the Subject

The results reported earlier in this thesis indicate that prior knowledge of the tutorial subject affects learners' behaviour on interaction with different screen designs.

The evidence of the experiments shows this difference in a number of ways:

(1) Tutorial Questionnaire

All three groups of learners revealed no significant difference to the experimental questions on their preferences as to the size and type of illustrations. The three groups of learners clearly showed a preference for realistic illustrations which occupy a quarter or half of the screen. Equally, all three groups of learners clearly do not like analogical illustrations.

There was, however, a significant difference between the responses of the beginners and expert groups to the post-tutorial questionnaire, although both groups were generally positive to the design of the materials. The beginners' responses show that they had a more positive attitude to the tutorial than the experts with the exception of colour, illustrations and the display of information. Beginners found the information displayed more difficult to understand than the experts.

(2) Use of Learning Material

There was a considerable difference in how the groups used the learning material (table 55). Experts (42%) were more likely to browse the material than either the intermediates (16%) or the beginners (8%). In a similar way, the experts (29%) were least likely to complete whole modules.

(3) Routes through Learning Material

The expert group demonstrated a more complex pattern of behaviour than either of the other groups who essentially followed a sequential pattern. Only initially did the experts behave sequentially.

(4) Self-Test Questions

The majority of learners (85%) chose to answer the quiz questions, but only the expert group (58%) used the test more than once. The other main difference between the groups was that all the experts used the Quiz, whereas both beginners and intermediates contained a minority, 21% and 24%, respectively, who did not attempt the test.

(5) Additional Modules

Beginners made the least use of the additional modules. However, the use of cross-references in the tutorial increased the use of the module with beginners more than with the other two groups of learners.

The three groups of learners showed considerable variations from where in the tutorial they chose to access the Quiz and Dictionary modules. The pattern of access was complex.

(6) Screen Design

The evidence suggests that the learners' reaction to different screen displays was influenced by their prior knowledge. For example:

- (i) Beginners spent significantly more time on the structured text displays than the complexity measure would predict. The experts did not.

- (ii) The beginner group spent more time on the full screen representative graphics than complexity would predict, but not significantly more so (table 81). The intermediate and expert groups show a significant difference, in that they spent less time on these screens.

Summary

Prior knowledge of the tutorial subject clearly does exert an influence on the behaviour (their use of computer-based learning materials) of the learners. The experts showed a willingness to browse and explore the tutorial. They made non-sequential choices of route in comparison to both the other groups, who tended to be sequential and to complete each section of the tutorial before moving to the next. This may be a result of the increased confidence which comes with greater knowledge and understanding of the subject. The experts were the only group in which every member accessed the Quiz - some of whom used it more than once. Both the other groups had a minority who did not access the Quiz and neither of these groups used it more than once.

The beginners were the group with the least understanding and thus probably the least confidence, a view supported by their use of the additional modules which was very limited and the lowest of all the groups. However, the subjects' behaviour could be influenced by devices such as cross-references which did increase their use of the additional modules.

Analysis of Structured Text

A comparison between text screens and screens displaying structured text, based on complexity measures, reveals that structured screens score 116% compared with 100% for text screens. Using these measures to predict how much time learners should spend on structured screens shows that in practice, they spend more time on structured displays than the time complexity would predict. It is therefore not surprising that the t test analysis shows that, overall, learners spent significantly more time on structured text than text-only displays (table 87). Both the beginner and intermediate groups spent significantly more time on the structured text screen.

However, the expert group did not.

This may be due to the nature of structured text which can be displayed in many ways (for example, number of columns, number of rows and abbreviations used). Learners would need to spend some time studying the format of the structure before they could concentrate on the meaning of the information shown. This need not be a negative factor, since the organisation of the information frequently provides a learner with an understanding of the relationship between different elements of the information. An expert would more quickly identify the information in the structured display as familiar knowledge and move on, thus spending less time on the display.

The extra time spent by learners does indicate that they are making an effort to understand the meaning of the knowledge displayed, both in the form of the structure and the text shown.

Table 87 Comparison of Degree of Significance against Learners Group for Text compared to Structured Text Screens (n/s = not significant)

| Subjects | Significance |
|-----------------|---------------------|
| Beginner | 0.10% |
| Intermediate | 0.10% |
| Expert | n/s |
| Overall | 0.01% |

Analysis of Graphic Screens

A comparison of text screens and screens displaying graphics based on complexity measures reveals that graphic screens score 90%, compared with 100% for text screens. Using these measures to predict how much time learners should spend on graphic screens shows that in practice, they spend less time than the time complexity would predict.

The explanation of why learners should spend less time on displays containing graphic images than the predicted time based on complexity measures is not

straightforward. Graphic displays group together a wide variety of images (representational, logical and analogical pictures) and therefore the effect may be due to a particular type of image or a combination of these effects. In simple terms, learners are spending less time on graphic displays and are possibly not assimilating all the information contained in the images displayed. A second explanation is that pictorial information may be easier to understand than textual information (a picture is worth a thousand words). All three groups of learners showed a significant difference between the time spent on graphics and text screens (see table 88 - Degree of Significance based on t test).

Table 88 Comparison of Degree of Significance against Learners Group for Text compared to Graphics Screens (t test)

| Subjects | Significance |
|-----------------|---------------------|
| Beginner | 0.01% |
| Intermediate | 0.05% |
| Expert | 0.001% |
| Overall | 0.001% |

Analysis of Representational Graphics

Representational images are displays showing a realistic picture connected with the subject - 'What you see is what you get' - and on a complexity basis, predicted time spent should be closer to actual time than with graphic displays in general, since the image structure (logical graphics) and links to other experience (analogical graphics) are not key factors. Overall, this was observed with learners spending 2.26 seconds less time than the prediction. The total graphics result was that learners spent 3.85 seconds less than the complexity prediction.

It is worthwhile considering whether the method of producing the comparison for graphics screens is reasonable. A text-only screen, which is the basis of the complexity comparison, does not contain any images. Perhaps a more valid

comparison would be that of graphic screens with representational displays. Table 89 below gives the predictions of time based on text and representational screens.

This approach certainly closes the gap between the actual time spent and the predicted time. However, the trend is still broadly similar to the pattern produced by the predictions based on the complexity of text screens. All three groups of learners showed a significant difference (see table 90 - Degree of Significance based on t test) in that they spent less time on representational graphics than on text screens. This was particularly marked with the beginners and expert groups.

Table 89 Predictions of Time based on Text and Representational Graphic Screens

| Screen Type | Complexity | Actual Time | Text Predication | Graphic Prediction |
|-------------------------------|------------|-------------|------------------|--------------------|
| Text | 39.43 | 22.36 | 22.36 | 19.68 |
| Representation Graphics | 33.12 | 16.56 | 18.82 | 15.56 |
| Logical Graphics | 56.00 | 19.28 | 31.82 | 27.99 |
| Analogue Graphics | 36.55 | 16.68 | 20.77 | 18.27 |
| Full-Screen Analogue Graphics | 36.00 | 14.13 | 20.46 | 17.99 |
| Part-Screen Analogue Graphics | 31.3 | 17.04 | 17.79 | 16.65 |

Table 90 Comparison of Degree of Significance against Learner Group for Text compared to Representational Graphics (t test)

| Subjects | Significance |
|--------------|--------------|
| Beginner | 0.01% |
| Intermediate | 0.02% |
| Expert | 0.001% |
| Overall | 0.001% |

Analysis of Full-Screen Representational Graphics

A comparison of text screens and screens displaying representational graphics occupying the majority of the display reveals that the graphic screens score 85% compared with 100% for text screens. Learners spent 2.28 seconds less on each screen than the comparison would predict. The comparison with a text-only display is probably not the most reliable, since these graphic screens will contain only a small amount of text.

The intermediate and expert groups showed a significant difference in that they spent less time on the full-screen representative graphics than the text-only screens. The beginners group spent more time on these screens than complexity would predict but the difference was not significant (see table 91).

Table 91 Comparison of Degree of Significance against Learner Group for Text compared to Full-Screen Representational Graphics (t test)

| Subjects | Significance |
|-----------------|---------------------|
| Beginner | n/s |
| Intermediate | 0.01% |
| Expert | 0.001% |
| Overall | 0.01% |

Analysis of Part-Screen Representational Graphics

A comparison of text screens and screens displaying representational graphics occupying only part of the display reveals that the graphic screens score 84% compared with 100% for text screens. Learners spent 1.34 seconds less on each screen than the comparison would predict. This is perhaps not surprising since the pictures occupy only a small area of the screen and text may form the majority of the display. The results of the analysis of the types of part-screen displays does not entirely support this conclusion.

If the graphics contribution is removed from the measurement of complexity, the result is:

- (1) one-part: 27.59;
- (2) two-part: 22.63;
- (3) three-part: 28.00; and
- (4) four-part: 29.00.

This would indicate that any difference in result from one, three or four part-screen representative graphics is likely to be due to the graphic images since these complexity scores are very similar.

1. One Part-Screen Representational Graphics

A comparison of text screens and screens displaying a single representational graphic occupying only part of the display reveals that the graphic screens score 92% compared with 100% for text screens. Learners spent 3.05 seconds less on each screen than the comparison would predict which is surprising, since the majority of the display is text and it would be reasonable to expect that the actual time spent and predicted would be closer to the general value of 1.34 seconds. This may be evidence that the information contained in a representative graphic is easier to extract than the equivalent text display. All three groups of learners spent significantly less time on this type of display than complexity would predict (see table 92).

Table 92 Comparison of Degree of Significance against Learner Group for Text compared to one Part-Screen Representational Graphics (t test)

| Subjects | Significance |
|-----------------|---------------------|
| Beginner | 0.001% |
| Intermediate | 0.05% |
| Expert | 0.02% |
| Overall | 0.001% |

2. Two Part-Screen Representational Graphics

A comparison of text screens and screens displaying two representational graphics occupying only part of the display reveals that the graphic screens score 77% compared with 100% for text screens. Learners spent 3.64 seconds more on each screen than the comparison would predict. The difference probably rests with the learner having to spend time interpreting two separate images which are not fully integrated together as a whole screen image would be. This task will occupy some time and may account for the extra time spent. It also indicates that learners are attempting to use both images as sources of information on the subject. The images appear to provide sufficient motivation to make the extra effort to use both of them.

The beginner and expert groups showed a significant difference in that they spent more time on the two part-screen representative graphics than the text-only screens. The intermediate group also spent more time on the display but the difference was not significant (see table 93).

Table 93 Comparison of Degree of Significance against Learner Group for Text compared to two Part-Screen Representational Graphics (t test)

| Subjects | Significance |
|-----------------|---------------------|
| Beginner | 0.02% |
| Intermediate | n/s |
| Expert | 0.10% |
| Overall | 0.20% |

3. Three Part-Screen Representational Graphics

A comparison of text screens and screens displaying three representational graphics occupying only part of the display reveals that the graphic screens score 91% compared with 100% for text screens. Learners spent 1.64 seconds more on each screen than the comparison would predict. This would tend to support the conclusion that each individual image requires effort to interpret. However, it is

reasonable to expect that since these displays have three images, learners should spend longer than on a display containing only two illustrations. If the original concept is correct, perhaps learners find three images too great a burden to interpret and only some of the illustrations are used.

4. Four Part-Screen Representational Graphics

A comparison of text screens and screens displaying four representational graphics occupying only part of the display reveals that the graphic screens score 104% compared with 100% for text screens. Learners spent 3.93 seconds less on each screen than the comparison would predict. If the analysis of two part-screen representative illustrations is correct, then learners should spend more time on four-screen images than the complexity measure would predict, that is, a learner would have to interpret four individual images which would only be partially integrated. However, this was not observed.

