INDIVIDUAL DIFFERENCES IN HUMAN-COMPUTER INTERACTION

STEPHEN JAMES WESTERMAN

Doctor of Philosophy

THE UNIVERSITY OF ASTON IN BIRMINGHAM

July 1993

Thesis Summary

The University of Aston in Birmingham

Individual Differences in Human-Computer Interaction

Stephen James Westerman - Doctor of Philosophy - 1993

This thesis initially presents an 'assay' of the literature pertaining to individual differences in human-computer interaction. A series of experiments is then reported, designed to investigate the association between a variety of individual characteristics and various computer task and interface factors. Predictor variables included age, computer expertise, and psychometric tests of spatial visualisation, spatial memory, logical reasoning, associative memory, and verbal ability. These were studied in relation to a variety of computer-based tasks, including: (i) word processing and its component elements; (ii) the location of target words within passages of text; (iii) the navigation of networks and menus; (iv) command generation using menus and command line interfaces; (v) the search and selection of icons and text labels; (vi) information retrieval. A measure of self-report workload was also included in several of these experiments.

The main experimental findings included: (i) an interaction between spatial ability and the manipulation of semantic but not spatial interface content; (ii) verbal ability being only predictive of certain task components of word processing; (iii) age differences in word processing and information retrieval speed but not accuracy; (iv) evidence of compensatory strategies being employed by older subjects; (v) evidence of performance strategy differences which disadvantaged high spatial subjects in conditions of low spatial information content; (vi) interactive effects of associative memory, expertise and command strategy; (vii) an association between logical reasoning and word processing but not information retrieval; (viii) an interaction between expertise and cognitive demand; and (ix) a stronger association between cognitive ability and novice performance than expert performance.

British Library Index Terms:

(i) Cognitive Ability

(ii) Ageing

(iii) Word processing

(iv) Information retrieval

Acknowledgements

I would like to thank Dr. Roy Davies for his highly valued supervision during the course of this research, and Dr. Rob Stammers and Dr. Gerry Matthews for their helpful advice.

I would also like to thank the E.S.R.C. for the Studentship award which made this research possible.

In addition, I am grateful to Robert Cotterill, Sarah Cox, Sarah McCarthy, Mandy Poole, and Elaine Weeks for their assistance in collecting some of the data for the experiments reported in Chapters 2, 4, and 6, and to Ray Taylor for programming the experiment reported in Chapter 6.

Finally, I would like to thank my parents for their support throughout.

List of Contents

List of Figures

List of Tables

Page

Chapter 1

Individual Differences: Previous Research and the Experimental Plan

1.1 The Importance of Individual Difference Research

The ongoing spread of information technology inevitably leads to a wider and more heterogeneous computer-user population. Computer technology has been introduced into a range of work tasks including word processing, record keeping, stock taking, and financial accounting. People are confronted with computer technology in banks, libraries, and at a variety of information points. Improved and expanded facilities are available in electronic mail, various educational software and hypertext systems, and computer printouts must regularly be deciphered in the form of bills and pay slips (Paxton and Turner, 1984). These developments lead to an increasing demand for computer technology which is accessible to, and understandable by, a wide range of non-specialist and infrequent users. This, in turn, emphasises the importance of understanding the nature of individual differences in human-computer interaction. Individual differences represent a substantial source of variance (Nielsen, 1989), more so than for most other work tasks (cf. Egan, 1988). As will be discussed below, there are a range of important dimensions of individual difference which relate to computer-based performance, including cognitive ability, age, and experience. An understanding of the mechanics underlying such individual differences would enable improvements to be made to the design of the interface, instructional materials, training methods, and methods of selection.

1.1.1 The design of the interface

The concept of 'usability' has assumed growing importance in relation to the process of software design. Increasing consideration is being given to the ease with which the user is able to interact with the computer. Shackel (1990) provides a historical perspective of 'usability', and identifies four key elements which must be considered: the user; the task; the system; and the environment. Whilst the present thesis is concerned with the characteristics of the user, of equal importance are the interactive effects of the other three elements. As computing power increases so does the scope to reduce the cognitive load upon the user (Woods and Roth, 1988). In order to do so Streitz (1987) proposes that consideration be given to the degree of 'cognitive compatibility' which exists between the user and the interface. This may relate to the match between the structure applied to the data by the computer and the cognitive model applied to that information by the user (Olson, 1985; cf. Hutchins, Hollan, and Norman, 1986). Alternatively, this process may be concerned with matching the mechanics of the user-computer dialogue to the capabilities and preferences of the user. Both of these issues are examined in more detail in subsequent chapters.

In order to develop 'user friendly' interfaces many major software companies are becoming concerned with the process of usability testing. Usability laboratories are being established in which the performance of representative samples of prospective users can be monitored (Microsoft, 1992). The ensuing data are incorporated into the design process. It is not sufficient for designers and programmers to be the sole judges of the usability of new software products, as has often been the case (Nielsen, 1992). Not only do they possess substantially greater expertise than many prospective users, but they may differ with respect to other important factors such as cognitive ability (cf. Vicente and Williges, 1988). Consequently, the ease with which they use and understand software cannot be held to be representative of that which the general user population will experience (Meister, 1989; Spinas, 1989). In order to better understand the requirements of the user much has been made of the process of 'user modelling' (Rich, 1983; Potosnak, 1984; Jones and Mitchell, 1987; Kass and Finin, 1988). This requires an understanding of "..the user's beliefs, goals and plans, preferences and attitudes, or capabilities..." (Kass and Finin, 1988, p. 145), which can then be used by the designer or by the system to facilitate interaction. The mechanics of such a process are discussed in more detail below. However, it should be noted that the completeness and accuracy of the user model are of fundamental importance. A problem which has been identified with many attempts at user modelling is that complex attributes tend to be simplified (Potosnak, Hayes, Rosson, Schneider, and Whiteside, 1986). The model must reflect all important individual characteristics which bear upon the interaction between user and computer. Equally, if a model developed in the usability laboratory is to be applied to a wider user population it is essential to base it upon a representative sample of that population. In a report on current practice in usability testing for computer documents Rosenbaum (1989) refers to some of these issues. Whilst age and previous computer experience were considered in the selection of a user sample, many other equally important predictors of computer-based performance, as described below, were omitted.

A number of approaches to interface design may be employed in order to improve the level of cognitive compatibility between user and computer. These include the use of robust, adaptive, and adaptable interfaces, each of which will each be discussed in turn.

1.1.1.1 Robust interfaces

A robust interface is one which minimises the variance associated with individual differences whilst maximising overall efficiency levels (Egan, 1988). 'High fliers' should not be disadvantaged by such changes to the interface, but the performance of less able or less experienced users should be facilitated. An example of a robust interface is provided in a series of experiments conducted by Egan and Gomez (1985). Early experiments found age to be strongly associated with word processing performance when using a line editor, with older subjects being at a performance disadvantage. However, in a later experiment a display editor was found to improve the response times of older subjects with no response time cost to younger subjects. Within this context the display editor can be seen to provide a robust interface. It should be noted, however, that in this latter experiment, whilst the error-rate for older subjects was also improved there was some deterioration in the performance of younger subjects. As this perhaps illustrates, there are practical difficulties associated with the creation of robust interfaces. Balancing the needs of all users may prove an intractable problem. An example, which is addressed in detail in Chapter 6, relates to the use of menus and command line interfaces by novice and expert users. Menus are often held to facilitate the performance of novice users by reducing the memory demands associated with the process of interaction. At the same time they may disadvantage expert users by virtue of decreased command flexibility (Shneiderman, 1987a, 1988). In these circumstances it would be difficult to create a robust interface in which only one mode of interaction was available. Given that there may be several important dimensions of individual differences with respect to computer-based performance, each of which interacts with unique interface elements, the lack of flexibility associated with robust interfaces may prove limiting.

1.1.1.2 Adaptive interfaces

An alternative method of accommodating individual differences is to create an interface which adapts to the characteristics of the user (Vaubel and Gettys, 1990). Most of the literature relating to adaptive interfaces has focused upon modelling different levels of expertise (Norcio and Stanley, 1989). One difficulty associated with the implementation of such systems concerns the mechanism by which the computer determines the characteristics of the user upon which interaction will be based. Experiments conducted by Vicente, Hayes, and Williges (1987; Vicente and Williges, 1988) and Elkerton and Williges (1984b, 1985, 1987) suggest that it may be possible to infer individual cognitive or expertise characteristics from the type of commands which the user selects. An alternative solution was examined in experiments conducted by Martin and Fuerst (1987, 1988) in which, prior to task performance, the user was required to rate their level of expertise. Based upon this response one of two interfaces was presented. One of these provided increased support and was thought to benefit novice users, and the other was thought to benefit expert users by virtue of being less cumbersome. As predicted, subjects benefited from using an interface tailored to their level of expertise. An example of an adaptive interface which is based upon a more flexible system of user modelling is provided by Trevellyan and Brown (1987). This small study (n=4) examined information retrieval performance when the accessibility of computer presented information was determined by previous patterns of information demand by the user, the principle being that the system can monitor and respond to its own success or failure in adapting to the user. Further examples are given by Benyon and Murray (1988) in an overview of work undertaken in their laboratory. An adaptive system was being developed, and within this context a number of dimensions of individual differences were being considered.

Whilst the potential benefits of adaptive systems are great, there are also disadvantages associated with their design and use. Firstly, as indicated above, they are difficult to implement. They depend upon the successful identification of the relevant dimensions of individual difference, a knowledge of the mechanics of accommodating these differences, and the capacity to employ these mechanisms. At the moment, empirically derived guidelines to support such designs are not readily available. Secondly, there are additional costs associated with adaptive interfaces, both in terms of the initial programming effort required, and also the additional computer processing requirements of running such systems (Norcio and Stanley, 1989). Finally, it has been argued that users will find it difficult to achieve a mental model of a system which is necessarily labile in its methods of interaction and presentation of information (Greenberg and Witten, 1985).

1.1.1.3 Adaptable interfaces

An alternative design strategy, which does not incorporate many of the difficulties associated with adaptive interfaces, is to provide the user with a choice of dialogue modes (Kantorowitz and Sudarsky, 1989) and / or customisable features (Rich, 1983). Whilst decisions regarding the nature and variety of the available options should be based upon an understanding of the differing demands associated with the user population, it is not necessary for the software to infer the characteristics of a particular user. In this way the adaptable interface eases the burden upon the designer and the system and places the user in control. There are, however, three disadvantages associated with this approach. The first relates to the level of expertise of the user. In order to customise the interface the user must know the customisation procedure and be aware of the implications of each alternative. Whilst the frequent, expert user may enjoy the flexibility provided by an adaptable system, the infrequent, novice user may be unable to make informed choices. A study by Jorgensen and

Sauer (1990) supports this position. It was found that experienced users tended to make use of customisation facilities whilst novice users did not. A second difficulty with these systems is that, for occasional users, the overhead associated with establishing preferred options may be unacceptable, particularly if the task to be performed is brief. Finally, adaptable systems assume that the user has a satisfactory understanding of their own capabilities and will be able to adopt the most efficient performance strategy. There are a number of reasons why this might not be the case, for example when such systems are used in co-operative working environments common practices may prevail (Potosnak, Hayes, Rosson, Schneider, and Whiteside, 1986; Wogalter and Frie, 1990).

1.1.2 The design of training programs and instructional materials

Another area which may benefit from a consideration of individual differences lies in the design of training programs and instructional materials. It may be possible to adapt training strategies and instructional materials to suit the characteristics of the individual thereby improving the rate of skill acquisition. Empirical support for this position is mixed. Cronbach and Snow (1981) present an overview of experiments, in a variety of fields, which have examined the effectiveness of various instructional methods using an Aptitude-Treatment Interaction (ATI) approach. Whilst there was some evidence to suggest that general intelligence interacted with treatment, with subjects of low general intelligence benefiting from more detailed instructional materials, there was little evidence to suggest that more specific abilities reliably benefited from being associated with particular instructional regimes or materials. Within a human-computer interaction (HCI) context the interaction between various types of training regimes and the age of the user have been investigated. The results of these experiments are mixed, but generally do not support an ATI. Zandri and Charness (1989) found a significant interaction between age group, the provision of a jargon sheet, and paired vs individual learning, in relation to performance using a multi-function computer application (see section 1.2.5). However, Gist, Rosen, and Schwoerer (1988) found no significant interaction between age and the use of a 'behavioural modelling' strategy with respect to the acquisition of spreadsheet skills. Similarly, Czaja, Hammond, Blascovich, and Swede (1989) found no interaction between age and the use of instructor-based, on-line, or manual-based instruction for word processing, and Charness, Schumann, and Boritz (1992) found no interaction between age and the provision of an advance organisor for a word processor, or between age and self-paced vs fixed-paced tutorial conditions. There is more compelling evidence for an ATI with respect to individual differences in learning style (see section 1.2.3.2). It would seem that learning style may interact with the structure of computer-based training (Stanton and Stammers, 1989), the type of model (analogical or abstract) presented to the user (Sein and Bostrom, 1989), and the encouragement of active exploration within training programs (Frese, Albrecht, Altmann, Lang, Papstein, Peyerl, Prumper, Schulte-Gocking, Wankmuller, and Wendel, 1988). Similarly, Carroll and Mack (1984; Carroll, 1985) demonstrated the importance of active exploration as a training strategy for novices. Spatial ability has also been found to interact with the type of model (abstract or analogical) used during training, with low spatial ability individuals being disadvantaged by the use of an abstract model (Sein, Bostrom, and Olfman, 1987; Sein and Bostrom, 1989). A review of some of the literature pertaining to training and computer-based performance, along with suggestions for further developments in this area are provided by Gattiker (1992).

1.1.3 Selection

Reliable associations between individual characteristics and computer-based performance can be used to improve the process of occupational selection, selection to educational courses (e.g. Campbell and McCabe, 1984; see Koubek, LeBold, and Salvendy, 1985, for a review) and the identification of training needs (e.g. Kennedy, 1975). Important dimensions of individual difference may include biographical information (Potosnak, 1984), the results of interest inventories, personality questionnaires (see section 1.2.7) or pyschometric tests of ability. A number of commercial ability tests are available for this purpose. Neuman and Nomoto (1990) review five tests which are designed to measure aptitude for general computer programming, computer operation, word processing, data entry and retrieval, and COBOL programming. With the exception of the latter test, each derives a composite score from several components. Whilst Neuman and Nomoto (1990) provide a brief overview of these components, insufficient detail is provided to assess the specific associated cognitive demands. Selected validation studies are also presented which suggest that these tests can make a valuable contribution to the process of occupational selection. Within the broad field of occupational selection, cognitive ability and work sample tests have been demonstrated to provide one of the most reliable methods of selection, with biographical information and personality tests also being stronger predictors than the more frequently used methods of interview and references (Drenth and Algera, 1987; Shackleton, 1989). However, it should be noted that the predictive validity of interviews may vary according to the selection requirements and the structure which is applied to the interview (Ribeaux and Poppleton, 1978; Roth and Campion, 1992). (see Shackleton and Newell, 1991, for data upon the relative frequency with which various methods of management selection are employed.)

1.1.4 The stability of individual differences

A factor which will influence the efficiency with which a user model may be employed in relation to interface design, training design, or selection relates to the stability of the individual differences in question (Van der Veer, 1989a). If differences are labile or subject to environmental influence, as may be the case with the users' level of computer anxiety, then training may be a very suitable intervention, whilst selection decisions based upon this criterion are not likely to prove successful. Similarly, the creation of adaptive interfaces in these circumstances may prove problematic because of the need for frequent re-evaluation of the user-interface match. However, it should be noted that some help systems attempt to deal with exactly these difficulties. Alternatively, if the individual characteristics of concern are relatively stable, such as differences in cognitive ability, then progress made by training interventions is likely to be less pronounced, whilst such characteristics may provide useful information upon which to base selection decisions. Stable individual differences also provide the opportunity of transferring user models between applications and systems (Benyon and Murray, 1988). This may be achieved by transferring a 'user profile' which is stored on disk (Murray, 1991).

1.2 An 'Assay' of Individual Difference Measures

1.2.1 The methodology

In one of the most influential studies within this area Egan and Gomez (1985) proposed a research strategy which involved assaying, isolating, and accommodating individual differences in computer-based performance. The initial 'assay' stage requires the identification of the best predictors of computer-based performance from a range of alternatives. The 'isolation' phase refers to a process of determining the relationship between these predictors and task/interface related variables. This information is required for the final phase of the strategy 'accommodation', which involves making changes to task/interface components in order to reduce the variance associated with individual differences. An example of the use of this strategy within an information retrieval context can be found in the experiments of Vicente, Hayes and Williges (1987; Vicente and Williges, 1988).