One possible explanation could be that the effort required to interpret four individual images is greater than most learners are prepared to make. The learners are therefore only spending their time interpreting some of the images. This explanation is supported by the results from the displays containing three representative graphics.

The intermediate and expert groups showed a significant difference in that they spent less time on the four part-screen representative graphics than the text-only screens. The beginner group also spent less time on the display but the difference was not significant (see table 94).

Table 94 Comparison of Degree of Significance against Learner Group for Text compared to four Part-Screen Representational Graphics (t test)

| Subjects | Significance |
|-----------------|---------------------|
| Beginner | n/s |
| Intermediate | 0.20% |
| Expert | 0.10% |
| Overall | 0.05% |

Conclusion of Representative Graphics

The evidence suggests that learners' behaviour in relation to representative graphics depends on the number of illustrations shown.

- (1) Learners spend less time on full-screen and single part-screen images than the complexity measure would predict. This leads to the conclusion that learners extract information from a single representative graphic more easily than the complexity measure would suggest. This would be confirmed by the simple consideration that a single illustration contains the same information no matter how large it is, while this method of measuring complexity would increase with the size of the image.
- (2) Considering displays with 2 or 3 representative images, the results suggest that learners are prepared to spend more time on these displays than complexity would predict (table 95). Perhaps multiple illustrations take longer to interpret due to some form of image interference. What is clear is that a learner is willing to make the extra effort to extract the information. The images may provide the motivation for this extra effort.

- (3) The display with four illustrations reveals that learners have reached the point at which they are not prepared to make the effort to use all the information contained in the four images. This may be due to the fact that the visual complexity of this type of display hinders a learner's ability or desire to separate out each image or perhaps that the display simply de-motivates the learners.

Table 95 Representative Graphics (Time Differences with Predictions)

| Types of Display | Time Differences with Prediction |
|-------------------------|---|
| Full-Screen | -2.28 seconds |
| One Part-Screen | -3.05 seconds |
| Two Part-Screen | +3.64 seconds |
| Three Part-Screen | +1.64 seconds |
| Four Part-Screen | -3.93 seconds |

Logical Graphics

A comparison of text screens and screens displaying logical graphics reveals that logical screens score 142% compared with 100% for text screens. Learners spent less time on logical graphic displays than the time complexity would predict - 12.54 seconds less time than the predicted time.

This may be due to the nature of logical graphics which are complex and need a reasonable effort to translate. In a similar way to displays using structured text, the graphics structure contains information which is valuable to a learner. Learners would need to spend some time studying the format of the structure before they could concentrate on the meaning of the information shown. This would lead to the assumption that learners would spend more time than the prediction. However, the results show the opposite, with learners spending considerably less time than that predicted.

The time spent by learners does indicate that they are not making the required effort to understand the meaning of the knowledge displayed, both in the form of the structure and the graphic shown. This is in direct contrast to structured text, which is the text equivalent of a logical illustration. All three groups of learners spent considerably less time on the logical displays than complexity would predict. That is:

- (1) beginners -16.56 seconds;
- (2) intermediates -8.85 seconds; and
- (3) experts -17.21 seconds.

This would indicate that the value of logical graphics to convey detailed complex information is being lost by the learners' lack of motivation to make the effort to decipher the logical graphic displays.

Analogical Graphics

Analogical graphics are illustrations, intended to link the learners' prior knowledge to the subject being studied. For example, showing a picture of a metal block might serve to make a learner recall memories of heavy objects. Learning should be reinforced by these types of connection between the learning material and other knowledge. If this is happening, the results should show more time being spent on analogical illustrations than the prediction based on complexity measures since learners are also considering their existing knowledge and making the links.

The results however do not support this analysis, since less time was spent on analogical images than the prediction (4.09 seconds). Learners in all three groups are probably not utilising the power of analogies to aid their learning.

Full-Screen Analogical Graphics

A comparison of text screens and screens displaying full-screen analogical graphics reveals that analogical screens score 92% compared with 100% for text screens. In practice, learners spent less time on full-screen analogical displays than the time

complexity would predict. They spent 6.33 seconds less time than the prediction. This seems to show that learners are not making the links with other memories. The potentially powerful aid to learning that analogies can serve is not being utilised by learners.

Part-Screen Analogical Graphics

A comparison of text screens and screens displaying part-screen analogical graphics reveals that analogical screens score 80% compared with 100% for text screens. In practice, however, learners spent less time than the time complexity would predict. They spent 0.75 seconds less time than the prediction. This seems to show that learners are not making the links with other memories but are considering the image in a similar way to a representative graphic (a simple factual illustration). If this is correct, then since the images are not intended to serve this purpose, they may well cause learners some confusion.

Learners' Choice of Route through the Tutorial

All three groups of learners - beginner, intermediate and expert - showed a clear preference to follow a sequential route through the tutorial. For example, only 8% of beginners accessed the seventh module and they all made this module their seventh choice. 60% of the intermediate group accessed the second module and they all chose it as their second preferred choice.

This tendency to follow the sequence of modules in the order they are displayed in the main menu was particularly strong in the beginner and intermediate groups with the first two modules. Table 96 shows the percentage of users who chose the module in the order of the main menu as compared to the total who accessed the module. This sequential behaviour of learners has been reported by a number of researchers (Elliot, 1975; Sasser and Moore, 1984; Bolton and Peck, 1991; Schuerman and Peck, 1991a; Schuerman and Peck, 1991b).

Table 96 Learners' Choice of Route through the Tutorial

| Module | Beginner | | Intermediate | | Expert | |
|--------|----------|-----|--------------|------|--------|-----|
| | Order | All | Order | All | Order | All |
| 1 | 71% | 79% | 100% | 100% | 100% | 95% |
| 2 | 50% | 71% | 60% | 60% | 42% | 56% |
| 3 | 58% | 79% | 52% | 68% | 21% | 73% |
| 4 | 50% | 79% | 46% | 60% | 8% | 66% |
| 5 | 21% | 50% | 25% | 44% | 8% | 47% |
| 6 | 8% | 29% | 25% | 44% | 0% | 38% |
| 7 | 8% | 8% | 25% | 44% | 0% | 27% |

The best prediction of the use of a module is the order of display of the main menu for all three groups. The majority of subjects in the beginner and intermediate groups did follow a sequential path through the tutorial. However, the expert group participants demonstrated more complex behaviour with only the first and second modules being accessed by the majority of learners sequentially. The remaining modules were accessed in sequence by a minority of the expert learners. In particular, modules six and seven were not accessed in order by any of the expert group.

The main difference between the groups was their knowledge of the subject. The experts' understanding of the subject may have allowed them to choose modules which particularly interest them and ignore others. This theory is partially supported by the evidence of no one module being accessed by 100% of the group and no expert learner accessing all seven modules. Each expert user chose to leave out at least one module. The other two groups were restricted in their choices due to their limited knowledge of the subject. They may well have needed to accept the designer's decision on the order of modules to access.

For a designer, this sequential behaviour is very important since it can be used to ensure that key elements are considered by the learners. The learning material can be presented in order of importance and the majority of learners will use the material

in the predicted order. The expert learner's behaviour, however, does not follow a direct sequential pattern after the initial modules but the use of the modules is predictable on the basis of their order. To conclude therefore:

- (1) learners will tend to follow the sequence of the tutorial as shown by the main menu;
- (2) learners with expert knowledge of the subject will only initially follow the sequence of the tutorial; and
- (3) the best predictor of the numbers of learners who will access a given module is the order the module appears on the main menu.

Additional Modules

There were five additional modules included in the Graphics experimental system - Examples, More Information, Dictionary, Quiz and Exit. All the modules were accessed in similar ways. The Dictionary, Quiz and Exit modules were accessed from both main and pop-up menus while Examples and More Information were only accessed from the pop-up menu from within the tutorial. The main difference in their use was perhaps the degree of prior knowledge of the users. It is probably reasonable to assume that the users, who are all computer literate, were familiar with the Dictionary and Exit facilities since these are common to many computer packages. The choice of these modules was probably a result of additional factors other than simply what the learners encounter using the system. The behaviour of the learners towards Examples, More Information and Quiz, however, was probably due to a greater extent on what they encountered in the tutorial.

The Examples module was signposted throughout the tutorial by a series of cross-references in the text which encouraged the learners to access the appropriate example. The Quiz was not cross-referenced but each tutorial module contains a question which may have provided encouragement to try the self-test. Perhaps equally important was the users' awareness that the material was intended to assist

their learning of the subject and a test of some sort is usually associated with education and training. This may have been a hindrance or an aid depending on the users' previous experience. The Quiz module was also available from both main and pop-up menus while Examples and More Information were only available from the pop-up menus. The More Information facility was not cross-referenced in any direct way except as an item on the pop-up menu. It could thus be used as a direct comparison with Examples to assess the effectiveness of tutorial cross-references.

Examples and More Information Modules

The tutorial contains 13 cross-references to the Examples module. Examples could be accessed from any frame with a reference by using the pop-up menus. In contrast, the More Information module had no cross-references to inform users of its existence but it was accessed in an identical manner to the Examples module. Table 97 gives the comparison of learners using Examples and More Information modules against all three groups of users. It clearly shows two effects:

- (1) the beginner group accessed the additional modules less than either of the other two groups; and
- (2) the use of cross-references increased access by the beginner group considerably more than the other two groups, even though the overall access was still less.

This shows that the use of Examples is more extensive than the More Information module. It does suggest that the use of references in the tutorials is an effective way of encouraging use of an additional module. There are, however, considerable differences in the rate of accessing the Examples module between the three groups. Even with the use of references, only a minority (30%) of the beginner group used the Examples module.

Table 97 Comparison of Examples and More Information Modules

| Subjects | Examples Access | More Information Access | Access Ratio |
|-----------------|------------------------|--------------------------------|---------------------|
| Expert | 60% | 27% | 2.22:1 |
| Intermediate | 58% | 24% | 2.42:1 |
| Beginner | 30% | 8% | 3.75:1 |

Only 8% of the beginner group used the More Information module, so the references did substantially increase the use of the additional modules although they did not guarantee that a majority of the group will access the module.

Approximately a quarter of the expert and intermediate groups (27% and 24%, respectively) used the More Information module. This increased to 60% of experts and 58% of the intermediate group using the Examples modules which does indicate that cross-references do increase the use of the additional modules.

The conclusion of this analysis is that the use of cross-references in the tutorial does increase the use of additional modules. However, it is certainly not a guarantee that the module will be accessed and a beginner is still more likely to ignore the additional module. Nevertheless, without cross-references it is doubtful whether the use of additional modules would be worthwhile for beginners. This is especially relevant when the number of references encountered is considered. Beginners encountered an average of 5.66 references during their use of the tutorial compared to 6.75 for experts and 4.25 for the intermediate groups. Multiple references still only achieved a 30% success rate for beginners. Table 98 shows the accumulated success rate of encountering cross-references. No member of the beginner or intermediate groups accessed Examples without first encountering a cross-reference. However, 10% of the expert group used Examples without encountering a cross-reference to the module. After encountering three cross-references or less, 30% of beginners, 58% of intermediates and 60% of experts had used Examples.

Table 98 Accumulative Effect of References

| Subjects | References | | | |
|--------------|------------|-------|--------|-------|
| | None | First | Second | Third |
| Beginner | 0% | 10% | 20% | 30% |
| Intermediate | 0% | 17% | 25% | 58% |
| Expert | 10% | 30% | 40% | 60% |
| Overall | 3% | 19% | 28% | 49% |

The use of the additional module did increase when users encountered multiple references, as shown in table 98. For a screen designer, this shows the need to provide many cross-references to the additional modules in order to ensure their use. However, this effect was only maintained for the first three cross-references encountered. The experts, intermediates and beginners who encountered four or more references did not access the additional modules. The value of the reference stopped once three had been encountered.

Additional modules are intended to provide a learner with additional choices and assistance but are expensive to produce. The beginner group would be expected to gain most from these extra facilities but this is the group that used them least. The sequential behaviour discussed earlier is consistent with this limited use of the additional modules. The useful role of aiding learning provided by optional facilities such as Examples will be lost if cross-references from the main tutorial are not employed.

There is a need for further research into the most effective type of cross-reference. This research has considered only one type of reference, a text clue.

Dictionary and Quiz Modules

The Dictionary provided a means of accessing a glossary of terms, while the Quiz allowed the learners to test their understanding of the tutorials. Both modules were

accessed in identical ways from both the main and pop-up menus. Both of these additional modules were used to a far greater extent than the earlier ones (Examples and More Information). 58% and 85%, respectively of the users of the system accessed the Dictionary and Quiz. Across all three groups of users, access to both modules was consistently high, probably reflecting the users' previous knowledge of the two types of module. The value of these facilities could be judged more easily than the additional modules relating to Examples or More Information.

The modules could be accessed from two different points. A comparison between the access points picked reveals a considerable difference between and within the two modules. The Dictionary was primarily entered from the main menu with almost 80% of users accessing it this way. All the beginners accessed the Dictionary module from the main menu and approximately 84% of the experts. Only the intermediate group showed any strong tendency to use the pop-up route. Even the intermediate group revealed approximately 63% of subjects accessing Dictionary from the main menu. This pattern of access suggests that learners do not use the Dictionary when they come to new terms during the tutorial but wait until they return to the main menu.

The access to the Quiz shows a more complex pattern. The beginner and intermediate users chose to enter the Quiz from the pop-up menu in ratios of 1.3 to 1 and 1.7 to 1, respectively. This is a reasonable preference for access from the tutorials. The expert group showed a complete reversal of this trend with a ratio in favour of the main menu of 3.5 to 1. The overall picture was almost an equal split of users choosing to enter the Quiz from either the main or the pop-up menus. Reasoning from a designer's view point, this would suggest that learners might wish to test themselves during a module or on completing a module. The overall evidence would support this conclusion.

Exit Module

The Exit module serves the simplest function of all the additional facilities. Its purpose is to allow users to leave the system but with the option to change their minds if they have accidentally chosen this facility. During the experiment, the

majority of the users (71%) correctly entered Exit and left the system. However, 29% of the subjects entered the Exit facility and then returned to the tutorial and 7% of users entered Exit more than once and chose to return. This serves to illustrate that some form of safety net needs to be provided to avoid users accidentally leaving the system.

An analysis of the Exit points showed that users opted to leave the system from a wide range of positions. The most common Exit point (41% of users) was from one of the tutorial modules. This is not surprising since this constitutes the major part of the system. A quarter of users chose to Exit from the main menu which does indicate that having completed a section they returned to the main menu and decided they had no further modules they wished to use. A similar number (22% of users) left from the Quiz, which again can be explained in terms of completing the tutorial as far as they wanted to go. The most interesting exit point is from the additional modules which, considering the level of their use, was chosen by 12% of users as their point of departure. The key design point would seem to be the provision of an Exit facility from all elements of the tutorial.

5.5 Operation of Tutorial

The system was used in three distinct ways. Learners in all three groups adopted the different approaches to varying degrees. The learners either browsed the tutorial, completed an entire module before moving to the next module or used a mixture of browsing and completing modules.

The choice of approach varied between the three groups of learners. Only 8% of the beginners and 16% of the intermediates chose to operate in an exclusively browsing mode, in contrast to 42% of the experts. Beginners and intermediates preferred to use the system by completing each module in turn or by a mixture of browsing and completing modules.

A number of researchers (O'Connor, 1985; McAleese, 1989) have commented that browsing is not a random process but rather is linked to the prior knowledge and browsing experience of the subject. It is a structured process which

is linked to the motivation of users and the initial factors presented to them, such as the computer display. Individuals browsing in a library do not randomly select books. Their choices are motivated by their own knowledge, understanding and experience.

The key difference between the groups was the degree of knowledge of the subject. The observed behaviour can be explained in terms of knowledge by equating browsing with the purposeful search for material with which the learner is not familiar. The experts are more likely to adopt this approach than the other groups since their level of understanding and knowledge of the subject is greatest. The use of browsing in conjunction with other approaches can be explained by learners seeking to identify the contents of a module in preparation for using it or seeking to gain an understanding of the tutorial structure. Other researchers (Akscyn et al, 1988) have identified the link between browsing and a user's desire to build up a picture of the material's structure, which will allow more informed choices to be made. This is analogous to the way that people dip into a book to decide if it is suitable for them. The mixed mode is therefore more likely to be used by learners who are new to a subject. The observations confirm this analysis by showing that 50% of beginners and 40% of intermediates adopted this approach compared to 29% of experts.

The key finding is that systems which rely on browsing for users to discover the contents of modules are likely to be most effective with learners who already have extensive knowledge of the subject.

5.6 Conclusions

Although there is a body of research related to screen design, it is not evenly spread between the three main elements of text, colour and graphics. Only text has a well-balanced body of evidence. Many of the sources of evidence related to colour and graphics come from experiments on media other than a computer screen, with no focus on learning. There are therefore doubts as to whether they can be transferred to the screen design of computer-based learning materials.

The Colour experiment produced a variety of results which suggest that

colour motivates users to spend more time studying a display. However, when too many colours are employed (more than seven colours) they begin to distract users who respond by accessing fewer frames and terminate the tutorial as soon as possible.

All users indicated a positive attitude to the design of the Graphics tutorial and successfully demonstrated their comprehension of the material through their test results. The comprehension test was limited but did provide a means of assessing the learners' understanding of the tutorial. This is positive evidence that the underlying design of the material is effective and supports the standard screen design guidelines employed. This is especially important in the use of colour which followed the results of the Colour experiment. The Graphics experimental results provide additional confirmation of the colour screen guidelines.

The review of computer-based learning materials provided a baseline of current practice in the use of colour, text and graphics. The Graphics experiment made more extensive use of graphics than current practice. The users' attitudes show that they supported this increase and would marginally welcome even more graphics. Current design practice and user preferences are therefore not in tune.

The users' preference is for representational graphics occupying a quarter to half of the screen. Equally, learners do not like analogical graphics and they do not appear to use them to make links with their prior experience. At best, the analogical images are considered by the learners as representational images which may well hinder rather than assist learning. These results are in line with current practice - as shown by the review of learning materials.

Complexity was found to be a useful means of predicting and comparing the users' behaviour when interacting with different screen displays. Overall, the complexity measures accurately predicted the time spent on the tutorial; however, there were considerable differences between the three groups. The intermediate group spent the greatest time, followed by the beginner group, with the expert group spending least. This again illustrates the importance of the users' previous understanding of the tutorial subject.

The behaviour of the three groups frequently followed the pattern of the users' prior knowledge, that is, the behaviour of the experts was frequently different

to the other two groups of users. In particular, it differed with respect to reaction to graphic displays, choices of routes through the material, accessing the additional modules, and in the operation of the tutorial. This is critical for a designer who must have a clear understanding of the nature of learners. A design which assumes that learners will adopt a particular mode of use may well be unsuitable if its users are predominantly experienced or novice learners.

All types of graphics were studied for less time than complexity predictions would expect. However, representative images motivated learners to study them more, unless too many illustrations were employed. Equally, logical and analogical graphics did not motivate learners to make the necessary effort to extract the information they contained. In contrast, structured text, which could be considered the equivalent of a logical graphic, was clearly used by learners.

The route chosen by the users was definitely sequential, with the best prediction of the learner's route being the order in which modules appeared in the main menu. Only the expert group revealed more complex behaviour and even this group initially followed a sequential pattern. The experts, however, did operate in a browsing mode to a far greater extent than the other two groups, who were more likely to complete a module before moving on.

Accessing the additional modules is again clearly linked to a learner's prior knowledge. The Dictionary and Quiz were extensively used compared with the other modules. Modules of this type frequently form part of learning materials and software and thus computer literate users may well have already encountered them. The unfamiliar modules need the use of devices such as cross-references to encourage their use. Even the use of cross-references will not guarantee that the majority of novice users will access the module (only 30% of the beginner group accessed the Examples module).

Chapter Six

CONCLUSIONS

6.0 Introduction

This chapter will consider the conclusions of the research presented earlier in the thesis. It will compare the hypotheses presented in chapter one with the evidence produced by the review of the research literature as well as the Colour and Graphics experiments.

6.1 Use of Colour

Three hypotheses were investigated in the Colour experiment. They were A1, A2 and A3 (section 1.10).

- (A1) The addition of colour will increase the time each learner will spend on the individual frames of the tutorial.

The results of the experiment show that the use of colour increases the mean time spent on each frame. This effect was apparent for all types of frames, that is: menu, question and tutorial frames.

- (A2) The addition of colour will increase the number of frames which a learner will explore during the tutorial.

Colour clearly motivates learners to study the tutorial. However, the excessive use of colour (more than 7 colours) will deter learners from exploring the material.

- (A3) The addition of colour will increase the overall time each learner will spend on the tutorial.

The overall time to complete the tutorial showed an interesting trend, in that it increased from the base level in the monochrome version up to and including version 3 (7 colours). In the last two versions (9 colours and 11 colours, respectively) the time taken fell below the time spent on the monochrome version. The users spent a greater time on each individual frame as the number of colours used increased but accessed fewer frames in the later versions using 9 and 11 colours, respectively.

Colour has two different effects on the behaviour of learners. Initially, it motivates and encourages them to spend longer on each display. It engages them to interact with the information displayed. However, as the use of colour increases it begins to interfere and distract learners so that although they are spending longer on each screen, it is for negative reasons. Learners find it more difficult to understand the information displayed and have to make an increased effort to do so. This is illustrated by learners exploring less of the material and by their responses to the post-experiment attitude questions relating to colour. The turnover point for the use of colour has been shown to be seven colours. It is likely that other factors not considered in this research may also influence this, such as colour preference, colour combinations and tutorial subject.

6.2 Complexity

The following hypothesis was investigated in the Graphics experiment:

- (B1) the time spent on each frame will increase as the complexity of the display increases.

The results of the experiment show there is an overall positive correlation between the time spent on a frame and the measure of complexity. There are however differences between the three groups of learners, that is, beginner, intermediate and expert.

Beginners tended to spend less time on a frame than the complexity measure would predict. Complexity was found to be an accurate predictor of the time the intermediate group would spend on a frame while experts tended to spend less time on a frame than complexity would predict.

6.3 Learners

Two hypotheses were investigated in the Graphics experiment relating to a learner's prior knowledge of the tutorial subject:

(B2) the time spent on each frame will vary according to a learner's prior knowledge of the tutorial subject; and

(B3) the route chosen by a learner through the tutorial will reflect the differences in that learner's prior knowledge of the tutorial subject.

The evidence supports the view that the nature of learners affects their behaviour with particular screen designs. In this case, the main difference between the groups was their prior knowledge of the subject being studied.

Route

All three groups of learners (beginner, intermediate and expert) showed a clear preference to follow a sequential route through the tutorial. However, the expert group demonstrated more complex behaviour than the other two.

The behaviour of the three groups seems to be governed by the following guidelines:

(1) learners will tend to follow the sequence of the tutorial as shown by the main menu;

- (2) learners with expert knowledge of the subject will only initially follow the sequence of the tutorial; and
- (3) the best predictor of the numbers of learners who will access a given module is the order the module appears on the main menu (more people will use module 1 than module 2).