This chapter reviews the literature pertaining to an assay of individual differences in computer-based performance. In keeping with the concept of an assay, individual characteristics are considered in turn, and one study which has examined several predictors of computer-based performance may be mentioned a number of times under several different sub-headings. Later chapters are concerned with 'isolating' the relationship between individual characteristics and computer-interface components, and within this context more detailed reports of many studies are provided.

In some of the experiments cited below the authors have reported results as significant if they exceed a 90% confidence interval. However, the more usual 95% confidence interval has been adopted as the criterion of statistical significance throughout this thesis and all results are reported upon this basis. Results which fall between these figures may be reported as trends in the data.

1.2.2 Cognitive ability

Whilst there are various approaches to the study of cognitive ability (cf. Sternberg and Frensch, 1990), experiments relating to human-computer interaction have generally been based upon either psychometric or, to a lesser extent, cognitive (information-processing) models of intelligence. The psychometric approach relies upon the use of correlational techniques such as factor analysis to identify, or 'differentiate', dimensions of individual difference in cognitive ability. The number of factors which are held to exist is subject to continued debate. Guilford (1967), for example, presents a model of intelligence, a 'structure-of-intellect', in which the three orthogonal facets of 'operations', 'products', and 'contents' are identified. This leads to the presumption that 120 cognitive ability factors exist. Guilford (1985) expanded this model to incorporate 150 factors of which he thought over 100 had been empirically established. This model has received much criticism (Carroll, 1986, 1993) on the grounds that it is unnecessarily complex, with many of the identified factors being the result of inappropriate methods of analysis or inappropriate selection of variables. Most of the psychometric tests used to examine individual differences in human-computer interaction have been drawn from the Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, and Harman, 1976), which identifies and provides marker tests for 23 cognitive factors. Carroll (1986) reports results which are similar to this factor solution, and which he suggests bear comparison to the earlier models of intelligence proposed by Cattell (1971; Cattell and Horn, 1978) and Vernon (1961). In a more recent review of the factor analytic literature, which includes a re-analysis of many of the major data sets upon which previous studies had been based, Carroll (1993) presents a hierarchical model of intelligence in which various first-order abilities load onto the eight second-order abilities of fluid intelligence (2F), crystallised intelligence (2C), an indeterminate combination of 2F and 2C (2H), visual perception (2V), auditory perception (2U), cognitive speededness (2S) retrieval ability (2R, and memory ability (2Y). These in turn load onto a third-order general intelligence factor (G). The existence of this general intelligence factor is the subject of some debate (cf. Horn, 1985), however this is not a central issue for the present study.

One criticism of the psychometric approach to the study of intelligence is that it assumes that individuals do not vary in the strategies which they employ when performing ability tests (Kyllonen, Lohman, and Snow, 1984). This concern is addressed by the cognitive (information-processing) approach, the main thrust of which is an attempt to understand the cognitive components and processes involved in the operationalisation of cognitive ability. Whereas the psychometric approach is concerned with the structure of intelligence, the cognitive approach is concerned with the mechanics of intelligence. Within this context individual variation may arise from differences in the efficiency of the processing system, or differences in the efficiency of the algorithm being applied (Hunt, 1986). Using a similar framework, Sternberg (1977, 1985) identified six primary sources of individual variation in information processing: (i) the cognitive components selected for use, (ii) the rules which are applied when combining components, (iii) the sequencing of components, (iv) the mode of processing (exhaustive or self-terminating), (v) the speed and accuracy of processing, and (vi) the nature of the mental representation upon which components act (spatial vs verbal). Examples of a cognitive approach to intelligence, some of which are discussed further in later chapters, can be found in relation to verbal ability (Hunt, 1978; McLeod, Hunt, and Matthews, 1978), spatial ability (Cooper, 1976; Mumaw and Pellegrino, 1984), and reasoning (Sternberg, 1977; Sternberg and Gardner, 1983).

As mentioned, a large number of HCI experiments have used tests from the Ekstrom et al. (1976) Factor Referenced Kit. When reviewing this literature, where possible, individual tests will be specified and related to the underlying factor structure of cognitive ability as proposed by Ekstrom et al. (1976) and Carroll (1993).

1.2.2.1 Verbal ability

There are a number of ability factors which have been identified relating to the use of language. Carroll (1993) points out that the factor structure is complex with many factors which are associated with language ability also being strongly related to memory and reasoning abilities. However, for the present purposes the position is simplified by the fact that HCI studies have broadly limited themselves to the use of the Nelson-Denny Reading Test (1973). This test comprises three sections: vocabulary; comprehension; and reading speed. Vocabulary and comprehension tests are associated with a first-order verbal ability factor (Ekstrom et al., 1976; Carroll, 1993). This, in turn, loads heavily upon a second-order factor of crystallised intelligence (Gc), and a third-order factor of general intelligence (G: Carroll, 1993). Reading speed is identified as a distinct first order factor (Carroll, 1993).

The Nelson-Denny Reading Test has been found to be predictive of the speed (Egan and Gomez, 1985) and accuracy (Sebrechts, Deck, Wagner, and Black, 1984; Egan and Gomez, 1985) of text editing performance. In the series of experiments conducted by Egan and Gomez (1985) this association was modest and only occasionally apparent when a line editor was being used, but strong when a display editor was used. Unsurprisingly, Nelson-Denny scores were also predictive of the time taken to read manuals during these experiments. In a study which examined the performance of younger and older word processing novices, Davies, Wong, Glendon, Stammers, and Matthews (1993) found verbal ability (V; Ekstrom et al., 1976) to be predictive of performance accuracy, but only for the younger group. Vicente, Hayes, and Williges (1987) examined the three sections of the Nelson-Denny Reading Test (1973) as separate predictors of information retrieval performance. Vocabulary test scores were significantly correlated with completion time and the total number of commands used, with high vocabulary subjects performing more quickly and efficiently. The comprehension section was correlated with performance time, although the association was weaker than for vocabulary, and the reading rate section was not significantly related to task performance. Jennings, Benyon, and Murray (1991) also found verbal ability (NFER-Nelson verbal ability test) to be predictive of completion time for an information retrieval task, but only for one of five interfaces which were examined. This interface required subjects to type answers to a series of prompts. A qualified association between verbal ability and performance was obtained by Greene, Gomez, and Devlin (1986) who found that Nelson-Denny scores were modestly but significantly related to the use of a database query language, although only for specific logical operators (language components). A subsequent report of this experiment (Greene, Devlin, Cannata, and Gomez, 1990) indicated that this association was also sensitive to interface factors. Seagull and Walker (1992), however, found no significant association between verbal ability (V4; Ekstrom et al., 1976) and the performance of a menu navigation task.

In summary, verbal ability has generally been found to be a comparatively good predictor of word processing and information retrieval performance, although the studies of Egan and Gomez (1985), Jennings et al. (1991), and Greene et al. (1986; Greene et al., 1990) suggest that this may depend upon interface variables, and the study of Davies et al. (1993) suggests that it may also be related to age. A contrary result was obtained by Seagull and Walker (1992). Worthy of particular note is the finding of a modest and spasmodic association between verbal ability and performance when using a line editor, but a strong association when using a display editor (Egan and Gomez, 1985).

1.2.2.2 Reasoning ability

Ekstrom et al. (1976) identify induction, logical reasoning, general reasoning, and integrative processes (although the evidence is weaker for this latter factor: Carroll, 1993) as first-order factors relating to the process of reasoning. Carroll (1993) identifies three similar first-order factors: (i) induction, (ii) sequential reasoning (encompassing tests of logical reasoning), and (iii) qualitative reasoning. The induction and sequential reasoning factors load more heavily upon a second-order fluid intelligence factor (Gf), whilst the quantitative reasoning factor loads more heavily upon a second-order factor of crystallised intelligence (Gc). However, associations between these factors are comparatively strong, and all are heavily loaded upon a third-order general intelligence factor. It has been proposed (Kyllonen and Christal, 1990) that individual differences in working memory account for virtually all of the variance in the performance of reasoning tasks. Carroll (1993) has also pointed out that many of the tasks used to assess reasoning performance will also be strongly associated with language, number, and spatial skills.

The Diagramming Relations Test (Ekstrom et al., 1976) is a test of logical reasoning ability which has been used in a number of studies of computer-based performance. This test has been found to be predictive of text editing performance (Egan and Gomez, 1985; Czaja, Joyce, and Hammond, 1989) with high logical reasoning scores being associated with quicker and more accurate editing performance. In the database query language study by Greene, Gomez, and Devlin (1986; Greene et al., 1990) logical reasoning was also strongly related to performance. Foreman (1988) examined the Letter Sets Test (I1; Ekstrom et al., 1976) and the Following Directions Test (IP2; Ekstrom et al., 1976) as predictors of a number of component elements of computer programming ability. The Letter Sets Test is a test of inductive reasoning, whilst the Following Directions Test loads on the questionable integrative processing factor. Both of these tests were significantly correlated with all computer programming component tests with the exception of a test of syntactic knowledge. Chen and Vecchio (1992) found a correlation between the abstract reasoning scale of the Differential Aptitude Test (Bennett, Seashore, and Wesman, 1972) and scores on one out of four programming comprehension tests. The results of a study by Garfein, Schaie, and Willis (1988) indicated that various measures of fluid intelligence and reasoning ability, derived from the Primary Mental Abilities Test (Thurstone, 1958) and the Culture Fair Test (IPAT, 1973) were significantly associated with the accuracy of spreadsheet performance. Woelfl (1984) found that subjects with high Symbolic Reasoning Test (Ruch and Ruch, 1957) scores performed an on-line information retrieval task with greater efficiency and greater accuracy. However, Saracevic and Kantor (1988), in a similar experiment, found no significant association.

In summary, logical reasoning has been found to be a comparatively good predictor of text editing and programming performance. However, the evidence as regards information retrieval tasks is limited and mixed.

1.2.2.3 Spatial ability

The number of factors associated with spatial ability is subject to some controversy. Lohman (1979) identified three factors: spatial relations; spatial visualisation; and spatial orientation. He contends that other factors such as closure speed, perceptual speed, and visual memory (spatial memory) are of minor importance and are only identifiable when a number of similar tests are included within the test battery. Following a review of the literature, McGee (1979), proposed a two factor solution, agreeing with Lohman (1979) in the identification of spatial visualisation and spatial orientation. On the basis of experimental evidence (n=150) Meade (1986) proposed a more parsimonious account of spatial ability in which a single factor accounts for all variance. She suggests that multi-factor solutions are the result of over extraction and relate to differences in test difficulty rather than different cognitive abilities. In his recent review of the literature and re analysis of the data Carroll (1993) identifies spatial relations, spatial visualisation, spatial memory, closure speed, closure flexibility, and perceptual speed as first-order factors. He also presents some evidence to indicate the existence of the additional factors of serial perceptual integration, spatial scanning, imagery, and length estimation. All of these factors load onto a second-order visual perception factor. However, in the case of spatial memory there is also a strong loading upon a second-order general memory factor, supported by the strong correlation between spatial memory and associative memory. Carroll points out that, whilst the spatial memory factor relies upon spatial encoding of information, many tests of spatial memory may be performed using alternative strategies and that this may be the source of further individual variation (cf. Cooper and Reagan, 1982). It is interesting to note that Carroll suggests that imagery may be differentiated from spatial visualisation. There are substantial similarities in the factor structure of spatial ability suggested by Carroll (1993) and those identified by Ekstrom et al. (1976), which were flexibility of closure, speed of closure, visual memory, perceptual speed, spatial orientation, spatial scanning, and visualisation.

In their study of information retrieval, Vicente, Hayes, and Williges (1987) examined a battery of spatial ability tests taken from the Ekstrom et al. (1976) Factor Referenced Kit. These included tests of spatial visualisation (VZ1 and VZ2), flexibility of closure (CF2 and CF3), spatial scanning (SS1 and SS2), spatial orientation (S1 and S2) and perceptual speed (P2 and P3). They found significant correlations between (i) spatial visualisation and completion times and command efficiency, with VZ2 (the Paper Folding Test) being the stronger predictor of performance; (ii) between spatial scanning (SS1) and completion times; and (iii) between flexibility of closure and completion times. Further evidence of an association between spatial visualisation (VZ2; Ekstrom et al., 1976) and information retrieval performance speed is provided by Campagnoni and Erlich (1989) and Seagull and Walker (1992). Jennings, Benyon, and Murray (1991) examined information retrieval performance with five alternative interfaces. Spatial ability (Saville and Holdsworth spatial ability test ST7) was predictive of performance when using two of these interfaces, both of which required typed responses as opposed to menu selection. Spatial memory was found to be predictive of the production of database queries using the 'AND' logical operator (Greene et al., 1986; Greene et al., 1990), and the speed and accuracy of performance of menu navigation task (Billingsley, 1982).

A battery of spatial ability tests, also drawn from the Ekstrom et al., (1976) Factor-Referenced Kit, was used by Davies et al. (1993) in their study of age differences in word processing. All tests were reported to be predictive of the speed and accuracy of performance of an older novice sample (45 years +). Gomez, Egan. Wheeler, Sharma, and Gruchacz (1983: also reported in Egan and Gomez, 1985) examined the Two-Dimensional Space Test as a predictor of text editing performance. Scores were significantly correlated with reading time (time spent reading the manual), execution time per successful change, cursor moves per successful change, and first try errors, with high spatial ability being associated with faster and more accurate performance. Spatial memory (MV2: Ekstrom et al., 1976) has also been found to be predictive of the speed and accuracy of text editing performance (Egan and Gomez, 1985; Czaja, Joyce, & Hammond, 1989), although not by Sebrechts, Deck, Wagner, and Black (1984). Bringelson and Eberts (1990) examined the impact of spatial visualisation (VZ2) upon goal structure during the performance of a text editing task. The level of workload experienced by subjects was manipulated using a secondary working memory task. A number of factors make it difficult to draw firm conclusions, although the results appear to suggest that high spatial subjects used more complex, and perhaps less efficient, goal strategies than low spatial subjects, a pattern which became more pronounced under increased workload.

In Foreman's (1988) study of computer programming skill, spatial ability (VZ2) was only found to be correlated with the debugging and modification component tasks. These were the tasks considered to be the most demanding. Other component tasks included program comprehension, program composition, and knowledge of syntax. Spatial visualisation (VZ2) has also been found to be associated with the performance of file manipulation tasks, with high spatial ability subjects performing better than low spatial ability subjects, particularly when using an abstract, as opposed to analogical, model of the task (Sein, Bostrom, and Olfman, 1987; Sein and Bostrom, 1989; Davis and Bostrom, 1993). An association between spatial ability and mental representation was also found by Van der Veer and Wijk (1988). Spatial ability, as measured using the DAT Space Relations Test (Bennet, Seashore, and Wesman, 1972) was associated with better knowledge of spreadsheet functions, following a course of training. Similarly, Van der Veer (1989a : n=10) found that subjects designated as 'imagers' had a better understanding of the semantics relating to the use of a computer-based 'office system', and Van der Veer (1989b) found high spatials to have a more complete mental representation of an electronic mail system (subjects were designated 'imagers' or 'non-imagers' using a procedure related to associative memory detailed in Van der Veer, Van Muylwijk, and Van der Wolde, 1978). Garfein, Shaie, and Willis (1988) examined the Culture Fair Test (IPAT, 1973) and the Primary Mental Abilities Test (PMA: Thurstone, 1958) in relation to the accuracy of spreadsheet performance. The Space measure of the PMA was one of the stronger predictors.

In summary, spatial ability has been found to be one of the strongest and most consistent predictors of computer-based performance. In particular, spatial visualisation appears to be strongly associated with the performance of information retrieval tasks, complex programming tasks, and file manipulation tasks. A number of studies have found spatial memory to be predictive of word processing performance, although a contrary result was obtained by Sebrechts et al. (1984). There is also some evidence to suggest that high spatial ability is associated with more comprehensive mental representation of computer-based systems.

1.2.2.4. Short-term memory and processing speed

This section reviews the evidence relating to measures of short-term memory along with evidence pertaining to processing speed. It should be noted that the term 'short-term memory' was selected to describe the contents of this section, as distinct from 'working memory', in order to encompass related psychometric tests. There are no grounds, however, to suggest that working memory (Baddeley, 1986) relates to memory factors identified in psychometric studies (Carroll, 1993).