Learners' behaviour in using the tutorial was classified into three types: browsing, completing the majority of a module before moving to new module, and a mixture of browsing and completing modules. There was a considerable difference between the three groups of users. The experts tended to browse the material to a far greater degree than the other two groups of learners. In contrast, the beginner and intermediate groups tended to complete a module before moving on to the next part of the tutorial.

The overall conclusion is that the learners' prior knowledge does influence their behaviour in using computer-based learning material.

6.4 Use of Graphics

The Graphics experiment considered the following hypothesis relating to the use of graphic illustrations in the tutorials:

- (B4) the time spent on each frame will relate to the type and size of graphic being displayed.

Preferences

The three groups of learners expressed preferences for both the size and type of illustration used in learning material. However, the strength of these preferences does remain in doubt. The beginner group appeared to be indifferent to the size of illustration, the experts ranked 'Do not care' as their third choice and only the intermediates seemed to have a firm preference for particular sizes of illustration.

These preferences were for illustrations which occupy a quarter or half of the screen.

The three groups of learners clearly showed preferences for different types of illustration. All three groups did not like the use of analogical illustrations. In direct contrast to this negative reaction was a preference for realistic images which was expressed by all three groups. The overall preferences, in order, were realistic, cartoon, line drawings, diagrams, charts and tables and analogies.

Graphic Illustrations

A learner's behaviour towards different types of graphic display certainly varied. The comparison based on the time actually spent compared to a prediction based on the complexity of the display shows that, overall, learners spent less time on displays containing graphic illustrations than would be expected. This could be explained in two main ways. Firstly, learners may find it easier to extract the information contained in an image rather than a text-only display. Secondly, learners may not be willing to invest the effort in extracting all the information contained in an illustration.

It is likely that both answers are correct, depending on the circumstances, such as the nature of the display and the type of illustration employed.

Learners generally spent less time than predicted on graphic images of all types with three exceptions:

- (1) beginners spent more time on full-screen representative images (although not statistically significant);
- (2) all learners spent more time than predicted on displays containing two part-screen representative images; and
- (3) all learners spent more time than predicted on displays containing three part-screen representative images.

The exceptions were all linked to representative displays which, combined with the learners' preference for this type of graphic image, would suggest that these are the most effective type of graphic to use in learning material.

Learners interacting with logical graphics spent significantly less time than predicted. This would indicate that learners were unwilling to make the effort to understand the structure of the graphic in order to obtain the information it contains. This is in contrast to the learners' behaviour with displays of structured text. Structured text is the direct equivalent of a logical graphic. Learners spent more time on structured text than predicted, indicating they were willing to make the required effort to understand the structure or that the structure was a lesser barrier.

Learners appear to treat analogical images (at least part-screen analogical graphics) in a similar way to representative images and are clearly not making the potential powerful links to other experiences which would aid their learning. They may be hindering their learning by treating analogies as realistic images.

The evidence supports the hypothesis (B4) in that the time spent on each frame will relate to the type and size of graphic being displayed.

6.5 Additional Modules

The Graphics experiment considered the following hypothesis related to the learners' access to the additional modules:

(B5) the additional modules which are cross-referenced in the tutorial will be accessed more frequently than the other additional modules.

There were five additional modules within the tutorial. They were Examples, More Information, Dictionary, Quiz and Exit. The modules can be compared in varying degrees, depending on the type of module, how they are accessed and the learners' prior experience of their function.

The Examples and More Information modules can be compared in order to assess the effect of using cross-references. Both modules were accessed in an identical manner. However, the text of the tutorial contained 13 cross-references to

the Examples module while the More Information module was not cross-referenced. The Examples module was accessed substantially more than the More Information module. In particular, almost four times more beginners used the Examples module than used the More Information module. This evidence supports the hypothesis.

Another comparison can be made between the Dictionary and Quiz modules. Neither of the modules was cross-referenced and both were accessed in an identical manner. Both these additional modules were used to a far greater extent than the Examples and More Information modules. This difference in access probably reflects the users' previous knowledge of the two types of module. The Dictionary module is a familiar part of computer software while a Quiz is often associated with learning activities. Learners appear to be more influenced by prior knowledge than by the use of cross-references to use additional modules.

6.6 Research Literature

The review of the research literature indicates that a large proportion of the work is not based on a computer screen, does not have a learning focus and does not use integrated displays. This casts doubt on whether it can be transferred to the design of computer-based learning displays. However, the guidelines developed from text research were shown to be effective in both experiments.

6.7 Next Steps

This research has shown the power of screen design to influence a learner's behaviour. It has also illustrated the need to continue to investigate screen design issues. One area indicated for further study is the investigation of the effects of higher quality images, since this research was limited to VGA quality graphics using sixteen colours. It did not consider higher resolution images. Although the size and number of graphic images were considered, the position of the image on a screen was not investigated. The results indicated that analogical graphics did not aid learners to make links with their prior knowledge. However, this research was limited to a single context, that is computer studies. The application of analogical images should

be appraised in other contexts, since the research literature does indicate that they can serve this purpose (make links with their prior learning). In a similar way, learners do not study logical graphics adequately to extract their information while structural text is studied sufficiently. This contrast should be considered in more depth. Finally, this research has shown the value and limits of a text cross-reference within the tutorial. This is basically applying a device used extensively in books. A computer offers a range of other cross-referencing devices such as sound, which may be more effective. These devices are worth investigating.

REFERENCES

- Alesandrini, K.L (1984), 'Pictures and adult learning', *Instructional Science*, 13, 63-77.
- Alesandrini, K.L (1985), 'Instructional graphics in commercial courseware', in proceedings of 25th Association for the Development of Computer-Based Instructional Systems Conference, Columbus, Ohio, 188-195.
- Alesandrini, K.L (1987), 'Computer graphics in learning and instruction ', in H.A Houghton and D.M Willows (editors), *The Psychology of Illustration*, Springer-Verlag, New York, 2, 159-188.
- Alesandrini, K.L and Rigney, J.W (1981), 'Pictorial practice and review strategies in science learning', *Journal of Research in Science Teaching*, 5, 465-474.
- Allen, W.H (1975), 'Intellectual abilities and instructional media design', *Audio Visual Communication Review*, 23(2), 139-170.
- Akscyn, R, McCracken, D.L, Yoder, E.A (1988), 'KMS, a distributed hypermedia system for managing knowledge in organisations', *Communication of the ACM*, 31(7), 820-835.
- Andrews, R.B (1949), 'Reading power unlimited', *Texas Outlook*, 33, 208-209.
- Anglin, G.J (1986), 'Prose relevant pictures and older learners recall of written prose', *Educational Communication and Technology Journal*, 34, 131-136.
- Anglin, G.J (1987), 'Effects of pictures on recall of written prose: How durable are picture effects?', *Educational Communication and Technology Journal*, 35(1), 25-30.
- Ausubel, D.P (1968), 'Educational psychology: A cognitive view', Holt, Rinehart and Winston, New York.
- Avons, S.E, Beveridge, M.C, Hickman, A.T and Hitch, G.J (1983), 'Teaching journey graphs with microcomputer animation', *Human Learning*, 2, 93-105.
- Ayerst, J.M (1987), 'The use of authoring languages for computer-based learning (CBL)', *Resources for Information Technology in Further Education*, Further Education Unit, London, page 46.
- Backs, R.W, Walrath, L.C and Hancock, G.A (1987), 'Comparison of horizontal and vertical menu formats', in proceedings of the Human Factors Society, 31st Annual Meeting, Santa Monica, California.

- Baek, Y.K and Layne, B.H (1988), 'Color, graphics, and animation in a computer-assisted learning tutorial lesson', *Journal of Computer-Based Instruction*, 15(4), 131-135.
- Barker, P (1987), 'Author Languages for CAL', Macmillan Education Ltd, London, page 97.
- Barfield, L (1993), 'The user interface: Concepts and design', Addison-Wesley, Wokingham.
- Bauer, D and Cavonius, C.R (1980), 'Improving legibility of visual display units through contrast reversal', in E Grandjean and E Vigliani (editors), *Ergonomic Aspects of Visual Display Terminals*, Taylor and Francis, London, 137-142.
- Beevers, C.E, Cheery, B.S.G, Foster, M.M.G and McGuire, G.R (1991), 'Software tools for computer-aided learning in mathematics', Avebury Technical, Aldershot, 12-13.
- Benest, I.D and Dukic, D (1989), 'High-level user-interface objects', in G Salvendy and M J Smith (editors), *Designing and Using Human-Computer Interfaces and Knowledge Based Systems*, Elsevier Science Publishers B.V, Amsterdam, 597-604.
- Bernard, R.M (1990), 'Using extended captions to improve learning from instructional text', *British Journal of Educational Technology*, 21(3), 215-225.
- Bolton, J.W and Peck, K.L (1991), 'Menu sequence, learner control and achievement in CAI', *Interactive Learning International*, 7, 95-100.
- Bonsiepe, G (1968), 'A method of quantifying order in typographic design', *Journal of Typographic Research*, 2, 203-220.
- Booher, H.R (1975), 'Relative comprehensibility of pictorial information and printed words in proceduralised instructions', *Human Factors*, 17(3), 266-277.
- Borg, W.R and Schuller, C.F (1979), 'Detail and background in audio visual lessons and their effect on learners', *Educational Communication and Technology Journal*, 27, 31-38.
- Bork, A (1982), 'Information display and screen design', paper presented at an American Education Research Association Conference, New York.
- Bork, A (1983), 'A preliminary taxonomy of ways of displaying text on screens', *Information Design Journal*, 3, 206-214.
- Bork, A (1987), 'Learning with personal computers', Harper and Row, New York.

Borland International (1988), 'Turbo Pascal 4.0', Borland International, 1800 Green Hills Road, PO Box 660001, Scotts Valley, California.

Borland International (1989), 'Turbo Pascal 5.5', Borland International, 1800 Green Hills Road, PO Box 660001, Scotts Valley, California.

Borland International (1991), 'Turbo Pascal 6.0', Borland International, 1800 Green Hills Road, PO Box 660001, Scotts Valley, California.

Bouma, H (1970), 'Interaction effects in parafoveal letter recognition', *Nature*, 226, 177-178.

Brooke, J.B and Duncan, K.D (1981), 'Effects of system display format on performance in a fault location task', *Ergonomics*, 24(3), 175-189.

Brown, C.M (1986), 'Human computer interface design guidelines', Ablex Publishing Corporation, Norwood, New Jersey.

Bruce, M and Foster, J.J (1982), 'Looking for entries in videotext tables: A comparison of four colour formats', *Journal of Applied Psychology*, 67(5), 611-615.

Caldwell, R.M (1980), 'Guidelines for developing basic skills instructional materials for use with microcomputer technology', *Educational Technology*, 21(10), 7-12.

Cahill, M.C and Carter, R.C (1976), 'Colour code size of search in displays of different density', *Human Factors*, 18(3), 273-280.

Campbell, D.T and Stanley, J.C (1963), 'Experimental and quasi-experimental designs for research', in N.L Gage (editor), *Handbook of Research on Teaching*, Rand McNally, Chicago.

Campbell, A.J, Marchetti, F.H and Mewhort, D.J.K (1981), 'Reading speed and text production: A note on right justification technique', *Ergonomics*, 24(8), 633-640.

Card, S.K (1982), 'User perception mechanisms in the search of computer command menus', in proceedings: *Human Factors in Computer Systems*, Gaithersburg, March, 190-196.

Carver, R.P (1970), 'Effect of chunked typography on reading rate and comprehension', *Journal of Applied Psychology*, 54, 288-296.

Chen, H.C and Tsoi, K.C (1988), 'Factors affecting the readability of moving text on a computer screen', *Human Factors*, 30, 25-33.

Chen, H.C, Chan, K.T and Tsoi, K C (1988), 'Reading self paced moving text on a computer display', *Human Factors*, 30(3), 285-291.

Christ, R.E (1975), 'Review and analysis of color research for visual display', *Human Factors*, 17(6), 542-570.

Chute, A.G (1979), 'Analysis of the instructional functions of color and monochrome cueing in media presentations', *Educational Communication and Technology Journal*, 27(4), 251-263.

Chute, A.G (1980), 'Effect of colour and monochrome versions of a film on incidental and task relevant learning', *Educational and Communications and Technology Journal*, 28(1), 10-18.

Clark, R.E (1983), 'Reconsidering research on learning from media', *Review of Educational Research*, 53(4), 445-459.

Clark, R.E (1985), 'Evidence for confounding in computer-based instructional studies: Analysing the meta-analyses', *Educational Communication and Technology Journal*, 33(4), 249-262.

Clarke, A (1989), 'The principles of screen design for computer-based learning materials', Employment Department, Moorfoot, Sheffield.

Clarke, A (1991), 'Authoring', in proceedings of Open Learning Conference, Budapest Technical University, Employment Department, Moorfoot, Sheffield, 43-52.

Clarke, A (1992), 'The principles of screen design for computer-based learning materials', 2nd edition, Employment Department, Moorfoot, Sheffield.

Condor Group (1993), 'The potential for Open and Flexible Learning and Technology-Based Training', unpublished report for Learning Methods Branch, Employment Department, Moorfoot, Sheffield.

Cushman, W.H (1984), 'Reading from microfiche, from VDT and from the printed page; subject fatigue and performance: A preliminary report', in E Grandjean (editor), *Ergonomics and Health in Modern Offices*, Taylor and Francis, London, 322-328.

Danchak, M.M (1976), 'CRT displays for power plants', *Instrumentation Technology*, 23(10), 29-36.

Davidson, R.E (1976), 'The role of metaphor and analogy in learning', in J.R Levin and V.L Allen (editors), *Cognitive Learning in Children: Theories and strategies*, Academic Press, New York, 135-162.

Dean, C (1993), 'An overview of authoring systems for CBT and interactive multimedia that are available in the UK', Dean Associates, Sheffield.

Dean, C and Whitlock, Q (1988), 'A handbook of computer-based training', Kogan Page, London, 98-107.

Dean, R.S and Enernok, P.A (1983), 'Pictorial organisation in prose learning', *Contemporary Educational Psychology*, 8, 20-27.

Dean, R.S and Kulhavy, R W (1981), 'Influence of spatial organisation in prose learning', *Journal of Educational Psychology*, 76, 57-61.

DeMars, S.A (1975), 'Human Factors considerations for the use of colour in display systems', NASA, John F Kennedy Space Centre, Florida.

Department of Education and Science (1991), 'Survey of information technology in schools', *Statistical Bulletin*, 11/91, Analytical Services Branch, Room 338, Mowden Hall, Staindrop Road, Darlington, DL3 9BG.

Department for Education (1993), 'DfE bulletin 6/93', Analytical Services Branch, Room 338, Mowden Hall, Staindrop Road, Darlington, DL3 9BG.

Descy, D.E (1981), 'Color in media', *International Journal of Instructional Media*, 8(3), 261-266.

Dodson, D.W and Shields, N.J (1978), 'Development of user guidelines for ECAS display design', Report No. NASA-CR-150877, Essex Corporation, Huntsville, Alabama, Volume 1.

Donahue, T.R (1973), 'Viewer perceptions of colour and black and white paid political advertising', *Journalism Quarterly*, 50(4), 660-665.

Duchastel, P.C (1980), 'Research on illustration in text: Issues and perspectives', *Educational Communication and Technology Journal*, 28, 283-287.

Duchastel, P.C (1988), 'Display and interaction features of instructional texts and computers', *British Journal of Educational Technology*, 19(1), 58-65.

Durrett, H.J and Stimmel, D.T (1987), 'Color and the instructional use of the computer', in H.J Durrett (editor), *Color and the Computer*, Academic Press, Boston, 241-254.

Dwyer, F.M (1970), 'Exploratory studies in the effectiveness of visual illustrations', *Audio Visual Communication Review*, 18, 235-249.

Dwyer, F.M (1978), 'Strategies for improving visual learning: a handbook for effective selection, design and use of visualised materials', State College, Pennsylvania State University, Learning Services Division, Pennsylvania.

- Dwyer, F.M and Lamberski, J.R (1983), 'A review of research on the effects of the use of color in the teaching-learning process', *International Journal of Instructional Media*, 10(4), 303-327.
- Dwyer, F.M and Joseph, J.H (1984), 'The effects of prior knowledge, presentation mode and visual realism on student achievement', *Journal of Experimental Education*, 52, 101-121.
- Eaker, C.W and Jacobs, E.L (1982), 'Computer animation of a chemical reaction', *Journal of Chemical Education*, 59(10), 939-941.
- Egeth, H, Jonides, J and Wall, S (1972), 'Parallel processing of multielement displays', *Cognitive Psychology*, 3, 674-698.
- Elliot, P.H.(1975), 'An exploratory study of adult learning styles', Doctoral Dissertation, University of Illinois at Urbana-Champaign, University Microfilms, DAH 76-06757
- Employment and Education Departments and the Welsh Office (1991), 'Education and Training for the 21st Century', HMSO, London.
- Erdelyi, M.H and Stein, J.B (1981), 'Recognition hypermnesia: The growth of recognition memory over time with repeated testings', *Cognition*, 9, 21-33.
- Erickson, T.D (1991), 'Creativity and design', in B Laurel (editor), *The Art of Human-Computer Interface Design*, Addison-Wesley, Reading, Massachusetts, 1-4.
- Faiola, A and DeBloois, M.L (1988), 'Designing a visual factors based interface', *Educational Technology*, 29(8), 12-21.
- Fisher, D.L and Tan, K.C (1989), 'Visual displays: The highlighting paradox', *Human Factors*, 31(1), 17-30.
- Flynn, M.L (1989), 'The potential of older adults response to computer-assisted instruction', *Journal of Educational Technology Systems*, 17, 231-241.
- Force 9 (1993), 'Open for learning marketing proposals', unpublished report for Learning Methods Branch, Employment Department, Moorfoot, Sheffield.
- Frase, L.T and Schwartz, B.J (1979), 'Typography cues that facilitate comprehension', *Journal of Educational Psychology*, 71, 197-206.
- Freedman, J and Haber, R.N (1974), 'One reason why we rarely forget a face', *Bulletin of Psychonomic Society*, 3, 107-109.
- Gagne, R.M (1985), 'The conditions of learning and theory of instruction ', 4th edition, Holt, Rinehart and Winston, New York.

- Galitz, W.O (1985), 'Handbook of screen format design', 2nd Edition, QED Information Services, Wellesley, Massachusetts.
- Galitz, W.O (1993), 'User-Interface Screen Design', QED Publishing Group, Wellesley, Massachusetts.
- Galitz, W.O (1994), 'Its time to clean your windows: Designing GUIs that work', John Wiley and Sons, New York.
- Gamberg, B.F (1985), 'Strategic and attentional factors of computer generated highlighting', unpublished M.A thesis, New Mexico State University, Las Cruces, New Mexico.
- Gardiner, M.M and Christie, B (1987), 'Applying cognitive psychology to user-interface design', John Wiley and Sons, Chichester.
- Gery, G (1987), 'Making CBT happen, Weingarten Publications, Boston, Massachusetts, 93-130.
- Gorman, D.A (1973), 'Effects of varying pictorial detail and presentation strategy on concept formation', *Audio Visual Communication Review*, 21, 337-350.
- Gould, J.D and Grischkowsky, N (1982), 'Doing the same work with CRT terminal and with hardcopy', in proceedings of Human Factors Society, 26th Annual Meeting, Santa Monica, California, 165-166.
- Gould, J.D and Grischkowsky, N (1984), 'Doing the same work on hard copy and with the CRT computer terminal', *Human Factors*, 26, 323-337.
- Gould, J.D, Alfaro, L, Barnes, V, Finn, R, Grischkowsky, N and Minuto, A (1987), 'Reading is slower from CRT displays than from paper: Attempts to isolate a single variable explanation', *Human Factors*, 29(3), 269-299.
- Grabinger, R.S and Albers, S (1988), 'The effect of CRT Screen Design on Learning', in proceedings of Convention of Association for Educational Communications and Technology, New Orleans, 283-300.
- Graf, R and Torrey, J.W (1966), 'Perception of clause structure in written language', in proceedings of the American Psychology Association, 74th Annual Meeting, 83-84.
- Granaas, M.M, McKay, T.D, LaHam, R.D, Hurt, L.D and Juola, J.F (1984), 'Reading moving text on a CRT screen', *Human Factors*, 26, 97-104.
- Greene, J and D'Oliveira, M (1982), 'Cognitive psychology methodology handbook', Part 1 and 2, Open University, D303 units 5 to 12, Milton Keynes.

- Gregory, M and Poulton, E.C (1970), 'Even versus uneven right-hand margins and the rate of comprehension in reading', *Ergonomics*, 13, 427-434
- Guastello, S.J, Traut, M and Korienek, G (1989), 'Verbal versus pictorial representations of objects in a human-computer interface', *International Journal of Man-Machine Studies*, 31, 99-120.
- Haber, R.N and Wilkinson, L (1982), 'Perceptual components of computer displays', *IEEE Computer Graphics and Applications*, 2, 23-35.
- Hamilton, A, Siverside, J, Torrance, A and Hann, H (1994), 'Designing CBT - Focusing on the issues', in proceedings of Association for Educational and Training Technology, 28th Conference, 11-13 April, Edinburgh, 132-138.
- Hannafin, M.J (1985), 'Empirical issues in the study of computer assisted interactive video', *Educational Communication and Technology Journal*, 33(4), 235-247.
- Haring, M.J and Fry, M.A (1979), 'Effect of pictures on children's comprehension of written text', *Educational Communication and Technology Journal*, 27, 185-190.
- Harrison, N (1991), 'How to design effective computer-based training', McGraw-Hill, 193-224.
- Hartley, J (1980a), 'Space and structure in instructional text, in *The Psychology of Written Communications*, Kogan Page, London, 127-144.
- Hartley, J (1980b), 'Spatial clues in text', *Visible Language*, 14, 62-79.
- Hathaway, M.D (1984), 'Variables of computer screen display and how they affect learning', *Educational Technology*, January, 7-11.
- Hativa, N and Teper, A (1988), 'Differential effectiveness of three colour treatments in learning geometric concepts via computer guided teaching', *Journal of Educational Computing Research*, 4(3), 303-320.
- Haubner, P and Newmann, F (1986), 'Structuring alphanumerically coded information on visual display units', in proceedings of International Scientific Conference : Work With Display Units, Stockholm, Sweden, May 12-15, 606-609.
- Hawkridge, D, Newton, W and Hall, C (1988), 'Computers in company training', Croom Helm, London.
- Heines, J.M (1984), 'Screen design strategies for computer-assisted instruction', Digital Press, Bedford, Massachusetts.