Given that empirical evidence indicates a strong association between working memory and measures of reasoning ability (Kyllonen and Christal, 1990), and that reasoning has been found to be predictive of computer-based performance, an association with working memory can also be predicted. Consistent with this position, Shute (1991) found a working memory factor (derived from a factor analysis of a battery of computer-based tests) predicted most of the variance in programming skill acquisition. Benyon and Murray (1988) refer to experiments conducted in their laboratory in which an unspecified measure of short-term memory was found to interact with dialogue speed. Knowledge acquisition was greater for those with poor short-term memory when dialogue speed was slow. Benyon and Murray conclude that compensation can be made for poor short-term memory by adapting dialogue structure. In the study of menu navigation by Billingsley (1982), mentioned above, a psychometric test of associative memory (MV3: Ekstrom et al., 1976) was found to be modestly predictive of performance speed but not accuracy. However, Gomez, Egan, and Bowers (1986) and Sebrechts, Deck, Wagner, and Black (1984) found no association between associative memory (MA2 and MA1 respectively; Ekstrom et al., 1976) and word processing performance. Similarly, Jennings, Benyon, and Murray (1991) examined an unspecified test of short-term memory in relation to the performance of an information retrieval task and found no significant correlation.

There is little direct evidence of an association between processing speed and computer-based performance. In the study mentioned above, Shute (1991) found a modest association between processing speed and programming performance. However, in the study of information retrieval performance by Vicente, Hayes and Williges (1987) no significant association was apparent between performance and a measure of choice reaction time. Similarly, Gomez, Egan, Wheeler, Sharma, and

Gruchacz (1983) found no significant association between a psychomotor test (the Tapping Test; Parker and Fleishman, 1961) and text editing performance. However, Seagull and Walker (1992) hypothesise that their finding of an association between spatial ability and information retrieval performance was in fact attributable to individual differences in processing speed. This is discussed further in Chapter 7.

In summary, there are too few studies which have related short term memory to computer-based performance for firm conclusions to be drawn. However, it would appear that working memory may be predictive of tasks with strong reasoning components such as computer programming. The association between short-term memory and computer-based performance, however, may interact with interface variables. There is some evidence to suggest a modest association between psychometric measures of short-term memory and the performance of information retrieval tasks, although studies which have examined word processing performance report no similar association. There is little empirical evidence to support processing speed as a predictor of performance. However, this may be heavily dependent upon the nature of the task and the skill level of the user. There are grounds to believe that processing speed may be more strongly associated with skilled performance which employs high levels of automatic processing (Ackerman, 1988: see below).

1.2.2.5 Computer programming tests

There are a number of tests which have been devised specifically to predict computer programming performance. The Aptitude Assessment Battery: Programming (Wolf, 1970) comprises five problems which require the "...manipulation of precisely defined symbols, rigid adherence to instructions, tight logical reasoning, and the use of flow charts." (DeNelsky and McKee, 1974, p.131). This test was found to be moderately predictive of training performance, as assessed by tests and supervisors rating of job performance (DeNelsky and McKee, 1974). The PAT (IBM Programmer's Aptitude Test) comprises three parts: (i) series completion, in which subjects are provided with a series of numbers or letters and required to provide the next number or letter to complete the series, (ii) figure classification, in which subjects are presented with a series of figures and required to select another figure to continue the series, and (iii) arithmetic reasoning, in which subjects must solve problems, selecting their answer from multiple choices. Mazlack (1980) found the PAT to be a relatively poor predictor of the acquisition of computing skills in an introductory course taught to an undergraduate sample.

1.2.2.6 Other related aptitudes

Borgman (1984, 1986) found academic major to be predictive of information retrieval performance, with science and engineering majors being more successful than humanities and social science majors. In an experiment designed to clarify the reasons for this association Borgman (1989) found that science and engineering majors demonstrated greater 'technical aptitude'. Gray, Barber, and Shasha (1991) found academic achievement to be predictive of the number of correct answers obtained during information retrieval. Overall performance time was not affected, although high academic achievement was associated with less efficient search paths. They suggested that high academic achievers were prepared to 'get lost' in the database knowing they were able to recover their position. Van der Veer (1989a) reports earlier experiments conducted in his laboratory in which level of mathematical ability was not found to be associated with learning a programming language. However, 'non-maths' subjects reported greater difficulty and wrote less readable programs. Subsequent experiments using a graphical programming environment overcame these problems. Similarly, Wong (1987) found no significant relationship between academic major and performance whilst using a multi-application package (including word processor and spreadsheet), and Mazlack (1980) found no significant relationship between academic discipline and the acquisition of programming skills.

Card, Moran, and Newell (1983) report a small (n=8) word processing experiment in which a performance superiority was found for 'technical' (had written at least one major piece of code) over 'non-technical' (no programming experience) users who took approximately 1.5 times longer to complete the task, however the significance level of this difference is not reported.

1.2.3 Cognitive style

An alternative approach to the study of individual differences is that of cognitive style. A number of dimensions of cognitive style have been identified (cf. Messick, 1984; Tiedemann, 1989) in which a common approach can be seen. Whereas, in the study of cognitive ability emphasis is placed upon "...the measurement of competencies in terms of maximal performance" (Tiedemann, 1989, p. 263), the study of cognitive style "....implies the measurement of propensities in terms of typical performance with the emphasis on a predominant or customary processing mode" (Tiedemann, 1989, p. 263). Whereas ability is regarded as a unipolar dimension with high ability the preferred capacity, cognitive styles are considered to be bipolar, with neither extreme necessarily more valuable. Whilst a thorough review of cognitive style is not attempted here (cf. Messick, 1984; Robertson, 1985; Tiedemann, 1989), the following sections present a number of studies which reflect a convergence of human-computer interaction research around a handful of these measures.

1.2.3.1 Field dependence-field independence

A number of studies have examined human-computer interaction in relation to Field Dependence-Field Independence (FI-FD : Witkin and Goodenough, 1981). This is purportedly a measure of cognitive style rather than cognitive ability, with neither extreme (FI or FD) being associated with necessarily better performance. Field dependents are said to adopt a holistic approach to problem solving whilst field independents adopt an analytical approach. In addition other personality variables are thought to relate to this construct, with field dependents being more social and conforming than field independents who are more autonomous and less concerned with social graces. However, it has been convincingly argued (McKenna, 1984; Robertson, 1985; Tiedemann, 1989) that measures of FI-FD in fact reflect differences in cognitive ability, and closely resemble measures of spatial ability or fluid intelligence. Indeed, Van der Veer and Van der Wolde (1982: cited in Waern, 1989) found that when intelligence is controlled, the problem solving difference associated with FI-FD disappears. The Group Embedded Figures Test (GEFT: Oltman, Ruskin, and Witkin, 1971), one of the most frequently used measures of FI-FD, loads heavily upon the flexibility of closure factor, mentioned above in relation to spatial ability (Ekstrom et al., 1976; Carroll, 1993). Consequently, it is argued that these findings can most appropriately be considered within the context of a second-order spatial ability factor, with field independent individuals having greater spatial ability than field dependent individuals.

A study by Wong (1987) illustrates the similarity of these constructs. Whilst purporting to relate cognitive style (FI-FD) to performance upon a multi-application package (including word processor and spreadsheet), FI-FD was measured using a test of flexibility of closure drawn from the Ekstrom et al. (1976) Factor-Referenced Kit. No association with performance was found. A similar result was obtained by Vicente et al. (1987) and Jennings et al. (1991) who examined FI-FD as a predictor of information retrieval performance. No significant correlations were found with either completion time (Vicente et al., 1987; Jennings et al., 1991) or command efficiency (Vicente et al, 1987).

However, when FI-FD has been found to be related to computer-based performance the direction of the association is consistent across studies, with field independent subjects proving to be more able computer users. Benbasat, Dexter & Masulis (1981) examined GEFT scores in relation to interface characteristics such as dialogue type (user-guided vs system-guided), type of output requested (e.g. number of graphs requested), and length of commands during the performance of a decision making game in which subjects assumed the position of manager of a company. More FIs completed the task, and there was a tendency for FIs to request more graphs during task performance. Stevens (1983) found FI-FD (GEFT) to be predictive of performance upon a college computing course, with field independent subjects performing better. Morrison and Noble (1987) considered FI-FD as a predictor of performance upon a 'videotex-type task', in which subjects were required to simulate the booking of aircraft tickets, including the use of a database and an electronic mail system. They found significant correlations between EFT scores and performance time and also the amount of the task achieved with field independent subjects performing better than field dependent subjects. Similarly, Foreman (1988) examined GEFT scores in relation to a number of component elements of computer programming competence. A significant positive correlation was found for all elements.

Whilst there is some evidence to suggest that FIs and FDs adopt different command strategies when performing computer-based tasks, this is not incompatible with the previous ability related interpretation of the construct. Ambardar (1984) examined the effects of FI-FD upon the performance of a task which simulated the maintenance of personal financial records. FDs were found to prefer the use of a sequential item number search mode, whilst FIs preferred a key-word search mode. Preferred search mode was apparently related to speed of target location, with both FDs and FI subjects retrieving more quickly whilst using their preferred search method. However, Ambardar does not report whether the magnitude of these effects was significant. Consistent with these results, Fowler, McCauley, and Fowler (1985) found FIs to prefer an unstructured method of generating commands whilst FDs preferred a sub-structured system. However, in both of these results FI was found to be associated with the more demanding command method, as would be predicted if FI was associated with greater cognitive ability. In a small study (n=7) of users learning an operating system, Coventry (1989) found FDs tended to seek help more frequently whilst FIs engaged in active exploration. However, the small sample size makes it difficult to interpret these results in the light of possible differences in ability.

In summary, the concept of FI-FD has been criticised upon the grounds that the most widely used tests appear to be measuring fluid or spatial ability. A number of studies have examined FI-FD in relation to computer-based performance. Results have been mixed, however when FI-FD is significantly related to performance the result is consistent with an ability interpretation, with FIs outperforming FDs.

1.2.3.2 Learning style

The most frequently used measure of learning style within an HCI context is the Learning Style Inventory which was developed by Kolb (1971). This measure assesses "...an individual's relative emphasis upon four learning abilities - Concrete Experience (CE), Reflective Observation (RO), Abstract Conceptualisation (AC), and Active experimentation (AE) - plus two combination scores that indicate the extent to which an individual emphasises abstractness over concreteness (AC - CE) and the extent to which an individual emphasises action over reflection (AE - RO)" (Kolb, 1981, p. 290). These scores are used to define a four cell matrix as shown in Table 1.01. This measure has been criticised as possessing poor psychometric properties, and consequently being unable to support the underlying theory (Freedman and Stumpf, 1980; Stumpf and Freedman, 1981: see also Kolb, 1981).

Table 1.01: Kolb's (1971) Learning style matrix

	Active Experimentation	Reflective Observation
Concrete Experience	Accomodator	Diverger
Abstract Conceptualisation	Converger	Assimilator

Woelfl (1984) found a modest effect of learning style (Kolb, 1971) in relation to on-line information retrieval performance. Reflective observation was associated with greater search effort, active experimentation with less. In a similar experiments Saracevic and Kantor (1988) found abstract conceptualisation to be associated with better performance, and Logan (1990) found that assimilators were more active in their searching whereas the reverse was true for accomodators. Vicente, Hayes, and Williges (1987) included the Abstract Orientation Scale (O'Connor, 1972) in their examination of information retrieval performance. This scale assesses people with respect to abstract or concrete thinking styles. No significant correlations with performance were found.

Olfman (1987) examined the effect of learning style (Kolb, 1971) upon perceived ease of use and post training knowledge of a spreadsheet package. No significant effects were found. In the file manipulation experiment of Sein and Bostrom (1989) abstract learners were found to perform better than concrete learners when task complexity was high. There was also an interaction between learning style and training condition such that abstract learners performed better when presented with an abstract model of the system, whilst concrete learners performed better when presented with an analogical model. Davis and Bostrom (1992) examined the effects of learning style upon a file manipulation task using command line and graphical user interfaces, however no significant differences were found.

A measure of learning style based upon Pask's (1977) theory was examined by Van der Veer (1989a) in relation to the acquisition of a programming language. This included the factors of inclination to learn (inclination to put effort into memorisation), operation learning (inclination to deduce specific rules and procedural details), and comprehension learning (inclination to induce general rules and descriptions). All three of these measures were significantly correlated with the time taken to learn a programming language, even when fluid intelligence (Ravens Advanced Progressive Matrices) was partialled out.

A number of other learning styles or learning strategies have been identified with respect to computer-based performance. Shute (1991) found an interactive effect of what she termed 'learning style' variables (e.g. frequency of help requests) and performance level (good vs poor) which was predictive of the acquisition of programming skills. Brindle (1981) used a model of cognitive style in which subjects were designated as receptive or perceptive with respect to information gathering, and systematic or intuitive with respect to information evaluation. However, the tests used to determine these types were drawn from the Educational Testing Service Factor Referenced Kit of cognitive ability tests (primarily relating to spatial and verbal ability). Results were not generally significant and no obvious patterns were apparent. Frese et al. (1988) examined the effects of a learning style (a measure of their own devising which assessed preference for learning by study as opposed to a preference for active exploration) upon text editing performance. Significant performance advantages were associated with a preference for active exploration. In an experiment which investigated computer assisted learning, Stanton and Stammers (1989) identified 'top down' (look at most important information first) and 'sequential' performance strategies. There was a training time advantage for the former strategy, although no difference in subsequent transfer performance.

In summary, Kolb's (1971) Learning style inventory has been used in a number of studies of computer-based performance. The results are mixed, with either no significant association with performance, or modest, but inconsistent differences between types. However, a promising interaction was obtained with the type of model (abstract or analogical) presented to users to describe the computer-based task (Sein and Bostrom, 1989). Various other measures of learning style and learning strategy have been examined. Again, no consistent pattern of results is apparent.

1.2.4 Computer expertise

1.2.4.1 The nature of expertise

Implicit within the nature of expertise is an increased task specific skill level which differentiates the expert from the non-expert. Whilst there is much evidence of such variance within the HCI field, in itself this is not of particular relevance to the present discussion and experimental evidence of expert performance superiority will therefore not be presented. What is of interest, however, is the mechanics which underpins these skill differences and the processes involved in the acquisition of expertise.

It has been proposed that expertise should be distinguished from experience (Salthouse, 1991a). Whilst these factors may be highly correlated they are not synonymous. Similarly, Fisher (1991), in a discussion of expertise and computer-based performance, distinguishes between novice and naive users and between experienced and expert users (see also Paxton and Turner, 1984). The difference between these groups is held to reflect an additional expertise dimension of 'system understanding' as well as the more usually considered dimension of frequency of use. Unfortunately, such distinctions have not generally been applied within the context of HCI. An example of a classification system which is more frequently applied is provided by Shneiderman (1987a) who suggests that users may be categorised as 'novice users' (no syntactic knowledge of the system; little semantic knowledge of computers; possible computer anxiety), 'knowledgeable intermittent users' (maintain semantic knowledge but have difficulty maintaining syntactic knowledge), or 'frequent users' ('power users' who are as familiar with both syntactic and semantic aspects of the system).

1.2.4.2 The acquisition of expertise

A number of studies have investigated the process of acquiring expertise in HCI. Card, Moran, and Newell (1983) demonstrated the application of the Power Law of Practice (Newell and Rosenbloom, 1981) to the use of four different text selection devices. In a series of experiments Carroll and Carrithers (1984; Carroll, 1985) identified active exploration as an important factor in the speed of word processing skill acquisition. Restricted interfaces which prevented certain types of errors, and minimised and adapted instructional materials were found to facilitate this process. Supportive results were obtained in a study of statistical computing by Green and Gilhooly (1990). This experiment examined individual differences in the types of performance strategy which were associated with effective learning. Results suggest that individuals who performed effectively used a structured, goal-directed approach which included active exploration with attention to feedback. A second experiment indicated that training in these methods facilitated performance.

A frequently cited model of skill acquisition is provided by Anderson (1982, 1983, 1987). This proposes that knowledge may be declarative, as defined by a consciously accessible propositional network, or procedural, as defined by production systems which comprise automated rules for skilled behaviour. These knowledge states can be compared to the distinction between controlled and automatic processing (Schneider and Shiffrin, 1977). Anderson suggests that skill is acquired through a process of compilation which comprises two elements: (i) proceduralisation, in which declarative knowledge is incorporated into productions; and (ii) composition, in which sequences of productions are collapsed into single productions. Applied examples of some of these principles, within an HCI setting, can be found in Anderson and Reiser's (1985; Anderson and Skwarecki's, 1986) description of the development of a LISP tutor, and Singley and Anderson's (1985) investigation of the transference of word processing skills. A review of some further recent approaches relating to the acquisition of programming skills is provided by Elsom-Cook (1989).

1.2.4.3 Novice - expert differences

The study of expertise has often focused upon a comparison of novice and expert performance. This has occurred in a number of fields, e.g. chess (deGroot, 1978; Simon and Chase, 1973; Charness, 1991), solving physics problems (Chi, Feltovich, and Glaser, 1981; Anzai, 1991), and music (Sloboda, 1991). However, computer programming is also one of the more frequently selected domains of study, and results are broadly consistent with those in other areas. It has been shown that expert computer programmers organise knowledge into more meaningful chunks (McKeithen, Reitman, Rueter, and Hirtle, 1981), rely more heavily upon semantic content (Bateson, Alexander, and Murphy, 1987), use automatic processing more frequently (Wiedenbeck, 1981), and are more efficient in structuring programming problems (Jeffries, Turner, Polson, and Atwood, 1981; Peio and Klein, 1984). Koubeck and Salvendy (1991) also found that the efficiency of representation remains a factor in the process of skill development when comparing 'super-experts' to experts, although the level of automatic processing appeared to have reached asymptote.

Given such differences between novices and experts in the structure and access of knowledge, it is not surprising that broader strategy differences have been identified relating to the performance of a range of HCI tasks. Differences in command use have been established within an information retrieval setting (Elkerton, Williges, Pittman, and Roach, 1982; Elkerton and Williges, 1984a). Experts were found to use fewer but more powerful commands. In subsequent experiments (Elkerton and Williges, 1984b, 1987) it was found that novice search performance could be improved by using an interface which modelled expert search behaviour. Contrary to these results, Rosson (1983) discovered that whilst experts were able to produce text editing file changes more quickly, with fewer commands, this was not due to the use of more powerful commands. Experts, however, did use a wider range of commands. Santhanam and Wiedenbeck (1993) compared the word processing performance of 'discretionary' (intermittent) users to that of novices and experts. Discretionary users were found to use a small command subset to perform tasks, within which their performance compared to experts. However, if extended beyond this subset their performance resembled that of novices. Rosson (1984a, 1984b) found that experts made greater use of customisation facilities (macros) when performing a text editing task, although this was also dependent upon the amount of expertise with similar applications. In a field study of clerical workers Prumper, Frese, Zapf, and Brodback (1991; Prumper, Zapf, Brodback, and Frese, 1992) discovered that experts made no fewer errors (other than knowledge errors) but were quicker to recover from these errors. A common finding is that novices will employ inappropriate task analogies drawn from non-computing domains (Mack, Lewis, and Carroll, 1987; Dewier and Karnas, 1991). Waern (1985) presents an analysis of the process involved in such transfer learning. She identifies situational similarity, strength of rule, level of automatic processing, and feasibility of old strategies as being important determinants of the transfer effect. A method of identifying differences in the performance of novices and experts using multidimensional scaling is presented by Schvanneveldt, Durso, Goldsmith, Breene, and Cooke (1985). Much has been made of the potential interaction between expertise and interface variables. An experimental example of such an interaction was given earlier (Martin and Fuerst, 1987, 1988) and these issues are discussed at greater length in Chapters 3 and 6.

1.2.4.4 The interaction between cognitive ability and expertise

In studies of information retrieval performance Vicente et al. (1987) and Vicente and Williges (1988) found that most of the variance associated with expertise could be accounted for by user differences in spatial ability. A similar association was found in the management simulation study, mentioned above, by Benbasat, Dexter, and Masulis (1981) in which subject expertise was found to be significantly related to mathematical expertise. It is not clear to what extent these results can be attributed to a process of self-selection, with individuals of low ability choosing not to use computers, or to an improvement in spatial skill as a result of computer use. Further research is needed to address this issue.

Ackerman (1988) proposes a model which links the process of skill acquisition to individual differences in cognitive ability. He relates a power-speed dimension of cognitive ability to three stages of skill acquisition (based upon Anderson, 1982, 1983): a declarative stage, a compilation stage, and a procedural stage. At the 'power' end of the cognitive ability dimension, broad abilities (e.g. general intelligence) are thought to be most strongly related to the declarative stage of skill acquisition. In the middle range of the cognitive ability dimension, perceptual speed factors are held to be most strongly related to the compilation stage. Finally, it is proposed that, at the speed end of the cognitive ability dimension, psychomotor skills are most strongly related to the procedural stage. This model has been applied to performance in an information retrieval setting by Seagull and Walker (1992 : see chapter 7). A broadly similar, although less refined, model of individual differences in skill acquisition was developed by Fleishman (1972). Within this context it can be seen that predictors of computer-based performance will vary as the user acquires expertise. Assuming that there is a degree of task consistency (cf. Schneider and Shiffrin, 1977; Fisk, 1987), whilst high level cognitive abilities will be predictive of early performance, perceptual speed and psychomotor skills will be predictive of later skilled performance when increased levels of automatic processing are used.

1.2.5 Age

There are a number of reasons to suppose that middle-aged and older adults will be at a disadvantage when using computer technology. Firstly, there are large cohort differences in exposure to computers. Whilst in recent years computers have been included in a range of school curriculum activities, and the proliferation of computer games serves to increase the computer literacy of the young, middle-aged and older adults have generally not been provided with the same opportunities or incentives to familiarise themselves with computer technology. On the basis of two surveys which they conducted, Breakwell and Fife-Schaw (1988) report that older adults feel less able to master new technology and are less willing to invest effort in doing so. Interestingly, however, this was not found to be attributable to a fear of new technology. Secondly, physiological changes such as age-related declines in vision and audition (cf. Stuart-Hamilton, 1991) will tend to disadvantage older users. Particular difficulties can arise, for example, from the use of bi-focal glasses which make it difficult to quickly glance from keyboard to screen and vice versa. Thirdly, there are a number of age-related changes in cognitive ability which can be predicted to interact with computer-based performance. Of particular interest within this context are changes in verbal ability, reasoning ability, spatial ability, working memory, and processing speed. Whilst performance upon many verbal ability tests remains constant or improves well into old age (Davies, Taylor, and Dorn, 1992) reasoning ability and spatial ability have been found to decline (Salthouse, Mitchell, Skovronek, and Babcock, 1989; Salthouse, Babcock, Skovronek, Mitchell, and Palmon, 1990; Salthouse and Mitchell, 1990). A similar, if overly simplistic, distinction can be drawn between the effects of aging upon Gc and Gf (cf. Horn and Hofer, 1992). Salthouse (1992) presents evidence which suggests that this decline in fluid ability is related to changes in the efficiency of working memory, and that this, in turn, is strongly related to age differences in processing speed (Salthouse, 1991b; Salthouse and Babcock, 1991). This position is consistent with the results of Hertzog (1989) who found that a substantial proportion of age-related variance in scores upon the Primary Mental Abilities Test could be accounted for by variance in perceptual speed. Differences in processing speed (Cerella, Poon, and Williams, 1980) and resource availability (Salthouse, 1991c) are also thought to account for an association between age and complexity such that the performance of older individuals is subject to proportionately greater deterioration as task complexity increases.

As may be predicted from this pattern of change, age has consistently been shown to be predictive of the speed with which computer-based tasks are performed. However, the pattern of results with respect to response accuracy is less clear. In a study relating to the use of word processors Hartley, Hartley, and Johnson (1984) found that after an initial training period there were no differences between young subjects (aged 18-30 years) and older subjects (aged 65-75 years) in the recall of information about the word processor or "in the correctness and efficiency with which computer operations were carried out". Older subjects, however, did take more time to perform certain components of the task, and required more assistance. In an examination of skill transference to a second word processor no age differences were found. Contrary results were obtained by Charness, Schumann, and Boritz (1992) in two experiments which investigated the effects of the provision of an advance organisor, and self-paced vs fixed paced tutoral conditions upon word processing performance. Older subjects (over 50 years) had lower scores on a post tutorial test of word processor functions than younger subjects (under 40 years) in both experiments. However, in common with the results of Hartley et al., older subjects took longer to complete the tutorial (in self-paced conditions), and requested more assistance. Elias, Elias, Robbins, and Gage (1987) also found a performance speed disadvantage for older subjects on a word processing task, but age differences were only apparent for specific types of errors. Czaja, Joyce, and Hammond (1989) examined the performance of a group of older, computer naive, adults (women aged 40-70 years) whilst using three different word processors. Age was predictive of both task completion times and accuracy of performance. Czaja, Hammond, Blankovich, and Swede (1989) found age differences in word processing performance time and also certain types of errors, with younger subjects performing better than older subjects. However, an examination of different training regimes produced no interactive effects. Egan and Gomez (1985) also found significant correlations between age and word processing performance time and accuracy when using a line editor. However, Gomez, Egan, Wheeler, Sharma, and Gruchacz (1983) found that age was only significantly associated with performance time when subjects used a display editor. In an experiment which is described in more detail in chapter 2, Egan and Gomez (1985) also examined the association between age and the component elements of text editing performance. Age was significantly correlated only with those elements concerned with the process of command generation. Egan and Gomez (1985) suggest that this pattern of results may relate to interactions between age and task complexity.

Wong (1987) found a significant effect of age (range 23-44 years) upon performance accuracy whilst using a multi-application package (including spreadsheet and word processor). Similarly, Zandri and Charness (1989) found that older adults (58-84 years) were slower in learning to operate a multi-function application (Borland 'Sidekick'), obtained worse test scores, and requested help more frequently than younger adults (20-38 years). There was, however, an interaction with training regime such that older subjects obtained higher test scores when learning individually and provided with a 'jargon sheet'. Davies, Wong, Glendon, Stammers, and Matthews (1993) also found an age related decline in word processing speed, however this appeared to be due to a speed-accuracy trade off, with older subjects performing more accurately. Gist, Rosen, and Schwoerer (1988) examined age differences in the acquisition of spreadsheet skills under two different training regimes. One provided an interactive computer-presented tutorial, whilst the other included a video tape demonstration coupled with computer-based practice. The dependent measure was the number of tasks performed correctly. The performance of older subjects was significantly more error prone with no interactive effects of training method. Garfein, Schaie, and Willis (1988) examined the effects of age and sex upon the performance of computer novices upon a spreadsheet task. Only accuracy data were collected. No significant differences were found for either age or sex. The age range of the sample in this experiment was very restricted, ranging from 49 - 67 years. Analyses involved splitting this group in two giving a younger (49-57) and an older (58-67) group. Given this restricted age range, the non significant effect of age is not surprising. Age (20 to 61 years : mean = 45 years) was found to be predictive of the use of a database query language by Greene, Gomez, and Devlin (1986). However, this was dependent upon the logical operator (language component) required for the query and the type of interface (Greene, Devlin, Cannata, and Gomez, 1990).

In summary, a number of experiments, mainly related to word processing or the use of spreadsheets have found a consistent age difference in performance speed, with older individuals performing more slowly. The pattern of results with respect to performance accuracy is less clear. Frequently older subjects perform at a similar level of accuracy to younger subjects. However, there is also some evidence to suggest that they may be prone to specific types of error.

1.2.6 Sex

A variable which has received comparatively little previous research attention within the context of HCI is that of gender. This has also been said to be true of the wider field of human factors research (Hudgens and Billingsley, 1978). There is much evidence to suggest a male advantage with respect to the performance of tasks which require spatial ability (Campos & Cofan, 1986; Sanders, Cohen, & Soares, 1986; Pepin, Beaulieu, Matte, & Leroux, 1985; McGee, 1979). This appears to be a relatively robust effect, although generally not of great magnitude (Burnett, 1986; Hiscock, 1986). It should be noted, however, that there are those who contest this view on the grounds that there is: (i) an under-reporting of non-significant results; (ii) an overestimation of the importance of positive findings; and (iii) a general lack of clarity with respect to the construct of spatial ability (Caplan, MacPherson, & Tobin, 1985, 1986).

Many HCI experiments have controlled for sex differences or not analysed the effects of this variable. The Egan and Gomez (1985) experiments controlled for sex by using only female subjects. In the Vicente, Hayes and Williges (1987) experiment no correlation was found between sex and performance. Similarly, in their *study* of spreadsheet performance Garfein, Schaie, and Willis (1988) found no sex differences in performance. Mazlack (1980) found no sex difference in the performance of undergraduates upon an introductory programming course. Evans and Simkin (1989) report that sex accounted for significant variance in an examination relating to an entry-level business computing class, with females performing better than males. Borgman (1986), in a study of on-line information retrieval, found sex differences in the description of the mental models used of the retrieval environment. Men scored higher on a index of system description and level of model abstraction. With respect to performance, men were found to make significantly more errors for simple tasks but not for complex tasks. Borgman (1986) also presents evidence to suggest sex differences in the search methods used for information retrieval. Feltler (1985) found a significant computer literacy advantage for sixth and twelfth grade boys as compared to girls. However, boys were also found to have more contact with computers and to have more positive attitudes towards computers. This is consistent with the view of Clarke (1990) who, following a review of the literature, concludes that whilst there is little evidence of a sex difference in ability to use computers males use computers more frequently.

Fulton (1985) suggests that there are gender differences in language which extend to human-computer interaction. Fulton goes on to point out that computer dialogue is generated by computer programmers who are predominantly male. In order to examine this hypothesis a series of experiments were conducted in which the semantic differential technique (Jenkins, Russell, and Suci, 1958) was employed. Subjects were required to perform a computerised punctuation test which was presented with either masculine or feminine computer dialogue (as defined by Key, 1975). Results demonstrated that subjects were sensitive to the manipulation of computer dialogue style. Female subjects associated 'masculine' computer dialogue with greater potency. Male subjects rated the 'female' computer dialogue as being more potent than female subjects did. Whilst these findings provide interesting evidence of sex differences in response to different styles of computer dialogue, they do not necessarily demonstrate that the gender of the computer dialogue is the causal variable.

1.2.7 Personality and affect factors

The factor structure of personality has been the subject of ongoing controversy in much the same way as mentioned above with respect to psychometric theories of intelligence. Eysenck (1967), for example, proposes that the orthogonal dimensions of extroversion and neuroticism, with the later inclusion of a psychoticism factor (Eysenck and Eysenck, 1985), adequately explain the structure of personality. Eysenck's theory includes a physiological underpinning which relates to the level of autonomic arousal. Cattell (1957, 1972, 1973), however, contends that a more complex factor structure is required and identifies 16 personality factors, as measured by the widely cited 16PF test. Recent factor studies suggest something of a compromise between these two positions in the form of a five factor solution. These being: (i) extroversion or surgency, (ii) agreeableness, (iii) conscientiousness, (iv) emotional stability, and (v) intellect or openness to experience (Digman and Inouye, 1986; Digman, 1990; De Raad, Hendricks, and Hofstee, 1992). However, many of the studies which have examined personality as a predictor of computer-based performance have used the Myers-Briggs Type Indicator (MBTI: Myers and MaCaulley, 1985). This personality test has been extremely widely used, particularly within occupational settings (Murray, 1990). It is based upon Jung's (1971) typology, and categorises users upon the basis of the four continuous scale scores: Extroversion-Introversion, Thinking-Feeling, Sensing-Intuitive, and Judging-Perceiving. However, the validity of this measure has been criticised upon the basis that it has not been developed through factor analytic methods, and successful classification into Jungian types is difficult to demonstrate (Kline, 1993).

1.2.7.1 Myers-Briggs Type Indicator and Eysenck's Personality Inventory

The MBTI was used by Bush and Schkade (1985) to examine the profiles of computer programmers and analysts within a company setting. Upon the basis that a process of occupational selection had taken place these scores may be indicative of success within these professions. They report that the predominant cognitive types were Introverted, Sensing, Thinking, and Judging. A similar study of computer professionals, albeit with a larger sample spread over a number of companies, is reported by Lyons (1985). Results were broadly consistent with those of Bush and Schkade (1985) with a predominance of Introversion, Thinking, and Judging. The scores for the Sensing-Intuitive dimension were fairly evenly balanced. The importance of introversion (as measured using the EPI) as a predictor of programming performance was also identified by Kagan and Douthat (1985). A significant, although very modest, negative association was found between the E-I scale and grades achieved during a programming course. However, Chen and Vecchio (1992) found a similarly modest positive association between extroversion (MBTI) and one of four programming comprehension tests. Sitton and Chmelir (1984) administered the MBTI to a group of 'data operators' and found that they could be categorised as Extroverted, Intuitive, Thinking, and Perceiving. However, when compared to population norms the proportion of intuitive types was high whilst the proportion of extrovert types was low. Evans and Simkins (1989) examined the MBTI as a predictor of performance upon an 'entry-level business computing class'. A number of separate tests were used as dependent measures in a series of multiple regression analyses. Introverted and Judging types were associated with better performance for one of these tests, whilst there were mixed results for the Sensing-Intuitive and Thinking-Feeling scales.