- Henney, M (1981), 'The effects of all-capital print versus regular mixed print, as displayed on a microcomputer screen, on reading speed and accuracy', ERIC Document Reproduction Service No.208 359.
- Heppner, F.H, Anderson, J.G.T, Farstrup, A.E and Weideman, N.H (1985), 'Reading performance on a standardised test is better from print than from computer display', *Journal of Reading*, 28, 321-325.
- Holliday, W.G (1975), 'The effects of verbal and adjunct pictorial-verbal information in science instruction', *Journal of Research in Science Teaching*, 12, 77-83.
- Holliday, W.G (1976), 'Teaching verbal chains using flow diagrams and texts', *Audio Visual Communication Review*, 24, 63-78.
- Holliday, W.G and Thursby, P.M (1977), 'Teaching visual concepts in the classroom', paper presented to Annual Convention of the Association for Educational and Communication and Technology, April, Miami.
- Holliday, W.G, Brunner, L.L and Donais, E.L (1977), 'Differential cognitive and effective responses to flow diagrams in science', *Journal of Research in Science Teaching*, 14, 129-138.
- Isaacs, G (1987), 'Text screen design for computer-assisted learning', *British Journal of Educational Technology*, 18(1), 41-61.
- Jackson, R, MacDonald, L and Freeman, K (1994), 'Computer generated color', *Wiley Professional Computing*, Chicester, 126-129.
- Kak, A.V (1981), 'Relationships between readability of printed and CRT displayed text', in proceedings of the Human Factors Society, 25th Annual Meeting, Santa Monica, California, 137-140.
- Kanner, J.H (1968), 'The instructional effectiveness of color in television: A Review of the evidence', Department of the Army, Office of Assistant Chief of Staff for Communication-Electronics, ERIC Document Reproduction Service ED 015675.
- Kanner, J.H and Rosenstein, A.J (1961), 'Television and army training: Color versus black and white', *Audio Visual Communication Review*, 9(1), 44-49.
- Katzman, N and Nyenhius, J (1972), 'Color versus black and white effects on learning, opinion, and attention', *Audio Visual Communication Review*, 20(1), 16-28.
- Keane, D.R, Norman, G.R, and Vickers, J (1991), 'The inadequacy of recent research on computer assisted instruction', *Academic Medicine*, 66(8), 444-448.
- Kearsley, G (1986), 'Authoring - A guide to the design of instructional software', Addison-Wesley, Reading, Massachusetts, 24-48.

Kearsley, G, Hunter, B and Seidel, R.J (1983), 'Two decades of computer-based instruction projects: What have we learned?', *Technological Horizons in Education*, 12(3), 90-94.

Kelley, J.P (1988), 'Color and highlighting effectiveness in text-based CBT', in proceedings of 6th Annual Computer-Based Training Conference, Weingarten Publications, Boston, Massachusetts, Session 506.

Klare, G.R, Nichols, W.H and Shuford, E.H (1957), 'The relationship of typographic arrangements to the learning of technical training material', *Journal of Applied Psychology*, 41, 41-45.

Knowlton, J.Q (1966), 'On the definition of picture', *Audio Visual Communication Review*, 14, 157-183.

Kolers, P.A, Duchnicky, R.L and Ferguson, D.C (1981), 'Eye movement measurement of readability of CRT Displays', *Human Factors*, 23, 517-527.

Koran, M.L and Koran, J.J (1980), 'Interaction of learner characteristics with picture adjuncts in learning from science text', *Journal of Research into Science Teaching*, 17, 477-483.

Kruk, R S and Muter, P (1984), 'Reading of continuous text on video screens', *Human Factors*, 26(3), 339-345.

Kulhavy, R.W, Schwartz, N.H and Shaha, S.H (1983), 'Spatial representation of maps', *American Journal of Psychology*, 96, 337-351.

Labour Force Survey (1988), Employment Department, Caxton House, London.

Lefrere, P (1984), 'Text processing', in T O'Shea and M Eisenstadt (editors), *Artificial Intelligence*, Harper and Row, New York, 400-422.

Leiblum, M.D (1992), 'Implementing CAL at a university', *Computers and Education*, 18(1 to 3), 109-118.

Levie, H and Lentz, R (1982), 'Effects of text illustrations: A review of research', *Educational Communication and Technology Journal*, 30, 195-232.

Levin, J.R (1989), 'A transfer-appropriate-processing perspective of pictures in prose', in H Mandl and J.R Levin (editors), *Knowledge Acquisition from Text and Prose*, Elsevier Science Publishers B V (North Holland), Amsterdam, 83.

Levin, J.R and Lesgold, A (1978), 'On pictures and prose', *Educational Communication and Technology Journal*, 26, 233-243.

Levin, J.R, Anglin, G and Carney, R (1987), 'On empirical validating function of pictures and prose', in D.M Willows and H.A Houghton (editors), *The Psychology of Illustration, Basic Research*, Springer-Verlag, New York, 1, 51-85.

Lewis, R and Paine, N (1985), 'How to Communicate with the Learner', Council for Educational Technology, London, 69-76.

Lloyd, D.C and Gressard, C (1984), 'The effects of sex, age, and computer experience on computer attitudes', *AEDS Journal*, 8(2), 67-77.

Long, T (1984), 'Human factors principles for the design of computer colour graphics displays', *British Telecom Technology Journal*, 2(3), 5-14.

Lucas, L (1991), 'Visually designing the computer learner interface', *Educational Technology*, July, 56-58.

Luder, C B (1984), 'Redundant colour coding on airbourne CRT display', *Human Factors*, 26(1), 19-52.

Macdonald, W.A and Cole, B.L (1988), 'Evaluating the role of colour in a flight information cpckpit display', *Ergonomics*, 31(1), 13-37.

Mahmood, M.A and Medewitz, J.N (1989), 'Assessing the effect of computer literacy on subjects attitudes, values and opinions towards information technology: an exploratory longitudinal investigation using the linear structure relations (LISREL) model', *Journal of Computer-Based Instruction*, 16(1), 20-28.

Marschalek, D.G (1988), 'Processing and memory of color, contour, and pattern found in computer digitized color pictures for elemantary children', *Journal of Educational Computing Research*, 4(4), 403-412.

Marshall, I.M, Samson, W.B and Dugard, P (1994), 'Predicting the development effort and cost of multimedia courseware', paper presented to Association for Educational and Training Technology, 28th Conference, 11-13 April, Edinburgh.

Mayer, R.E (1975), 'Different problem solving strategies established in learning computer programming with and without useful models', *Journal of Educational Psychology*, 67, 725-734.

Mayer, R.E (1980), 'Elaboration techniques that increase the meaningfulness of technical text: An experimental test of the learning strategy hypothesis', *Journal of Educational Psychology*, 72, 770-784.

Mayer, R.E (1982), 'Learning, *Encyclopedia of educational research*', Free Press, New York.

- Mayer, R.E and Anderson, R.B (1991), 'Animation need narrations: an experimental test of a dual coding hypothesis', *Journal of Educational Psychology*, 83(4), 484-490.
- McAleese, R (1989), 'Navigation and browsing in hypertext', in R McAleese (editor), *Hypertext: theory into practice*, Blackwell Scientific Publications, Oxford, 6-44.
- Miller, G (1956), 'The magical number seven, plus or minus two', *Psychological Review*, 63, 81-97.
- Mills, C.B and Weldon, L.J (1984), 'Readings from computer screens', Technical report CAR-TR-94, College Park, MD, University of Maryland, Human-Computer Interaction Laboratory.
- Mitchell, J and Shneiderman, B (1989), 'Dynamic versus static menu: An exploratory comparison', *ACM SIGCHI Bulletin*, 20(4), 33-37.
- Moore, P.J and Skinner, M.J (1985), 'The effects of illustrations on childrens' comprehension of abstract and concrete passages', *Journal of Research in Reading*, 8, 45-56.
- Moran, T (1981), 'An applied psychology of the user', *ACM Computing Surveys*, 13(1), 1-11.
- Morris, D.C (1989), 'A survey of age and attitudes towards computers', *Journal of Educational Technology Systems*, 17, 73-78.
- Moyer, J.C, Sowder, L, Threadgold-Sowder, J and Moyer, M.B (1984), 'Story problem formats: Drawn versus verbal versus telgraphic', *Journal of Research in Mathematics Education*, 15, 342-351.
- Muncer, S J, Gorman, B S, Gorman, S and Bibel, D (1986), 'Right is wrong: an examination of the effect of right justification on reading', *British Journal of Educational Technology*, 17(1), 5-10.
- Murray, K (1992), 'The Graphics Coach', New Riders Publication, Carmel, Indiana, 39-46.
- Muter, P, Latremouille, S.A, Treurniet, W.C and Beam, P (1982), 'Extended readings of continuous text on television screens', *Human Factors*, 24, 501-508.
- Myatt, B and Carter, J.M (1979), 'Picture preferences of children and young adults', *Education Communication and Technology Journal*, 27, 45-53.
- Nilsson, L.G, Ronnberg, J and Ohlsson, K (1983), 'Rating of colour combinations of text and background presented by means of a videotext system', *Umea Psychological Reports*, No.166, University of Umea.

NOP Social and Political (1992), 'Employer flexible access', unpublished report for Learning Technology Unit, Employment Department, Moorfoot, Sheffield.

North, A.J and Jenkins, L.B (1951), 'Reading speed and comprehension as a function of typography', *Journal of Applied Psychology*, 35, 225-228.

O'Connor, B.C (1985), 'Access to moving image documents: Background concepts and proposals for surrogates for film and video works', *Journal of Documentation*, 41, 209-220.

Oleron, G and Tardieu, H (1978), 'Influence of scrolling on the recall of text', in M.M Gruneberg, P.E Morris and R.N Sykes (editors), *Practical Aspects of Memory*, Academic Press, London.

Paivio, A (1971), 'Imagery and verbal processes', Holt, Rinehart and Winston, New York.

Paivio, A (1975), 'Imagery and long term memory', in A Kennedy and A Wilkes (editors), *Studies in Long Term Memory*, John Wiley and Sons, New York, 57-85.

Parkhurst, P.E and Dwyer, F.M (1983), 'An experimental assessment of student IQ level and their ability to profit from visual instruction', *Journal of Instructional Psychology*, 10, 9-20.

Parkinson, S.R and Sisson, N (1985), 'Organisation of broad computer menu displays', *International Journal of Man-Machine Studies*, 23, 689-697.

Paterson, D.G and Tinker, M.A (1940), 'How to make type readable', Harper and Brothers, London.

Patrick, J and Fitzgibbon, L (1988), 'Structural displays as learning aids', *International Journal of Man-Machine Studies*, 28, 625-635.

Peeck, J (1987), 'The role of illustrations in processing and remembering illustrated text', in D.M Willows and H.A Houghton (editors), *The Psychology of Illustration*, Springer-Verlag, New York, 1, 115-151.

Pond, D.J (1982), 'Gender differences in response to ambient color, task complexity and presence of evaluative others', in proceedings of the Human Factors Society, 26th Annual Meeting, Santa Monica, California, 590-594.

Preece, J, Sharp, H, Benyon, D, Holland, S and Carey, T (1994), 'Human-Computer Interaction', Addison-Wesley, Wokingham, page 727.

Public Domain Software Library (1992a), 'Iconnew', The Public Domain Software Library, Winscombe House, Beacon Road, Crowborough, Sussex.

Public Domain Software Library (1992b), 'Raster Master 2.1', The Public Domain Software Library, Winscombe House, Beacon Road, Crowborough, Sussex.

Race, P and Brown, S (1994), 'Getting the wording right in computer-based learning materials', in proceedings of Association of Educational and Training Technology, 28th Conference, 11-13 April, Edinburgh, 75-77.

Reed, S.K (1985), 'Effect of computer graphics on improving estimates to algebra word problems', *Journal of Educational Psychology*, 77(3), 285-298.

Reid, D.J and Wicks, S (1988), 'Young children's perception of colour on a School Microcomputer', *Programmed Learning and Educational Technology*, 25(4), 384-353.