The MBTI has been examined in a number of experiments in relation to performance using various computer applications. However, no clear pattern of results is apparent. Wong (1987) used the MBTI in an examination of different training methods for a multi-application package (including word processor and spreadsheet). The Thinking-Feeling dimension was associated with post workshop (demonstration and hands on training) word-processing performance, but not with performance following a computer-based tutorial or hard-copy documentation training methods. Feeling types exhibited the better performance. However, Ramurthy, King, and Premkumar (1992) found that Introverted, Sensing, and Thinking were the types associated with better performance when using a decision support system. The MBTI was included as a predictor in a study by Matta and Kern (1991) which compared spreadsheet learning following either interactive videodisk or classroom instruction. A negative association between extroversion and performance was found (introverts perform better). There was also a significant interaction between instruction mode and the Sensing-Intuitive scale such that there was a comparatively strong association with performance following videodisk instruction, with sensing individuals performing better than intuitive individuals, but little association following classroom instruction. Jennings et al. (1991) examined the Myers-Briggs as a predictor of information retrieval performance and found no significant association. Van Hoe, Poupeye, Vandierendonck, and De Soete (1990) found a modest positive correlation between extroversion and response time upon a menu search task. High neuroticism was also associated with slower responses in the early stage of task performance.

In summary, the MBTI has been investigated in relation to a range of computer-based tasks. Results have been mixed, and when the MBTI has been found to be significantly associated with performance it has only been at a modest level. The MBTI has been used to determine the typical 'personality profile' of computer programmers and computer operators. These individuals tend to be introverted, thinking, and judging types.

1.2.7.2 Locus of control

Measures of locus of control (Rotter, 1966) relate to the attribution of event outcomes. If an individual believes that the outcome of events which concern them is contingent upon their own actions they can be said to have internal locus of control. If however, they believe that outcome such events is contingent upon external factors then they have an external locus of control. In their study of menu search Van Hoe et al. (1990) found no significant association between locus of control and performance. However, in an information retrieval study, Gray, Barber, and Shasha (1991) found that subjects with internal locus of control (Rotter, 1966) performed more quickly. Similarly, Gray (1989) found a performance advantage for 'internals' in the use of computer-based training software.

1.2.7.3 Risk taking behaviour

In the management simulation study of Benbasat, Dexter, and Masulis (1981) risk taking behaviour was assessed (Kogan and Wallach, 1964). A tendency was found for subjects with higher risk scores to take less time to make decisions and to request less ordered history reports. Neal (1987) also reports a tendency for 'high risk' subjects to actively explore a help system, whilst 'low risk' individuals were often tentative in the process of command generation.

1.2.7.4 Computer anxiety

Egan and Gomez (1985) found that a measure of attitudes towards computers was not predictive of performance. Similarly, Vicente, Hayes, and Williges (1987) included a measure of computer anxiety in their study of information retrieval. This was not found to be significantly associated with performance. Will (1992) found no main effects of state anxiety (Spielberger, Gorsuch, Lushene, Vagg, and Jacobs, 1983) upon performance using an expert system. However, Chen and Vecchio (1992) found computer anxiety (using a scale adapted from the Mathematics Attitude Inventory: Sandman, 1979) to be negatively associated with scores on three out of four programming comprehension tests. Wong (1987) found a negative correlation between computer attitude and the acquisition of spreadsheet skills in a computer-based tutorial condition, but not for workshop or documentation conditions. A number of studies have examined sex differences in attitudes to computers and computer anxiety. The findings can be likened to those for spatial ability. A significant difference is not always apparent (Howard, 1986; Loyd and Gressard, 1986; Morrow, Prell and McElroy, 1986; Parasuraman and Igbaria, 1990; Ray and

Minch, 1990), but when it is present then it is such that males have more positive attitudes to computers and are less computer anxious (Raub, 1981; Dambrot, Watkins-Malek, Silling, Marshall, and Garver, 1985; Gilroy and Desai, 1986; Dukes, Discenza, and Cougar, 1989; Igbaria and Chakrabarti, 1990).

1.2.8 Summary of assay

Verbal ability, logical reasoning, and spatial ability have all been found to be predictive of computer based performance. Whilst there are grounds to suggest that verbal ability may interact with interface factors, and there is uncertainty as to the task specificity of the association with logical reasoning, spatial ability has consistently been associated with the performance of a range of computer-based tasks. It would appear that spatial visualisation (generally as assessed by VZ2) may be the stronger predictor of information retrieval performance (Vicente, Hayes, and Williges, 1987), whilst spatial memory (MV2) may be the stronger predictor of word processing performance (Gomez, Egan, Wheeler, Sharma and Gruchacz, 1983). Short-term memory has received comparatively little attention in this area, but there would appear to an association between working memory and tasks demanding reasoning, such as programming. Psychometric tests of associative memory have also been found to be modestly related to information retrieval performance, but not word processing performance. A number of related aptitudes (e.g. academic ability, and technical ability) have also been found to be predictive of performance, primarily in relation to information retrieval.

Field dependence appears to be modestly predictive of many aspects of computer-based performance, however, the validity of this construct has been criticised. The relationship between FI-FD and computer based performance is consistent with an ability related interpretation of the GEFT, with high scores being associated with better performance. Studies which have included both FI-FD and spatial ability as predictors (Vicente et al., 1987; Jennings, et al., 1991) have found spatial ability to be more strongly associated with performance. The results with respect to the Learning Style Inventory (Kolb, 1971) are somewhat mixed. However, it would appear that learning style may interact with the type of system model (analogical vs abstract) presented to the user.

A number of differences in the mental representation of information have been found to differentiate novice and expert performance. Whilst these studies primarily relate to the task of programming, differences in strategy have also been demonstrated in relation to a number of computer tasks. However, within an information retrieval setting, spatial ability has been found to account for almost all of the variance associated with expertise (Vicente et al., 1987; Vicente and Williges, 1988). Age has been found to be a consistently strong predictor of word processing and spreadsheet performance, although this applies primarily to the speed of performance. The pattern of results with respect to error data is mixed and will be discussed further in Chapter 2. Sex differences in computer-based performance have received comparatively little research attention. There is some evidence to support a small male superiority when performing tasks which require spatial ability. However, there is little to suggest that this advantage transfers to computer-based tasks.

The MBTI has been used in a variety of experiments. Generally only modest correlations with performance have been obtained. The pattern of results would suggest that introverted, thinking, and judging types tend to exhibit better computer-based performance.

1.3 The Research Plan

Whilst some patterns are beginning to emerge with respect to which individual differences are particularly important in the prediction of computer-based performance, there is little understanding of the facets of the interface or the tasks which are particularly sensitive to such variance. Very little focused research attention has been given to identifying such interactive effects. This means that there are no clear guidelines which can be applied to the process of interface design (Potosnak, Hayes, Rosson, Scheider, and Whiteside, 1986). As an eventual ideal position, Van Der Veer, Tauber, Waern, and Muyiwijk (1985) suggest a set of design rules which stipulate that "...IF user (i,j...) AND task (k,l....) THEN apply design principles (m,n...)". This research project was concerned to make progress towards this ultimate goal. Subsequent chapters present experiments which examine various task and interface factors in relation to predictors of demonstrated worth (see above). They are concerned with the 'isolation' phase of the Egan and Gomez (1985) experimental methodology. Many of the studies mentioned above will be reported in greater detail, with an appropriate change of emphasis.

1.3.1 Word-processing skills

Chapter 2 is concerned with individual differences in word processing performance. In particular it investigates the performance of a number of component elements of word processing along with the performance of a complete task. This is related to various measures of cognitive ability and to age. Unlike previous studies of individual differences in word processing, expert users are examined to see if the predictors which have previously been established (e.g. Egan and Gomez, 1985; Czaja, Joyce, and Hammond, 1989) are associated with skilled performance.

1.3.2 The 'finding' component

Chapters 3 and 4 examine what may be defined as the 'finding' component of computer task performance. Egan and Gomez (1985) identified this component in relation to word processing, however it can equally be applied to many other computer-based tasks. The experimental task used in Chapter 3 involves the location of edits within computer-presented text upon the basis of a marked 'hardcopy'. Interactive effects of interface dimensions, such as window size and the formatting of the text, and individual differences in cognitive ability are considered. Chapter 4 examines individual differences in the navigation of networks and menus. These tasks are prevalent in a vast range of computer-based applications, particularly those related to information retrieval. This chapter considers the mental representation of the task environment and relates this particularly to differences in spatial and verbal ability.

1.3.3 The 'generating' component

Chapters 5 and 6 are concerned with the process of command generation. Chapter 5 examines individual differences relating to the use of a menu or command line interface. A number of sources suggest that novices will benefit from menu interfaces whilst experts benefit from command line interface (Shneiderman, 1987a, 1988). However, studies to date have failed to support this principle. This chapter re-examines these interactive effects in the context of command syntax complexity and also considers a number of cognitive ability measures. Chapter 6 also considers individual differences in command generation, but in relation to the use of either icons or text. This is an area which has received little attention. A start is made in this direction using a task in which subjects are required to search and select matching stimuli.

1.3.4 Information retrieval skills

Chapter 7 is the final experimental chapter, and examines individual differences in information retrieval performance. In particular, four different data structures are examined. Age and a range of cognitive ability tests are examined as predictors of performance.

Chapter 8 provides a summary of the main findings of the previous chapters. The implications of these findings in relation to the process of interface design, selection, and training are considered, and suggestions for future research are made.

1.3.5 Strategy differences

In addition to examining the relationship between individual differences in ability and performance levels, individual differences in strategy selection will also be considered in some experiments. As discussed earlier, many interfaces present the user with a number of performance options which achieve the same goal. Consequently, individual differences in performance may not be due to ability, but may be due to either differences in strategy selection or a complex interaction between ability and strategy. Existing research does not adequately address the question of the efficiency with which ability and strategy are matched.

1.3.6 Workload

It has been demonstrated that differences in cognitive performance may be masked by differences in attentional resource allocation (Gopher and Donchin, 1986). It is possible for individuals to achieve similar levels of performance with very different levels of invested effort. Consequently the study of individual differences in performance may be undervalued, and its implications for selection, interface design, training, and levels of stress experienced by computer users (Hockey, Briner, Tattershall and Wiethoff, 1989) given insufficient weight. Currently there is very little research which specifically examines the attentional resource demands of non-specialist computer-based tasks.

Most of the experiments reported in this thesis examined self-report workload using the NASA TLX (Vidulich and Tsang, 1986). This method of workload measurement was selected for two main reasons. Firstly self-report measures of workload have been found to be sensitive to differences in spatial processing demand (Eggemeier and Stadler, 1984). As spatial ability has been demonstrated to be predictive of computer-based performance this was an important factor. Secondly, this measure of workload measurement is comparatively unobtrusive. If this method of workload assessment were to prove useful it might easily be applied to the process of usability testing or a variety of other experimental designs with no major task disruption and no additional software programming overhead.

1.3.7 Measures of Individual Differences

As mentioned, these experiments will generally use predictors of proven worth. Rather than repeat detailed descriptions in each chapter these measures are described here. Not all measures were used in all experiments.

1. The Building Memory Test (MV2 - Ekstrom et al., 1976). This test of spatial memory has been found to be predictive of performance upon text editing tasks (Egan and Gomez, 1985; Czaja, Joyce, and Hammond, 1989). In the hierarchical model of Carroll (1993) it loads upon second order factors of spatial ability and general memory ability. This test presents subjects with a street map showing the location of a number of buildings. After a study period of four minutes subjects are presented with the same street map but with the buildings omitted. Areas within the map are marked with letters. Subjects are required to indicate location of each building by selecting the letter which is placed in the area of each building using a multiple choice response format.

2. The Paper Folding Test (VZ2 - Ekstrom et al, 1976). This test of spatial visualisation has been found to be predictive of performance upon information retrieval tasks (Vicente, Hayes, and Williges, 1987; Vicente and Williges, 1988; Campagnoni and Erlich, 1989; Seagull and Walker, 1992). It loads upon a second order factor of visual perception (Carroll, 1993). Each item in this test presents subjects with a series of diagrams indicating the folding of a piece of paper. The final diagram shows a hole being punched through all thicknesses of the folded paper. The response requires subjects to select, from five alternatives, a diagram illustrating the positions of the holes when the paper is unfolded.

3. Logical Reasoning (RL2 : Ekstrom et al., 1976). This test has been found to be predictive of word processing (Egan and Gomez (1985; Czaja, Joyce and Hammond, 1989) and the use of a database query language (Greene et al., 1986; Greene et al., 1990). It loads upon a second order factor of fluid intelligence (Carroll, 1993). Each item presents subjects with the names of three objects. Subjects are then required to select from five alternatives a Venn diagram which correctly describes the interrelationship of these objects.

4. Associative Memory (MA2 : Ekstrom et al., 1976). This test has been found to be predictive of menu navigation performance (Billingsley, 1982). It loads upon a second-order general memory factor (Carroll, 1993). Subjects are presented with a list of word-number pairs. Following a period of memorisation, subjects turn to a test

page upon which the objects are listed in a different sequence. They are required to recall the number with which each object was originally paired.

5. The Nelson-Denny Reading Test (Vocabulary section). The short form (11 mins.) of this test was administered. This test has been found to be predictive of performance of both information retrieval (Vicente, Hayes, and Williges, 1987) and text editing performance (Egan and Gomez, 1985). The vocabulary section of this test loads upon the second-order factor of verbal ability in Carroll's (1993) hierarchical model of cognitive ability. Each of the 100 items in this test present subjects with a definition followed by five alternative words from which to select a response.

6. The computer literacy sub-test of the CALIP (Computer Aptitude, Literacy, and Interest Profile; Poplin, Drew, and Gable, 1984). This has been shown to distinguish groups on the basis of computer experience. Levinson (1986) offers a critical appraisal of the CALIP in which he makes several criticisms of its structure. These criticisms, however, are confined to the Aptitude sub-tests (e.g. overly high inter-correlation of sub-tests). The Literacy section of the CALIP consists of 30 multiple choice items. The first 20 items require one answer to be selected from five alternatives, whilst the last 10 items require the selection of two responses from five alternatives. Items are placed in order of ascending difficulty. Poplin et al. (1984) recognise that the content of this section of the CALIP is necessarily somewhat limited, given the restricted number of items, but suggest that in conjunction with a further survey relating to type of computer experience an appraisal of computer learning may be realistically achieved. When examining the reliability of the CALIP Poplin et al. (1984) used four age groups: 12-14; 15-19; 20-29; and 30-60. Split half reliability for the Literacy section equalled or exceeded .90 for each of the age groups with the exception of the youngest for which it was .76. Test-retest reliability for the Literacy section was .82 ($n= 86$; $p<.0001$). Scores on the Literacy sub-test were found to increase with age, and there were significant sex differences ($p<.001$) when the effects of age were controlled, with males scoring higher. When controlling for the effects of age, the Literacy sub-test showed a significant difference in scores between expert programmers and random others ($p<.001$), between those who know two or more computer languages vs those who know one or none ($p<.001$), between those who have written intermediate / advanced programs vs others ($p<.001$), and between those who have taken two or more computer courses vs those who have taken one or no computer courses ($p<.001$).

1.3.8 Computer programming

With the exception of the experiment reported in Chapter 6, all experiments were programmed in C and Assembler (without the use of proprietary toolkits), which allowed flexibility and precision in experimental program design, and minimised routine execution times. During each experimental trial, care was taken to avoid the timing delay associated with hard disk access, and all necessary program data was stored in RAM. Improved screen display response times were achieved by using memory move functions which directly accessed the video RAM. Timing resolution is sometimes a problem on the IBM PC, with the timer interrupt provided by the operating system occurring at 18.2 times per second. Some of the experimental tasks used in this research required comparatively brief presentation and / or response intervals. It was therefore necessary to improve the timing resolution by modifying an Assembler routine written by Sheppard (1987), to allow repeated calls, and incorporating this routine into all programs. Screen presentation timing was latched to the screen refresh cycle in order to control for this additional source of timing error (cf. Tischer, 1989). As a result, given a consistent placement of stimuli upon the CRT, the variance associated with the screen refresh becomes a constant. These measures provided an eventual timing accuracy of ± 5 ms, with the remaining error relating to the priority given to the keyboard interrupt by the operating system.

Chapter 2

Word Processing Skills

2.1 Introduction

This chapter presents an empirical study of individual differences in word processing. Using an 'expert' sample, various cognitive abilities and age were examined in relation to the performance of a complete editing task, and also a number of component tasks. This experimental design allowed task components which are particularly sensitive to the effects of individual differences to be identified. This introductory section presents a review of studies which have examined individual differences in word processing.

Word processing, or text editing, is one of the most frequently performed and frequently studied computer-based tasks (Roberts, 1980). Experimental programs have been devised to investigate methods of evaluating word processors (Roberts, 1980; Roberts and Moran, 1983; Borenstein, 1985; Czaja, Joyce, and Hammond, 1989), the acquisition of word processing skills (Mack, Lewis, and Carroll, 1983; Ross and Moran, 1983; Carroll and Mack, 1985; Frese, Albrecht, Altman, Lang, Papstein, Peyerl, Prumper, Schulte-Gocking, Wankmuller, and Wendel, 1988), models of user performance (Card, Moran, and Newell, 1983; Keiras and Polson, 1985; Polson and Keiras, 1985), and the selection of command names (Barnard, Hammond, McLean, and Morton, 1982; Grudin and Barnard, 1985). Individual differences in performance have also been investigated in a number of studies, and several promising predictors have been identified.

In a widely cited series of experiments Egan and Gomez (1985: see also Egan, Bowers, and Gomez, 1982; Egan and Gomez, 1982; Gomez, Egan, Wheeler, Sharma, and Gruchacz, 1983; Gomez, Egan, and Bowers, 1986) attempted to assay, isolate and accommodate individual differences in text-editing performance. Subjects were female computer novices, aged 26-63 years, and were required to learn how to make revisions to a marked 'hardcopy' using a text-editor. Dependent measures for computer-based performance were reading time (time spent reading instruction manual), execution time per successful change, and first-try errors. Over the series of four experiments the predictor variables examined included age, spatial memory (MV2; Ekstrom, French and Harman, 1976), spatial visualisation (Two Dimensional Space Test), associational fluency (FA1; Ekstrom et al., 1976), verbal ability (Nelson-Denny Reading Test, 1973), logical reasoning (LR2; Ekstrom et al., 1976), associative memory (MA2; Ekstrom et al., 1976), attitude towards computers, knowledge of text editing terms, knowledge of / desire to use computers, and estimated typing speed.

The first two of these experiments used a line editor, and found spatial memory and age to be consistently strong predictors of speed and accuracy of performance. Following stepwise multiple regression analyses in which neither logical reasoning nor associative memory could account for the variance associated with spatial memory, Egan and Gomez (1985) concluded that the correlation between spatial memory and performance was not an artefact of a more general intelligence factor or a more general memory factor. Verbal ability was generally only predictive of the reading time element of task performance. In the second of these experiments a more detailed analysis of error data was undertaken. Age was found to be significantly correlated with errors of omission ("One or more elements of a command missing", p. 191) and multiple errors ("A combination of two or more errors", p. 191), whilst spatial memory was found to be significantly correlated with pattern errors ("The order or spacing of command elements is wrong", p. 191) and multiple errors.

During the isolating phase of these experiments (the third experiment) Egan and Gomez (1985) decomposed the task of text editing into the constituent processes of 'finding' (locating edits), 'counting' (counting the number of lines to the next edit), and 'generating' (generating the sequence of commands to make the edit), and examined subject performance upon pencil and paper versions of these component elements. Spatial memory was found to be consistently predictive of both 'finding' and 'generating' components of task performance, whilst age was only associated with the process of command generation.

For the accommodation phase of these experiments (the fourth experiment) Egan and Gomez (1985) used a display editor as opposed to the line editor used in experiments one and two. It was hypothesised that this would reduce the predictive power of both spatial memory and age by virtue of increased spatial information and reduced command complexity. The effect of age was much reduced, accounting for less variance in performance time and showing a slight negative association with errors. Spatial memory, however, remained predictive of both speed and accuracy of performance. Additionally logical reasoning was found to be predictive of execution time and verbal ability was found to be the strongest predictor for all dependent measures. A measure of spatial visualisation did not account for unique variance within the regression and it was therefore concluded that the effects of spatial memory could not be attributed to a more general spatial ability factor.

Czaja, Joyce, & Hammond (1989) confirmed many of Egan and Gomez's (1985) results in an experiment which examined the performance of a group of computer naive women (aged 40-70 years) whilst using three different word processors.

Experimental tasks included typing a letter, and making amendments to existing documents using a range of text editing commands (including insertion, deletion, boldfacing, and moving text). Consistent with Egan and Gomez (1985), spatial memory (MV2; Ekstrom et al., 1976), logical reasoning (LR2; Ekstrom et al., 1976) and age were all found to be significantly correlated with both task completion times and accuracy of performance for both typing and editing tasks, with spatial memory being the most strongly related. In an experiment concerned with the effects of age and different training regimes, Czaja, Hammond, Blascovich, and Swede (1989) studied the performance of three age groups (25-39, 40-54, and 55-70 years). Tasks were performed on a display editor and included the typing and editing of memos, letters, reports and address lists. There was a significant effect of age upon time per task with performance of the oldest group being significantly worse than that of the youngest group. The commission of errors followed a similar pattern. However further analyses revealed that there was no significant age effect upon spelling or structural errors (concerned with the mechanics of editing), but there was a significant effect of age upon the commission of format errors (concerned with the layout of the page). There were no interactive effects of age with training regime.

The importance of the association between spatial memory and word processing performance was not confirmed by Sebrechts, Deck, Wagner, and Black (1984). The predictors used in this experiment included typing ability (as measured with a typing test), the Computer Operator Aptitude Battery (Science Research Associates, 1973 - this test includes three sub-parts - sequence recognition, format checking, and logical thinking), spatial memory (MV2; Ekstrom et al., 1976), associative memory (MA1; Ekstrom et al., 1976), and verbal ability (Nelson-Denny Reading Test, 1973). Subjects (n=24: word processing novices of unspecified age although mostly students), attended five experimental sessions, during which time they were instructed in a wide range of word processing skills. The dependent measure was "a composite of the number of modifications made" (p. 235) whilst editing two multi-page documents. Consistent with the results of Egan and Gomez (1985) with respect to the use of a display editor, the Nelson-Denny Reading Test was the predictor most strongly correlated with performance, although significant correlations were also obtained for the sequence recognition and format checking sub-tests of the Computer Aptitude Battery. Spatial memory, however, was not a significant predictor of performance.

Whilst a positive association between age and response time is a consistent finding, the relationship between age and the commission of errors appears to be more complex. In an examination of the effects of age upon word processing performance, Hartley, Hartley, and Johnson (1984) found that after an initial training period there were no differences between younger subjects (aged 18-30 years) and older subjects (aged 65-75 years) in the recall of information about the editor or "in the correctness and efficiency with which computer operations were carried out". Older subjects, however, did take more time to perform certain components of the editing task, and required more assistance. Charness, Schumann, and Boritz (1992) conducted two experiments which investigated the performance of younger (under 40 years) and older (over 50 years) subjects whilst using a word processing tutorial. The effects of the provision of an advance organisor, and self-paced vs fixed-paced tutorial conditions were examined. Age did not interact with training condition in either experiment. In keeping with the results of Hartley et al., older subjects were slower to complete the self-paced tutorial conditions, and requested more assistance. However, in constrast, older subjects also performed more poorly upon a post tutorial test of knowledge of word processor functions.

As with the study of Czaja, Hammond, Blascovich, and Swede (1989), mentioned above, Elias, Elias, Robbins, and Gage (1987) found age to be related only to specific types of error. This experiment studied age differences in response to a training program for word processing. Subjects were recruited to three age groups (18-28 ; 37-48 ; and 55-67 years), and were required to reach a minimal typing level at the first training session. Performance was monitored over seven training sessions. Significant effects of age were found with respect to task completion times in all but the first training session. No significant age effects were found for overall errors. However, when errors of omission, commission, or command were analysed separately for each training session, older subjects were found to make more errors of commission in certain sessions. Consistent with the findings of Charness et al. (1992), in a test of knowledge of word processing techniques older subjects were found to perform significantly worse.

In an experiment by Davies, Wong, Glendon, Stammers, and Matthews (1993) an age-related speed accuracy trade-off was found. The performance of 30 younger (18-30) and 30 older (45-60) women who were experienced typists, but novice word processor users was compared. In addition to the effects of age, the relationship between word processing performance and spatial memory, spatial scanning, spatial visualisation, vocabulary, perceptual closure, number comparison, paired associate memory, and working memory were considered. Performance was monitored over five trials, during which subjects were required to edit computer presented files on the basis of marked 'hardcopies' of the text. The effects of age appeared, in part, to take the form of a speed accuracy trade off, with older subjects performing more slowly but more accurately. However, there was also a significant age by trials interaction for both speed and accuracy, with the older group showing a greater performance improvement over time for both dependent measures. Response time differences appeared to be the result of increased time per keystroke on the part of the older groups, as opposed to the use of less efficient editing strategies. Following the Egan and Gomez (1985) strategy for task decomposition, four component elements of performance were examined: text finding; text generating; menu finding; and menu generating. There was no significant interaction between age and the relative time spent performing each component element. With respect to individual differences in cognitive ability, there was a modest correlation between vocabulary scores and accuracy but only for the younger group. All tests relating to spatial ability were strongly correlated with both speed and accuracy of performance, with subjects high in spatial ability performing more quickly and producing fewer errors.

From the results of these studies, it would appear that age is consistently predictive of word processing performance. Whether this relationship applies to both the speed and accuracy of performance is less clear. All the above studies have found performance times to decline with age. However, whilst some studies have found performance accuracy to follow a similar pattern (e.g. Czaja, Joyce, and Hammond, 1989) others have found either no difference (e.g. Hartley et al., 1984), or a speed accuracy trade off with older subjects apparently substituting reduced speed for increased accuracy (Davies et al., 1993). Whilst there is evidence from studies of other computer-based tasks to support an age-related maintenance of performance accuracy (Garfein, Schaie, and Willis, 1988), these diverse results may relate to the type of errors which are being considered. The studies of Egan and Gomez (1985), Czaja et al. (1989), and Elias et al. (1987) all indicate that age is only related to certain categories of error. There is no clear pattern as to the types of error which are particularly age sensitive, however this may reflect differences in the methods used to categorise errors.

With respect to cognitive ability, most of the above studies have found spatial ability, and particularly spatial memory, to be predictive of performance. This appears to be consistent across different types of editor (Egan and Gomez, 1985), although Sebrechts et al.(1984) found no significant correlation between spatial memory and performance in an experiment which used a display editor. However, the importance of verbal ability (as assessed by tests of vocabulary) appears to be related to editor design. Whilst Egan and Gomez (1985) found no strong relationship between verbal ability and performance whilst using a line editor, both Egan and Gomez (1985) and Sebrechts et al. (1984) found a significant relationship between verbal ability and the use of a display editor. The results of Davies et al. (1993) suggest that this may also interact with age. Logical reasoning also appears to be associated with speed of performance whilst using a display editor.

2.1.1 Experimental aims

Despite the convergence of these studies in a number of areas relating to the prediction of word processing performance, many areas remain unresolved. One such question concerns the relationship between individual differences in ability and the component parts of the word processing task. If it is possible to establish which individual differences are most strongly related to which word processing components then this would enable a focused approach to be applied to the processes of training and interface design. Individual training programs could be devised such that different phases within the training regime matched the cognitive capabilities of the individual, maximising strengths and minimising weaknesses. Similarly, this information could be used to enable interface characteristics to be tailored to the individual. It could provide interface designers with valuable data as to the relative importance of various sub-tasks which could then be used to construct appropriate adaptive or adaptable elements within the interface. Egan and Gomez (1985) began to address this issue in their third experiment, which attempted to isolate the effects of individual differences in ability with respect to the 'finding', 'counting', and 'generating' components of task performance. However, it would appear that there are a number of shortcomings in their approach to the problem. Firstly, the tasks used for this experiment were paper and pencil exercises. Consequently, despite the moderate correlations which were found with the computer-based performance of their first two experiments, validity is an issue. Secondly, the finding component which was used in this experiment presented subjects with the same text, typed on opposite sides of a piece of paper. On one side of the paper words (edits) were circled in ink. Subjects were required to locate the same words on the reverse side of the paper. However, the formatting of the text on the two sides was different. The abilities required in the performance of this task cannot therefore be held to generalise to display-based text editing where hardcopy and screen based text are presented in the same format (WYSIWYG). Thirdly, there was obviously no cursor movement requirement, which may be one of the most fundamental components of the process of target word location within a display based editor. Further to this, the counting section of this experiment is not relevant to display based text editing. Finally, in the generating section of this experiment subjects were presented with a number of short phrases which required the use of the substitution command. Whilst the formatting of the command varied between trials, no command selection process was required. The method of command generation was similar to that used in a line editor, and dissimilar to the direct manipulation of text performed in a display based editor.

The present experiment attempts to build upon some of these issues, and examine in more detail the relationship between individual differences and the various sub-components of word processing performance. A process of task analysis (Card, Moran, and Newell, 1983) was used to decompose two 'core' word processing tasks (Roberts, 1980) into a number of component elements. These 'core' tasks involved the insertion and deletion of words within passages of text. The component elements were performed on a computer, and included: (i) two tasks designed to examine the process of matching hardcopy edit location to the screen edit location; (ii) a task concerned with the positioning of the cursor at the appropriate word; (iii) a task related to the process of command decision; and (iv) tasks related to the process of generating insertion and deletion command keystrokes. The relationship between individual differences and the performance of each of these component elements was then examined along with the performance of a complete text editing task. On the basis of the findings of Egan and Gomez (1985) it could be predicted that spatial memory would be related to all component elements of performance, whilst age would only be related to those components involved in the process of command generation. Additionally, as the current experiment used a display editor, it could be predicted that verbal ability would be related to performance, although the relationship with component elements was unclear.

Of particular interest in this process of analysis was the relationship between subject age and the performance of component and complete tasks. As described above, there is evidence to suggest that age differences in word processing performance may be most substantial when subjects are unfamiliar with the task and that these differences are reduced with practice (Davies et al., 1993). A potential explanation for this effect involves a process of 'compensation' (Dixon, 1992), by which older individuals compensate for the effects of specific age-related deficiencies by deploying other cognitive abilities at which they are more able. One method which has been used to examine this hypothesis is that of molar equivalence-molecular decomposition (Charness, 1981a, 1981b; Salthouse, 1984). This involves comparing the component task performance of individuals of different ages who are of similar molar task ability. If some or all of the molecular tasks show age related deficits then this may indicate a process of compensation. Such information enables deductions to be made as to which particular components are age sensitive, or how task components are combined to produce equivalent molar performance (Charness and

Bosman, 1990). It is possible that through extensive practice, and a process of compilation (Anderson, 1983), that whilst age differences are apparent in molecular task performance, molar task performance is maintained. This would be consistent with the suggestion of Hasher and Zacks (1979) that automatic processing is insensitive to the effects of age. However, following a review of the relevant literature, Salthouse (1991c) concludes that there is limited support for this position. A particularly relevant example of the use of the molar equivalence-molecular decomposition technique is provided by Salthouse (1984) who examined age differences in typing skill. In two studies, using skilled typists, a negative correlation was found between choice reaction time and age but no association between age and typing speed. Using an experimental design which controlled the number of characters visible to the subject Salthouse (1984) was able to demonstrate that older typists 'looked ahead' further (in Salthouse's terms they had a larger 'eye-hand span'). Salthouse (1984, p. 369) suggests that this enabled them to engage in a process of compensation by "...planning further ahead, in effect scheduling around bottlenecks in the system...".

In order to be able to use such an experimental method it is necessary that subjects are experienced in the molar task. Obviously, if individuals are unfamiliar with a task they will be less able to assess the associated cognitive demands and therefore less able to engage in a process of compensation. Consequently, such experiments require that either a large number of repeated trials are performed, or that experienced subjects are recruited. In this experiment the latter alternative was employed in order that such analyses could be conducted. Whilst subjects were initially recruited regardless of molar task ability, for certain analyses a sub-sample was used in which the molar performance of younger and older subjects was comparable. It could be predicted that a process of compensation would result in older subjects performing specific task elements with greater efficiency in order to negate the effect of reduced efficiency in other components.

The use of experienced subjects also allowed an interesting comparison with previous research which had exclusively focused upon the performance of novice word processor users. Certain theories of skill acquisition predict that cognitive ability is relatively influential in the early stages of skill acquisition but that this influence is reduced as further practice takes place (Ackerman, 1988). Obviously, such a relationship would have important implications for the use of such predictors in relation to computer-based performance. For example, the application of current research findings in this area to the process of selection would be limited, and the use of adaptive interfaces to accommodate such individual differences would have to incorporate sufficient flexibility to ensure that the effects of practice could be accommodated. Attention might be most usefully focused upon training regimes, and situations where computers are subject to relatively infrequent use by occasional or novice users. Conversely, if it were established that certain cognitive abilities are important predictors of performance regardless of expertise this would strengthen the case for the application of this research in processes of selection and lasting elements of interface design.

2.2 Method

2.2.1 Subjects

64 subjects were recruited from the students and staff of Aston university and from the local community. 32 were aged between 18 and 30 years (mean 20.25 years), and 32 were aged 45 years or older (mean 50.19 years). Equal numbers of males and females were recruited to all experimental conditions. All subjects had over 100 hours of experience in using a word processor, were right handed, reported vision which was normal or corrected to normal, and spoke English as their first language.