Reynolds, L (1979), 'Teletext and viewdata - a new challenge for the designer', *Information Design Journal*, 1, 2-14.

Richey, R.C (1991), 'Adult attitudes towards alternative delivery systems and industrial training outcomes', in M.R Simonson (editor), in proceedings of Selected Research Paper Presentations at the 1991 Annual Convention of the Association of Educational Communications and Technology, Ames, University of Iowa, 635-680.

Riding, R.J (1992), 'Performance on, and attitudes to, computer-based training within service industries', Learning Methods Project Report, No. OL 137, Employment Department, Moorfoot, Sheffield.

Rieber, L.P (1989), 'The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson', *Journal of Educational Computing Research*, 5(4), 431-444.

Rieber, L.P (1990), 'Using computer animated graphics in science instruction with children', *Journal of Educational Psychology*, 82(1), 135-140.

Rieber, L.P (1991), 'Animation incidental learning and continuing motivation', *Journal of Educational Psychology*, 83(3), 318-328.

Rieber, L.P and Hannafin, M.J (1988), 'Effects of textual and animated activities and practice on learning from computer-based instruction', *Computers in Schools*, 5(1), 77-89.

Rieber, L.P, Boyce, M.J and Assad, C (1990), 'The effects of computer animation on adult learning and retrieval tasks', *Journal of Computer-Based Instruction*, 17(2), 46-52.

Rigney, J.W and Lutz, K.A (1976), 'Effect of graphic analogies of concepts in chemistry on learning and attitude', *Journal of Educational Psychology*, 68, 305-311.

- Rivlin, C, Lewis, R, and Davies-Cooper, R (1990), 'Guidelines for Screen Design', Blackwell Scientific Publications, Oxford.
- Ross, S.M, Morrison, G.R and O'Dell, J.K (1988), 'Obtaining more out of less text in CBI: Effects of varied text density levels as a function of learner characteristics and control strategy', *Educational Communication and Technology Journal*, 36(3), 131-142.
- Rowntree, D (1992), 'Exploring Open and Distance Learning, Kogan Page, London, 175-200.
- Royer, J.M and Cable, G.W (1975), 'Facilitated learning in connected discourse', *Journal of Educational Psychology*, 67, 116-123.
- Royer, J.M and Cable, G.W (1976), 'Illustrations, analogies and facilitative transfer in prose learning', *Journal of Educational Psychology*, 68, 205-209.
- Rudnick, M.F, Porter, M.C and Suydam, E.L (1973), 'Pictorial Stimulus Variables', *Viewpoints*, 49(2), 21-28.
- Sasser, M.F and Moore, D.M (1984), 'A study of the relationship between learner control patterns and course completion in CAI', *Programmed Learning and Educational Technology*, 21(1), 28-33.
- Scanlon, J.T (1967), 'Colour television: New language?', *Journalism Quarterly*, 44(2), 225-230.
- Schwarz, E, Beldie, I P and Pastoor, S (1983), 'A comparison of paging and scrolling for changing screen contents by inexperienced users, *Human Factors*, 25(3), 279-282.
- Schuerman, R and Peck, K (1991a), 'Menus as lesson structuring devices: The effects of on-menu cueing on sequentiality', *Interactive Learning International*, 7(4), 293-304.
- Schuerman, R and Peck, K (1991b), 'Pull down menus, menu design and usage patterns in CAI', *Journal of Computer-Based Instruction*, 18(3), 93-98.
- Seddon, G.M and Shubber, K.E (1985), 'The effects of colour in teaching the visualisation of rotations in diagrams of three dimensional structures', *British Journal of Educational Technology*, 11(3), 227-239.
- Sewell, E.H and Moore, R.L (1980), 'Cartoon embellishment in informative presentations', *Educational Communication and Technology Journal*, 28(1), 39-46.
- Sherr, S (1993), 'Electronic Displays', 2nd Edition, John Wiley and Sons, New York.

- Simone, L (1992), 'The motion is the message', *PC Magazine*, 11(2), 435-467.
- Simpson, A (1994), 'Mastering Wordperfect 6.0 for Windows', Sybex, San Francisco, 925-927.
- Smith, S.L (1962), 'Colour coding and visual search', *Journal of Experimental Psychology*, 24(5), 434-440.
- Smith, S.L (1963), 'Colour coding and visual separability in information displays', *Journal of Applied Psychology*, 47, 358-364.
- Smith, S.L and Thomas, D.W (1964), 'Colour versus shape coding in information displays', *Journal of Applied Psychology*, 48(3), 137-146.
- Spannaus, T.W (1985), 'Screen design for CAI', in proceedings of 26th Conference Association for the Development of Computer-Based Instructional Systems, Philadelphia, 34-37.
- Spaulding, S (1955), 'Research on pictorial illustration', *Audio Visual Communications Review*, 3, 22-45.
- Standing, L (1973), 'Learning 10000 pictures', *Quarterly Journal of Experimental Psychology*, 25(2), 207-222.
- Standing, L, Conezio, J and Haber, R.N (1970), 'Perception and memory for pictures: Single trial Learning of 2500 visual stimuli', *Psychonomic Science*, 19, 73-74.
- Stefanko, D, Tonkin, D, Silver, M and Whitman, G (1990), 'Education technology - A corporate weapon in the 1990s market place', Unisys Corporation, Internal Education and Development, Box 110, Princeton, New Jersey.
- Stoddard, M.L (1985), 'Research on user interface design as related to computer-based instruction', in proceedings of 26th Association for the Development of Computer-Based Instructional Systems Conference, Philadelphia, 251-255.
- Surber, J.R and Leeder J.A (1988), 'The effect of graphic feedback on student motivation', Association for the Development of Computer-Based Instructional Systems, *Journal of Computer-Based Instruction*, 15(1), 14-17.
- Sutherland, N.S (1980), 'Prestel and the user', Central Office of Information, London.
- Sweeters, W.G (1985), 'Screen design guidelines', in proceedings of 26th Conference of Association for the Development of Computer-Based Instructional Systems, Philadelphia, 42-45.

Switchenko, D (1984), 'Reading from CRT versus paper: The CRT-disadvantage hypothesis re-examined', in proceedings of the Human factors Society 28th Annual Meeting, Santa Monica, California, 429.

Tan, K.C (1987), 'Highlighting parameters of computer generated display', unpublished M.A thesis, University of Massachusetts, Amherst, Massachusetts.

TechnoJock Software (1989), 'TechnoJock's Turbo Toolkit', TechnoJock Software, Inc, PO Box 820927, Houston, Texas.

TechnoJock Software (1991), 'TechnoJock's Object Toolkit', TechnoJock Software, Inc, PO Box 820927, Houston, Texas.

Thimbleby, H (1990), 'User interface design', ACM Press, New York.

Tillman, M.H and Glynn, S.M (1987), 'Writing text that teaches: Historic overview', Educational Technology, October, 41-45.

Tinker, M.A (1960), 'Legibility of mathematical tables', Journal of Applied Psychology, 44, 83-87.

Tinker, M.A (1963), 'Legibility of print', Iowa State University Press, Ames, Iowa.

Training Agency (1989), 'Training in Britain: a study of funding activity and attitudes', HMSO, London.

Training Agency (1990), 'Labour market and skills trends 1991/92', Employment Department, Moorfoot, Sheffield.

Travers, R.M and Alvarado, V (1970), 'The design of pictures for teaching children in elementary school', Audio Visual Communication Review, 18, 47-64.

Treisman, A (1982), 'Perception groupings and attention in visual search for features and for objects', Journal of Experimental Psychology: Human Perception and Performance, 8(2), 194-214.

Tritz, G.J (1986), 'Computer modeling of microbiological experiments in the teaching laboratory: animation techniques', Journal of Computers in Mathematics and Science Teaching, 6(2), 44-48.

Tullis, T.S (1981), 'An evaluation of alphanumeric, graphic and color information displays', Human Factors, 23, 541-550.

Tullis, T.S (1983), 'Predicting the usability of alphanumeric displays', PhD dissertation, The Report Store, Rice University, Lawrence, Kansas.

- VanderMeer, A.W (1954), 'Color versus black and white in instructional films', *Audio Visual Communication Review*, 2(2), 121-134.
- Virzi, R.A (1992), 'Refining the test phase of usability evaluation: How many subjects is enough?', *Human Factors*, 34(4), 457-468.
- Whalley, P.C and Fleming, R.W (1975), 'An experiment with a simple recorder of reading behavior', *Programmed Learning and Educational Technology*, 12, 120-123.
- Willows, D.M (1978), 'A picture is not worth always worth a thousand words: Pictures as distractors in reading', *Journal of Educational Psychology*, 70, 255-262.
- Willows, D.M (1980), 'Effects of pictures salience on reading comprehension of illustrated and non-illustrated aspects of text', paper presented to the Annual Meeting of the American Educational Research Association, Boston.
- Winn, W.D (1986), 'Simultaneously and successive processing of circuit diagrams having different amounts of detail', paper presented to the Annual Meeting of the American Educational Research Association, San Francisco.
- Winn, W.D (1987), 'Charts, graphs and diagrams in educational materials', in H.A Houghton and D.M Willows (editors), *The Psychology of Illustration*, Springer-Verlag, New York, 1, 152-198.
- Winn, W.D and Holliday, W.G (1981), 'Learning from diagrams: Theoretical and instructional considerations', in proceedings of Association of Educational Communications and Technology Conference, Philadelphia.
- Witte, R.S (1989), 'Statistics', Holt, Rinehart and Winston, Orlando, Florida, 217-228.
- Wright, P and Fox, K (1970), 'Presenting information in tables", *Applied Ergonomics*, 1, 234-242.
- Wright, P and Lickorish, A (1983), 'Proof Reading texts on screen and paper', *Behaviour and Information Technology*, 2(3), 227-235.
- Zavotka, S.L (1987), 'Three dimensional computer animated graphics: a tool for spatial skill instruction', *Educational Communication and Technology Journal*, 35(3), 133-144.

Appendix One

Main Program for Graphics Experiment

Program Shell;

{SD+}

{L+}

{F+}

{M 65520,0,65520}

USES

Overlay,Dos,Crt,totFAST,totMENU,totMSG,totIO1,totLOOK,tot Str,
Str ,Shelltext, ShellGraphics, ShellFrames, ShellDictionary, ShellQuestion,
ShellMore, ShellExamples;

{O Shelltext}

{O ShellGraphics}

{O ShellFrames}

{O ShellDictionary}

{O ShellQuestion}

{O ShellExamples}

{O ShellMore}

var

Ch : char;

next:string;

frame:integer;

version:integer;

procedure Controls(var frame:integer); forward;

procedure Exit; forward;

procedure Move(var frame:integer); forward;

procedure Icon; forward;

procedure MainMenu;

var

main : MenuOBJ;

choice : integer;

frame :integer;

next:string;

number:integer;

begin

frame:=0;

number:=0;

```

AddToFile(TheFile,Line,frame,Hour,Minute,Second,Sec);
Screen.ClearText(22,4,70,4);
Screen.ClearText(1,8,79,25);
pagezero(Titlecolours,textcolours);
with main do
begin
  Init;
  ShadowTOT^.SetShadowSize(0,0);
  Win^.SetRemove(true);
  SetStyleTitle(4,' Main Menu ');
  AddItem(""); AddItem(""); AddItem("");
  AddItem(' 1 COMPUTER SYSTEM ');
  AddItem(' 2 INPUT ');
  AddItem(' 3 COMPUTER DEVELOPMENT ');
  AddItem(' 4 DATA REPRESENTATION ');
  AddItem(' 5 STORAGE ');
  AddItem(' 6 PROCESSING ');
  AddItem(' 7 OUTPUT ');
  AddItem(' 8 DICTIONARY ');
  AddItem(' 9 QUIZ ');
  AddItem(' 10 EXIT ');
  AddItem(""); AddItem(""); AddItem("");
  SetMenuXY(21,7);
  Win^.SetColors(94,95,80,80);
  SetGap(6);
  SetActiveItem(2);
  Choice := Activate;
  Done;
  Win^.Done;
end;
Case Choice of
4:begin {***** Computer System *****}
  frame:=200;
  AddToFile(TheFile,line,frame,Hour,Minute,Second, Sec);
  GraphOpen; {***** Illustration *****}
  Icon;
  frame:=1;
  frame1;
  Controls(frame);
end;
5:begin {***** Input *****}
  frame:=201;
  AddToFile(TheFile,line,frame,Hour,Minute,Second, Sec);
  GraphOpen1; {***** Illustration *****}
  Icon;
  frame:=10;
  frame10;

```

```

    Controls(frame);
end;
6:begin      {***** Computer Development *****}
    frame:=202;
    AddToFile(TheFile,line,frame,Hour,Minute,Second, Sec );
    GraphOpen7; {***** Illustration *****}
    Icon;
    frame:=25;
    frame25;
    Controls(frame);
end;
7:begin      {***** Data Representation *****}
    frame:=203;
    AddToFile(TheFile,line,frame,Hour,Minute,Second, Sec );
    GraphOpen3; {***** Illustration *****}
    Icon;
    frame:=33;
    frame33;
    Controls(frame);
end;
8:begin      {***** Storage *****}
    frame:=204;
    AddToFile(TheFile,line,frame,Hour,Minute,Second, Sec );
    GraphOpen5; {***** Illustration *****}
    Icon;
    frame:=42;
    frame42;
    Controls(frame);
end;
9:begin      {***** Processing *****}
    frame:=205;
    AddToFile(TheFile,line,frame,Hour,Minute,Second, Sec);
    GraphOpen6; {***** Illustration *****}
    Icon;
    frame:=50;
    frame50;
    Controls(frame);
end;
10:begin     {***** Output *****}
    frame:=206;
    AddToFile(TheFile,line,frame,Hour,Minute,Second, Sec );
    GraphOpen4; {***** Illustration *****}
    Icon;
    frame:=58;
    frame58;
    Controls(frame);
end;

```

```

11:begin      {***** Dictionary *****}
    frame:=207;
    AddToFile(TheFile,line,frame,Hour,Minute,Second, Sec);
    GraphOpenDictionary; {***** Illustration *****}
    number:=frame;
    Search;    {Dictionary;}
    Icon;
    Move(number);
    Controls(frame);
end;
12:begin      {***** Quiz *****}
    frame:=208;
    AddToFile(TheFile,line,frame,Hour,Minute,Second, Sec);
    GraphOpenQuestion;  {***** Illustration *****}
    number:=frame;
    Icon;
    Questionone;
    Icon;
    Move(number);
    Controls(frame);
end;
13:begin
    Exit;
    MainMenu;
end;
end;
procedure Move(var frame:integer);
begin
case frame of
  0:MainMenu;
  1:frame1;
  2:frame2;
  3:frame3;
  4:frame4;
  5:frame5;
  6:frame6;
  7:frame7;
  8:frame8;
  9:begin
    frame9;
    MainMenu;
    end;
  10:frame10;
  11:frame11;
  12:frame12;
  13:frame13;

```

```
14:frame14;
15:frame15;
16:frame16;
17:frame17;
18:frame18;
19:frame19;
20:frame20;
21:frame21;
22:frame22;
23:frame23;
24:begin
    frame24;
    MainMenu;
end;
25:frame25;
26:frame26;
27:frame27;
28:frame28;
29:frame29;
30:frame30;
31:frame31;
32:begin
    frame32;
    MainMenu;
end;
33:frame33;
34:frame34;
35:frame35;
36:frame36;
37:frame37;
38:frame38;
39:frame39;
40:frame40;
41:begin
    frame41;
    MainMenu;
end;
42:frame42;
43:frame43;
44:frame44;
45:frame45;
46:frame46;
47:frame47;
48:frame48;
49:begin
    frame49;
    MainMenu;
```

```

    end;
50:frame50;
51:frame51;
52:frame52;
53:frame53;
54:frame54;
55:frame55;
56:frame56;
57:begin
    frame57;
    MainMenu;
    end;
58:frame58;
59:frame59;
60:frame60;
61:frame61;
62:frame62;
63:frame63;
64:frame64;
65:frame65;
66:frame66;
67:frame67;
68:frame68;
69:frame69;
70:frame70;
71:MainMenu;
end;
end;
procedure Exit;
var
    MsgWin : PromptOBJ;
    ActionCode :tAction;
    ID : byte;
begin
    frame:=700;
    AddToFile(TheFile,line,frame,Hour,Minute,Second,Sec);
    with MsgWIN do
    begin
        Init(1,' Warning ');
        AddLine("");
        AddLine(' Do you really want to exit?');
        AddLine(' please confirm by pressing ');
        AddLine(' EXIT  or ABORT to return');
        AddLine(' to the tutorial. ');
        AddLine("");
        SetOption(1,' ~E~XIT ',69,Finished);
        SetOption(2,' ~A~BORT ',65,Escaped);
    end;
end;

```



```

    ActionCode := Show;
    Done;
end;
If ActionCode = Finished then Halt
else
    Write;
end;
procedure Controls(var frame:integer);
var
    Main,Load:MenuOBJ;
    Choice:byte;
    NumStr:string;
    number:integer;
begin
    number:=0;
    with Load do
    begin
        Init;
        Win^.SetRemove(true);
        Win^.SetClose(false);
        AddItem("");
        AddFullItem('~N~ext ',11,78,"nil");
        AddFullItem('~B~ack ',12,66,"nil");
        AddFullItem('~M~enu ',13,77,"nil");
        AddFullItem('~N~ext Section ',14,78,"nil");
        AddFullItem('~E~xample ',15,69,"nil");
        AddFullItem('~D~ictionary ',16,68,"nil");
        AddFullItem('~Q~uiz ',17,81,"nil");
        AddFullItem('~M~ore ',18,84,"nil");
        AddFullItem('~E~xit ',19,69,"nil");
        SetStyleTitle(1,'Controls');
        SetMessageXY(25,25);
        SetGap(1);
        SetActiveItem(2);
    end;
    with Main do
    begin
        Init;
        Win^.SetRemove(true); Win^.SetClose(false);
        ShadowTOT^.SetShadowSize(0,0);
        AddFullItem('~C~ontrols',1,67,'Press C or Click on box',@Load);
        SetMenuXY(63,22);
        SetStyleTitle(2,"");
        SetMessageXY(25,25);
        SetGap(1);
        Choice:=Activate;
        Done;
    end;
end;

```

```

Load.Done;
Win^.Done;
end;
case Choice of
11:begin
    frame:=frame+1;
    Move(frame);
    Controls(frame);
    end;
12:begin
    frame:=frame-1;
    Move(frame);
    Controls(frame);
    end;
13:begin
    MainMenu;
    end;
14:begin
    if frame < 10 then
    begin
        frame:=10;
        Move(frame);
        Controls(frame);
    end;
    if frame < 25 then
    begin
        frame:=25;
        Move(frame);
        Controls(frame);
    end;
    if frame < 33 then
    begin
        frame:=33;
        Move(frame);
        Controls(frame);
    end;
    if frame < 42 then
    begin
        frame:=42;
        Move(frame);
        Controls(frame);
    end;
    if frame < 50 then
    begin
        frame:=50;
        Move(frame);
        Controls(frame);
    end;
end;

```

```

end;
if frame < 57 then
begin
    frame:=58;
    Move(frame);
    Controls(frame);
end;
if frame < 70 then
begin
    frame:=1;
    Move(frame);
    Controls(frame);
end;
end;
15:begin
    number:=frame;
    SearchExample;
    Icon;
    Move(number);
    Controls(frame);
end;
16:begin
    number:=frame;
    Search;
    Icon;
    Move(number);
    Controls(frame);
end;
17:begin    {***** Quiz *****}
    number:=frame;
    GraphOpenQuestion;
    Icon;
    Questionone;
    Icon;
    Move(number);
    Controls(frame);
end;
18:begin
    number:=frame;
    SearchMore;
    Icon;
    Move(number);
    Controls(frame);
end;
19:Begin
    Exit;
    Controls(frame);

```

```

    end;
  end;
end;
procedure Initialise; {***** initialise variables }
var
  w:integer;
begin
  {*****Set colour display *****}
  Begin
    {** background introduction frames **}
    {** background question frames ** }
    {** colours are 0 = black,1 = blue, 2 = green,3 = cyan **}
    {** 4 = red,5 = magenta,6 = brown,and 7 = grey **}
    {** 8 = darkgray,9 = LightBlue, 10 = LightGreen **}
    {** 11 = LightCyan,12 = LightRed, 13 = LightMagenta**}
    {** 14 Yellow,15 = White**}
    {** 0 to 7 Foreground and background colours**}
    {** 8 to 15 Foreground ** }
    {** loads font for title screen **}
    {** background introduction frames **}
    {** background question frames ** }
  end;{***** Select Colours *****}
end;
procedure Icon;
begin
  ClrScr; TextBackground(Cyan); ClrScr;
  with Screen do
  begin
    FillBox(4,2,15,5,113,3);    { 112 = blue,lightGray}
    FillBox(2,6,17,7,20,2);    { 20 = red,blue}
    WriteBetween(5,15,3,113,'Computer');
    WriteBetween(5,15,4,113,'Studies ');
    FillBox(20,3,75,5,30,2);    { 30 = yellow,blue}
  end;
end;
begin
  OvrInit('Shell.OVR');
  Version:=3;
  frame:=1000;
  OpenFile(TheFile,version,frame,Line,Hour,Minute,Second, Sec);
  TitleGraphics;
  Icon;
  Mainmenu;
  begin
    Ch:= ReadKey;
  end;
end.

```


Appendix 2

Timescale of Research

| | |
|---------------------------------|---------------------------------------|
| September 1989 to March 1990 | Literature Survey |
| January 1990 to January 1991 | Design of Colour Experimental Shell |
| February 1992 to August 1992 | Colour Experiment |
| September 1992 to December 1992 | Review of Colour Experimental shell |
| January 1993 to June 1993 | Design of Graphics Experimental Shell |
| July 1993 to January 1994 | Graphics Experiment |
| February 1994 to October 1994 | Writing Thesis |
| November 1994 | Submission |
| May 1995 to February 1996 | Amendments |
| March 1996 | Revised Thesis |

Appendix 3

Publications and Presentations of Research Results

The Principles of Screen design for Computer Based Learning Materials, 1989, Training Agency

Workshops on Authoring Tools, Open and Flexible Learning Conference, Budapest, June, 1991

Authoring Tools, The British Association of Open Learning, Seminar, Newcastle Upon Tyne, September, 1991

Multimedia Workshop, Nottingham Polytechnic, January 1992

Inter-University Information Council Conference, Oxford, April 1992, Workshops on screen design

Status Conference, Screen Design Guidelines, York, November 1992

The Principles of Screen design for Computer Based Learning Materials, 1992, Second Edition, Employment Department

How are graphics used in Computer Based Learning, British Journal of Educational Technology, Volume 23(3), 1993

Screen Design Workshop, Multimedia Conference, Liverpool, John Moore's University, June 1994

Screen Design of Computer Based Learning, Association for Educational and Training Technology, University of South West, Plymouth, April 1995