2.2.2 Measures of individual differences

Subjects initially completed tests of spatial visualisation (VZ2; Ekstrom et al., 1976), spatial memory (MV2; Ekstrom et al., 1976), verbal ability (Nelson-Denny vocabulary test, 1973), logical reasoning (LR2; Ekstrom et al, 1976), and computer literacy (CALIP; Poplin, Drew, and Gable, 1984). In addition, a short test of typing speed was completed. The typing test required subjects to type a series of three word sentences. Each trial began with a sentence being displayed upon the CRT for memorisation. Subjects were then required to press the spacebar at which point the sentence was removed from the screen and the timer started. Subjects then typed the sentence as quickly as possible and pressed the ENTER key to signal that they had finished, at which point the timer stopped. 16 trials were presented, of which the final eight were recorded. All words were three and four letters long and controlled for frequency using the same criteria as those selected for editing in the component and molar tasks (see below).

2.2.3 Experimental task overview

Computer-based tasks were presented on a PCAT 80286 computer with a Hercules screen. Software was purpose written in C and Assembler, timing was ± 5ms., text mode was used with a maximum display of 25 X 80 characters.

The molar task was concerned with the performance of basic text editing skills. Two simple, 'core' (Roberts, 1980) word processing tasks were selected, these being the insertion and deletion of words from within text. However, the performance of such tasks can also be seen in a variety of other computer applications. Following a process of task analysis (Card, Moran, and Newell, 1983) seven component or molecular task elements were designed. Some compromise was inevitable at this stage in balancing the conflicting demands of a thorough task decomposition with the need to retain workable molecular components. Given the number of component tasks and the requirement for the random presentation of text passages within each, a partially counterbalanced sequence of presentation was adopted which approximated the order in which component tasks would occur within the molar task. With all subjects being skilled word processor users this was seen as an acceptable compromise. However, the molar task was always performed last in order that strategy selection should not be influenced by unequal amounts of practice upon the component elements. The texts used for all tasks were drawn from the Aston University Library Business Magazine Database, and were matched for reading ease as, far as possible, using the Flesch formula (Flesch, 1949). Separate texts were used for each trial, and were presented in a random sequence. Subjects completed the NASA TLX measure of self-report workload following each experimental condition.

2.2.4 Component tasks

The molar tasks were decomposed into seven component tasks. Three of these were associated with the process of 'finding' targets, and four were associated with the process of 'generating' commands (cf. Egan and Gomez, 1985).

2.2.4.1 Coarse search (page task and scroll task)

These tasks were designed to represent that portion of the molar task which is concerned with the initial visual location of the target word including an orientation between the 'hardcopy' and the computer screen.

Subjects were presented with a 'hardcopy' of a piece of text upon which a number of words had been circled in red ink. Target words were all four letters long, contained one syllable, were each used only once as a target, and occur in normal language with a frequency greater than 50 times per million (Thorndike and Lorge, 1972). Exactly the same text was presented on the computer screen. Subjects were required to locate visually the target words marked on the 'hardcopy' within the text displayed upon the CRT. Two counterbalanced conditions were used, one in which subjects could use only the up and down scroll keys, and one in which subjects could use only the page keys. Subjects were not required to place the cursor at the exact position of the edit but merely to display the edit upon the screen. They then pressed the ENTER key to signal that they had found the edit, at which point the target word was highlighted on the screen. If they missed the target a message to that effect was displayed at the bottom of the screen and the computer 'beeped'. A practice piece of text (in which 16 edits were marked over 6 pages) preceded each condition. Performance was measured for 32 edits marked over a 12-page document. Of these 32 edits, 16 were designated as short edits and were less than 20 lines apart, and 16 were designated as long edits and were 21 lines or greater from the previous edit. Of these long and short edit conditions equal numbers were designated as inner edits (contained within the body of the paragraph of text), and outer edits (contained within the top or bottom lines of the paragraph). A window of 21 lines was used. The page key moved the text up or down by 20 lines (one less than the size of the window).

2.2.4.2 Fine search

This task was designed to represent the fine movement of the cursor to the position of a word which has already been located by a process of coarse search.

Subjects were presented with a screenful of text which included a highlighted (reverse video) four letter word. Target words conformed to the same criteria as described above. The cursor was initially positioned in the upper left hand corner of the screen. Subjects were instructed to first visually locate the target word and then to press the ENTER key at which point the timer was started. They were then required to move the cursor to this target word using the up, down, left and right cursor keys. Once the cursor was located upon any one of the letters of the target word the subject pressed the ENTER key to signal that they had finished, at which point the timer stopped. Subjects were presented with eight practice trials, followed by 16 performance trials. Eight of the performance targets were 'inner' targets (within the body of a paragraph) and eight were 'outer' targets (at the outer edges of a paragraph).

2.2.4.3 Decision task

This task was designed to represent that portion of the molar task which is concerned with deciding which command should be generated on the basis of the symbols used upon the marked hardcopy. This was the most difficult task element to operationalise, but it was felt that it represented an important component of molar task performance. In order to facilitate timing, all elements of this task were computer presented.

Symbols were used on the computer screen to represent the insert and delete symbols used on the hardcopy. A down arrow was used to represent INSERT whilst a cross was used to represent DELETE. Subjects were presented with a text description of one of the tasks ('insert' or 'delete'). The screen was then cleared and subjects were presented with a row of insert and delete symbols (max. 12 : min. 6). Subjects were required to count the number of symbols which represented the target command and to press the ENTER key at which point the screen cleared, the timer stopped, and they were prompted to enter the number of symbols counted into the computer. Subjects were presented with eight practice trials followed by 16 performance trials.

2.2.4.4 Insert, Delete and Backspace Tasks

These tasks were designed to represent those portions of the molar task concerned with command execution. They required subjects to make the final fine adjustments to the cursor position and to insert or delete a specified word. It was not practicable nor appropriate to separate the final adjustments of cursor position from the processes of insertion or deletion because alternative cursor placements are possible, and in some cases required, for each task.

Subjects were presented with a screenful of text. In a three-line box at the bottom of the screen subjects were instructed as to which word should be inserted/deleted and the position of the insert/deletion. The cursor was always positioned on the same line as the required edit and no more than two spaces from the required starting position. This was determined by a pseudo random sequence. On-screen instructions stated that the subject should locate and plan the required edit before pressing the ENTER key, at which point the time would start. This was in order that the reading time and task comprehension elements which were peculiar to these component tasks were not included. Having successfully completed the edit, subjects were again required to press the ENTER key at which point the time stopped and the screen cleared prior to the presentation of the next trial.

The insert and delete/backspace conditions were presented in a counterbalanced sequence. In each condition subjects were presented with eight practice trials followed by 16 trials during which performance was recorded. All words used in the insert condition were four letters long and were selected on the basis of the previously stated criteria.

The delete/backspace condition required the use of both the delete key (deleting at the position of the cursor and leaving the cursor in the same place after each command) and the backspace key (deleting the letter to the left of the cursor position and moving the cursor one place to the left during the execution of each command). At the beginning of each trial, on-screen instructions stated which command should be used. Of the 16 recorded trials eight required the use of the delete key and eight required the use of the backspace key, with the sequence of presentation being random. The words used in this condition were either three or five letters long presented in random sequence in order to prevent subjects from making deletions by counting the number of key presses.

2.2.5 Complete editing task

As mentioned above, the molar tasks performed in this experiment required the insertion and deletion of single words within a passage of text. Subjects were presented with a passage of text which had been marked with a number of insertions and deletions. They were required to make these edits using whichever command methods they preferred. All of the command options contained within the component tasks were available to them. Subjects were initially presented with a practice piece of text which was six pages long and contained 16 edits. Following this they were timed during the editing of a 12 page document which contained 32 edits.

2.2.6 Dependent measures

Dependent measures for component tasks were response time per task component, number of errors, and self-report workload. Dependent measures for the molar task were total time to complete the editing task, number of errors, and self-report workload.

2.3 Results

2.3.1 Missing data

One subject made an error in each of the backspace component task trials. Response time data were therefore not available for this subject for this component task. The type of error committed in all cases was leaving too many spaces between words after the target word had been removed. The details of this subject are shown in Table 2.01. The data for this subject were included in all analyses which did not involve the backspace component task.

Table 2.01 : Details of subject who made an error in every backspace component task trial	
Age	46 yrs
Sex	Female
Typing test	43.29
Computer literacy	14
Spatial visualisation	5.00
Spatial memory	8.75
Nelson-Denny vocabulary	32
Logical Reasoning	1.00

2.3.2 Distribution of scores

Whilst the distribution of scores with respect to response times / completion times approximated a normal distribution for each of the tasks, the error distributions were positively skewed. A series of Kolmogorov-Smirnov tests indicated that the error distribution for each task significantly deviated from a normal distribution ($p<.01$). Consequently non-parametric analyses were conducted as far as possible where error data were concerned.

2.3.3 The relationship between component and molar task performance

Table 2.02 shows the correlation matrix for component tasks and the complete editing task response times. As can be seen, correlations between component tasks and the complete editing task are all highly significant, with only the decision task falling below .50. However, the validity of this component as one which is related to a process of command generation is supported by the fact that correlations with the other command generation tasks are all .50 or greater. Generally the correlations within finding and generating component tasks are stronger than those between these two groups of tasks. However, there are consistently strong correlations between the fine search task and the generating component tasks. This seems concordant with the discussion above relating the requirement of final cursor positioning to the process of command generation.

Table 2.02 : Correlation matrix of response times for component and complete tasks

	SCRL	FINE	DEC	INST	BSPC	DELT	MOL
PAGE	.76 ***	.45 ***	.53 ***	.52 ***	.31 **	.34 **	.60 ***
SCRL		.45 ***	.60 ***	.48 ***	.31 **	.38 ***	.65 ***
FINE			.50 ***	.74 ***	.50 ***	.65 ***	.54 ***
DEC				.50 ***	.56 ***	.56 ***	.45 ***
INST					.71 ***	.72 ***	.72 ***
BSPC						.73 ***	.54 ***
DELT							.50 ***

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

The correlation matrix for the error performance in the component tasks and the complete editing task are shown in Table 2.03. Correlations are consistently much smaller than for response times. Of note is the comparatively large correlation between the number of errors for the scroll task and that for the delete task, combined with the significant correlations between these two component tasks and error performance for the molar task.

Table 2.03 : Correlation matrix of error performance for component and complete tasks

	SCRL	FINE	DEC	INST	BSPC	DELT	MOL
PAGE	.36 **	.25 *	.29 *	.17 ns	.25 *	.32 **	-.12 ns
SCRL		.01 ns	.06 ns	.33 **	.05 ns	.40 ***	.32 *
FINE			.15 ns	.34 **	.27 *	-.04 ns	-.02 ns
DEC				.14 ns	.08 ns	-.03 ns	.02 ns
INST					.14 ns	.13 ns	.19 ns
BSPC						.36 *	-.09 ns
DELT							.33 **

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

In order to facilitate an examination of individual differences in performance with respect to Egan and Gomez's (1985) finding and generating task components two additional composite measures were created. These were a 'Finding' composite which was based upon performance for the page, scroll and fine search tasks, and a 'Generating' composite which was based upon performance for the decision, insert, delete, and backspace tasks. Response times and error data for each component task were transformed into z scores, summed to form a total for each composite (one for response times and one for errors), and then divided by the number of elements in the respective composites.

2.3.4 Age differences in performance

Age differences in typing speed and computer literacy were examined using unrelated t-tests. The mean scores and results of these analyses are shown in Table 2.04. As can be seen, whilst in both cases there is a small performance advantage for the younger subject group, this difference is not significant.

Table 2.04 : T-tests results of differences between younger and older groups for typing speed and computer literacy

	Typing speed (secs)	Computer literacy
Younger	35.05	28.44
Older	38.40	25.91
t	-1.04	1.45
p	ns	ns

Table 2.05 : T-tests of response times (secs.) for younger and older subjects (one-tailed sig.)

	Page	Scroll	Fine	Decis.	Insert	Bspce	Delete	Molar
Young	5.86	6.18	4.08	2.19	4.13	3.61	3.95	509.27
Old	7.32	8.46	5.31	2.96	5.66	6.46	6.59	717.09
t	-2.82	-3.53	-4.19	-5.74	-4.02	-5.30	-5.45	-4.89
p	<.01	.001	<.001	<.001	<.001	<.001	<.001	<.001

A series of t-tests were used to examine the relative response times for younger and older age groups for each component task and for the complete editing task. As can be seen from Table 2.05 there were significant differences in response times for all tasks. Whilst the magnitude of these differences is slightly smaller for the coarse search tasks, it would appear that the effects of age were broadly similar across all component tasks.

A series of Mann-Whitney U tests was conducted upon the proportion of errors for component and molar tasks for younger and older subject groups. The results of these tests along with mean scores are shown in Table 2.06. As can be seen, the scroll task was the only task for which a significant difference was found, with older subjects generating more errors than younger subjects. The above mentioned response time advantage for the younger subject group cannot be attributed to a speed accuracy trade-off as, whilst error rates were low for both groups, the young subject group tended to perform more accurately in most of the tasks.

Age differences in self-report workload were also examined. There were no significant differences in workload for any of the task components or for the molar task, although mean scores for the older group were lower for each task condition.

Table 2.06 : Mean scores and results of Mann-Whitney U tests for the proportion of errors committed by younger and older subjects

	Page	Scroll	Fine	Decis.	Insert	Bspce	Delete	Molar
Young	.061	.002	.068	.039	.043	.070	.043	.010
Old	.063	.052	.059	.029	.057	.090	.079	.021
U	494.5	344	455.5	478	463.5	452	435.5	457.5
p	>.1	<.05	>.1	>.1	>.1	>.1	>.1	>.1

2.3.5 Cognitive predictors of performance

Table 2.07 presents the correlation matrix for the cognitive ability measures. As can be seen, intercorrelations between spatial visualisation, spatial memory, and logical reasoning are strong.

Table 2.07 : Correlation matrix for cognitive ability measures

	SM		ND		LR	
SV	.60	***	.27	*	.72	***
SM			.23	ns	.62	***
ND					.50	***

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

The correlations between each of the predictor variables and each of the component tasks which can be designated as 'finding', i.e. are concerned with the location of the target word, are shown in Table 2.08. With the exception of the correlation between vocabulary and fine search all of these correlations are highly significant and in the predicted direction, with high scores on the cognitive ability test being associated with fast performance in each of the component tasks. The strongest correlations for the page and scroll tasks are with logical reasoning, whilst the strongest correlation for the fine search task is with spatial memory.

Table 2.08 : Correlations between predictor variables and response time for 'finding' task components

	Page	Scroll	Fine Search
SV	-0.30 **	-0.35 **	-0.39 ***
SM	-0.35 **	-0.44 ***	-0.51 ***
ND	-0.35 **	-0.37 ***	-0.14
LR	-0.49 ***	-0.55 ***	-0.45 ***
LIT	-0.39 ***	-0.38 ***	-0.38 ***

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

Spearman correlation coefficients were calculated using the error data for the same component tasks and these are presented in Table 2.09. The only significant correlations were obtained for the scroll component task, for which spatial visualisation, spatial memory, and logical reasoning were all significantly related to performance. There was no evidence of a speed accuracy trade-off involving cognitive ability, with subjects high in cognitive ability also tending to perform more accurately than those low in cognitive ability. The scroll task dependent measures are particularly complementary suggesting a stronger association between cognitive ability and performance than is evidenced by response time data alone.

Table 2.09 : Correlations between predictor variables and number of errors for 'finding' task components

	Page	Scroll	Fine Search
SV	-.11	-.41 ***	0.00
SM	-.17	-0.36 ***	.02
ND	-.04	-0.19	-.02
LR	-.09	-.43 ***	.05
LIT	-.04	-02	.12

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

Table 2.10 shows the correlations between the predictors and response time performance upon the component tasks which can be classed as the 'generating' elements of word processing. Spatial visualisation, spatial memory, and logical reasoning all show consistently high correlations, whilst the correlations for computer literacy are more modest, and the association with verbal ability (Nelson-Denny) is generally fairly weak.

Table 2.10 : Correlations between predictor variables and response times for 'generating' task components

	Decision	Insert	Backspace	Delete
SV	-0.48 ***	-0.43 ***	-0.54 ***	-0.47 ***
SM	-0.63 ***	-0.46 ***	-0.53 ***	-0.51 ***
ND	-0.23 *	-0.21	-0.15	-0.16
LR	-0.55 ***	-0.48 ***	-0.52 ***	-0.49 ***
LIT	-0.37 ***	-0.23 *	-0.30 **	-0.23 *

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

Spearman correlation coefficients for cognitive ability and the generating component tasks are presented in Table 2.11. Correlations are weak, with the exception of the delete component task for which there are modest significant correlations with spatial memory, verbal ability, logical reasoning and computer literacy. In all cases where there is a significant correlation the relationship between ability and error performance is negative. Therefore, as with the finding component tasks there is no evidence of an ability related speed accuracy trade-off.

Table 2.11 : Correlations between predictor variables and number of errors for 'generating' task components

	Decision	Insert	Backspace	Delete
SV	-.24 *	-.11	.03	-.15
SM	-.16	-.03	.04	-.29 **
ND	-.05	-.01	-.03	-.21 *
LR	-.21	-.01	-.08	-.29 *
LIT	-.07	.04	-.11	-.27 *

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

Table 2.12 shows the correlations between each of the predictor variables and completion times for the molar task and the finding and generating composites. With the exception of computer literacy all predictors are highly significantly correlated with molar performance, with logical reasoning showing the strongest association. Verbal ability (Nelson-Denny, 1973) shows the weakest associations with the composite measures, with the correlation failing to reach significance for the generating component.

Table 2.12 : Correlations between predictor variables and response time for the finding and generating composites and for the molar task

	Finding composite	Generating Composite	Molar Task
SV	-.42 ***	-.56 ***	-0.39 ***
SM	-.52 ***	-.62 ***	-0.44 ***
ND	-.35 **	-.20 ns	-0.39 ***
LR	-.59 ***	-.60 ***	-0.52 ***
LIT	-.45 ***	-.32 *	-0.21

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

Table 2.13 presents Spearman correlation coefficients for cognitive ability with the finding composite, generating composite, and the molar task. Correlations are generally small and negative, with no significant associations with molar task performance.

Table 2.13 : Correlations between predictor variables and number of errors for the molar task

	Finding composite	Generating composite	Molar task
SV	-.29 *	-.28 *	-.09
SM	-.21	-.24	-.16
ND	.06	-.19	-.02
LR	-.22	-.26 *	-.20
LIT	.00	-.13	-.15

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

2.3.6 Age differences in the relationship between cognitive ability and performance

The results of a series of t-tests which examined differences between younger and older subject groups upon each of the cognitive ability tests are shown in Table 2.14. Given the pattern of cognitive ageing, as presented earlier, one-tailed tests were used with respect to scores upon tests of spatial visualisation, spatial memory, and logical reasoning, upon the grounds that older subjects could be predicted to be less able than younger subjects. The results of these tests supported this prediction with a highly significant performance advantage for the younger group in each case. However, there was no theoretical or empirical basis for predicting such a difference with respect to verbal ability (Nelson-Denny vocabulary test) scores and consequently a two-tailed test was used. Whilst a performance superiority was apparent for the younger group this difference was not significant.

Table 2.14 : Age differences in cognitive ability

	SV	SM	ND	LR
Younger	13.36	18.15	72.78	21.78
Older	7.43	11.63	65.69	12.79
t	6.00	5.59	1.81	5.52
p	<.001	<.001	ns	<.001

In order to assess the possible interactive effects of age with each of the cognitive predictors a series of multiple regression analyses were conducted for each task condition. Response times for the molar task, the finding and generating composites, and each component task were regressed separately upon each cognitive ability measure, age group, and a vector representing the interaction between age and cognitive ability (cf. Pedhazur, 1982). The main effects of cognitive ability and age are not reported as these replicate previous analyses. Significant interactions were found between age and computer literacy for the page task ($F(1,60)=4.40 : p<.05$), and between age and vocabulary scores for the page task ($F(1,60)=4.74 : p<.05$), the scroll task ($F(1,60)=8.55 : p<.01$), the 'finding' composite ($F(1,60)=7.76 : p<.01$), and the molar task ($F(1,60)=12.98 : p<.001$). The nature of all of these interactions was very similar, and those for the page, scroll and molar tasks are shown in Figures (2.01 to 2.04). As can be seen, verbal ability and computer literacy are predictive of performance for the older age group, with high ability individuals performing more quickly. However, there appears to be no association between these variables and the task performance of the younger group. In the case of the molar task, the association is in the opposite direction with high ability younger individuals performing more slowly.

Fig. 2.01 : Page task : Regression of response times upon computer literacy for the younger and older age groups

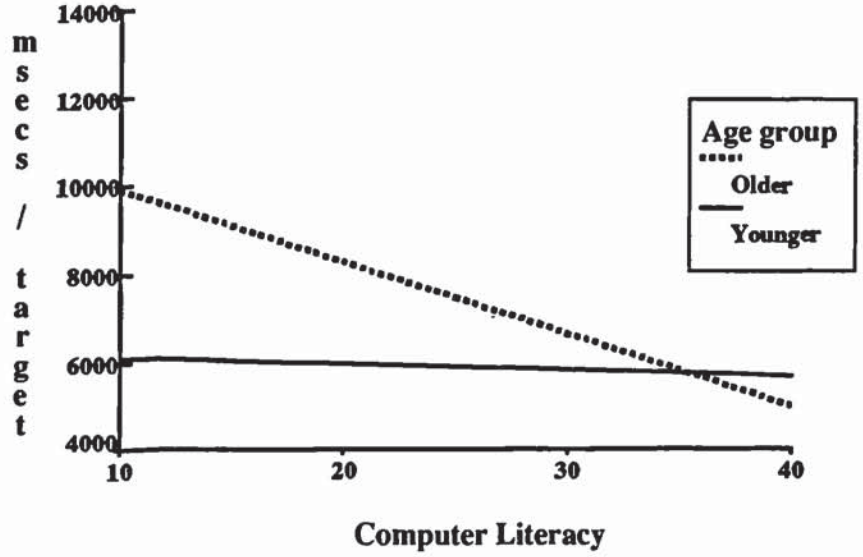


Fig. 2.02 : Page task : Regression of response times upon verbal ability for younger and older age groups

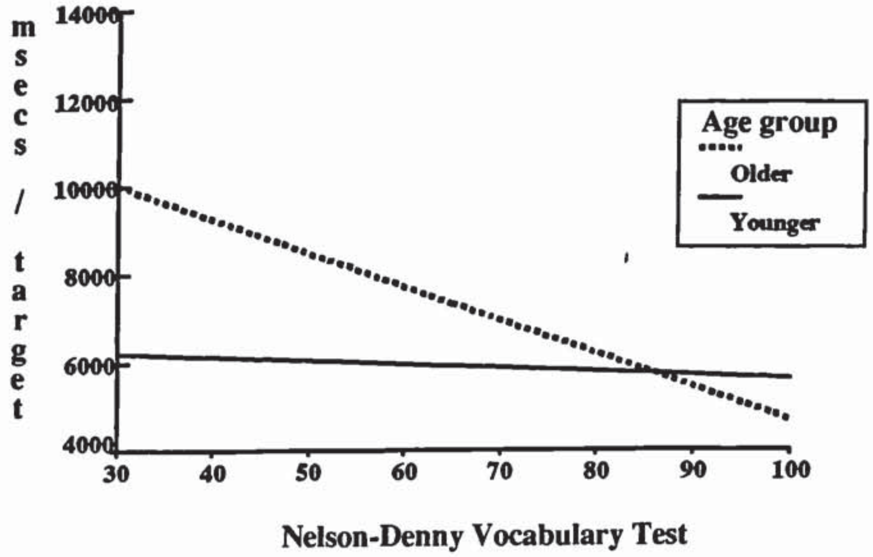


Fig. 2.03 : Scroll task : Regression of response times upon verbal ability for younger and older age groups

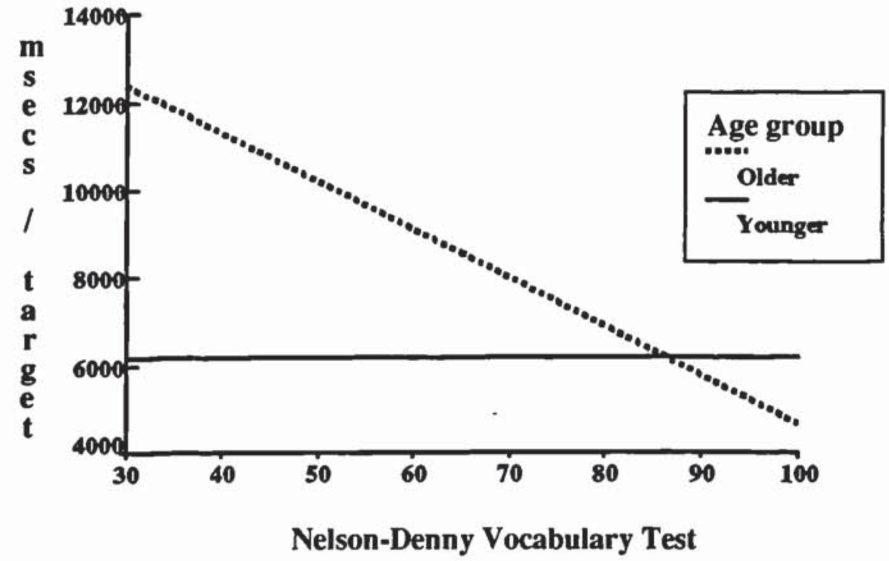
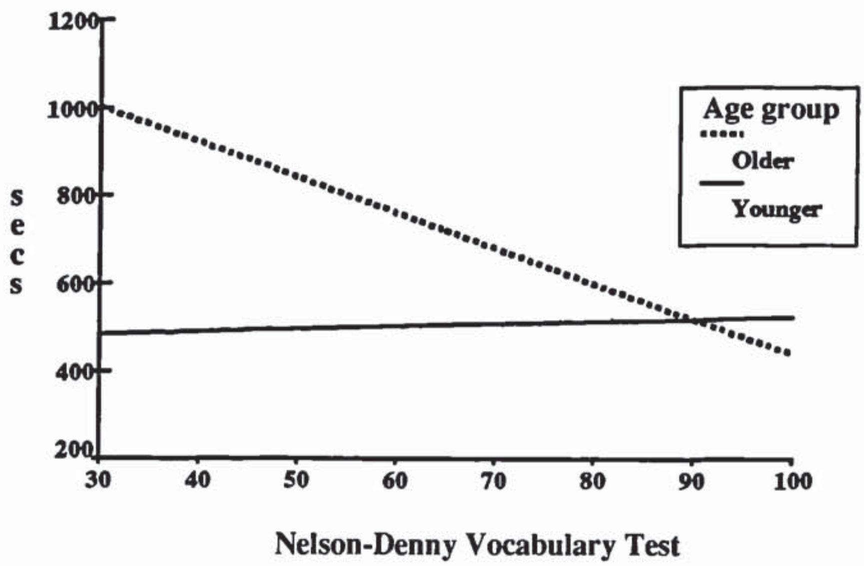


Fig. 2.04 : Molar task : Regression of completion times upon verbal ability for younger and older age groups



Whilst the skewed distribution of error scores makes the interpretation of regression analyses somewhat uncertain, similar analyses were conducted as for error data. The only significant interaction was with spatial memory for the fine search task

($F(1,60)=6.40$: $p<.05$). Once again, cognitive ability was more strongly associated with performance for the older subject group, with high ability subjects committing fewer errors.

2.3.7 The relative importance of predictors of task performance

In order to examine the relative importance of each of the predictor variables to each of the component tasks and the molar task, a series of stepwise regression analyses was performed. Variables entered into the equation included age group, each of the cognitive predictors, computer literacy, and self-report workload. The results of these analyses are shown in Tables 2.15 to 2.22.

Table 2.15 : Stepwise regression equation for response times for the page task

Mult. R	0.493	F	19.858
R square	0.243	p	<.0001
Variable	Beta	t	p
LR	-0.493	-4.456	<.0001

Table 2.16 : Stepwise regression equation for response times for the scroll task

Mult. R	0.547	F	26.533
R square	0.300	p	<.0001
Variable	Beta	t	p
LR	-0.547	-5.151	<.0001

Table 2.17 : Stepwise regression equation for response times for the fine search task			
Mult. R	0.571	F	14.786
R square	0.327	p	<.0001
Variable	Beta	t	p
SM	-0.445	-4.104	0.0001
LIT	-0.267	-2.464	0.0166

Table 2.18 : Stepwise regression equation for response times for the decision task			
Mult. R	0.715	F	20.929
R square	0.511	p	<.0001
Variable	Beta	t	p
SM	-0.383	-3.412	0.0012
AGE	0.329	2.964	0.0043
LIT	-0.213	-2.288	0.0257

Table 2.19 : Stepwise regression equation for response times for the insert task			
Mult. R	0.528	F	11.82
R square	0.279	p	<.0001
Variable	Beta	t	p
LR	-0.329	-2.482	0.0158
AGE	0.265	2	0.05

Table 2.20 : Stepwise regression equation for response times for the backspace task			
Mult. R	0.617	F	18.442
R square	0.381	p	<.0001
Variable	Beta	t	p
AGE	0.383	3.096	0.003
SM	-0.312	-2.516	0.0146

Table 2.21 : Stepwise regression equation for response times for the delete task			
Mult. R	0.647	F	14.467
R square	0.42	p	<.0001
Variable	Beta	t	p
AGE	0.43	3.562	0.0007
WL	0.223	2.262	0.0273
SM	-0.27	-2.24	0.0288

Table 2.22 : Stepwise regression equation for completion times for the molar task			
Mult. R	0.595	F	16.746
R square	0.354	p	<.0001
Variable	Beta	t	p
AGE	0.464	4.401	<.0001
ND	-0.283	-2.68	0.0095

As can be seen, age only remains in the regression equations for tasks which are generating components, whilst logical reasoning is the only variable to consistently remain in the equation for the finding tasks. Spatial memory accounted for the greatest proportion of variance associated with performance of the fine search and decision tasks. With respect to the molar task, age and verbal ability are the only variables to remain in the equation.

2.3.8 Age differences in component task performance for subjects of similar molar task ability

In order to examine the possibility that some older subjects maintain comparable molar task performance by virtue of superior performance of particular component tasks (Charness, 1981a, 1981b; Salthouse, 1984), a sub-sample of both younger and older age groups were selected for which molar task performance was very similar. By excluding the fourteen best performers from the younger subject group and the fourteen worst performers from the older subject group very similar group means for molar task performance were obtained (see Table 2.23). These reduced samples were then examined for differences in performance upon the component tasks and differences in cognitive ability (see Table 2.24).

Table 2.23 : T-tests of reaction times (secs.) for molar equivalent younger and older subjects (one-tailed sig.)

	Page	Scroll	Fine	Decis.	Insert	Bspce	Delete	Molar
Young	6.16	6.64	4.42	2.32	4.57	4.05	4.29	568.88
Older	6.30	7.13	4.87	2.83	4.51	5.37	5.90	580.89
t	-.27	-.71	-1.32	-2.48	.17	-2.67	-3.04	-.46
p	ns	ns	ns	<.01	ns	<.01	<.01	ns

Table 2.24 : T-tests of scores for molar equivalent younger and older subjects on typing, computer literacy and cognitive ability tests

	Type (secs.)	Lit	SV	SM	ND	LR
Younger	34.77	28.61	13.13	17.44	75.22	22.32
Older	29.21	26.72	8.75	12.78	70.44	15.61
t	1.51	.83	3.25	2.91	1.06	2.89
p	ns	ns	<.01	<.01	ns	<.01

Age differences in performance upon the 'finding' component tasks (page, scroll, and fine search) were no longer significant (see Table 2.23), although there was a tendency for younger subjects to perform more quickly. There was also no significant age difference in response times for the insert component task, and examination of means indicated marginally faster performance for the older group. There were still significant age differences for the decision task, the backspace task, and the delete task, although the magnitude of these effects was much reduced. There were no significant age differences with respect to component or molar task errors. These changes are therefore not attributable to a speed accuracy trade-off on the part of the older age group. Similarly, age group differences in self-report workload scores for each of the tasks were not significant. As with the full sample, the older group tended to report lower workload than the younger group. The only exception to this was the fine search task for which workload scores were very similar. Consequently, this would suggest that the improvement in the performance of the older group cannot be attributed to increased attentional resource allocation on the part of the individuals in this sub-sample.

Table 2.24 show the results of a series of unrelated t-tests which examined age group differences for this reduced sample upon the typing test, the computer literacy test, and each of the cognitive ability measures. As can be seen the older age group tend to be faster at typing than the younger group. Whilst the original performance advantage for the younger group remained for each of the other tests, the magnitude of this advantage was reduced and there was no significant difference between groups for the computer literacy or verbal ability tests.

2.4 Discussion

2.4.1 The effects of age

There was a highly significant age difference in the performance times of all the component tasks with the older subject group performing more slowly than the younger subject group. Whilst this is consistent with the results of Davies at al. (1993) it is contrary to the findings of Egan and Gomez (1985) who suggested that age was only related to the 'generating' task components. However, in support of this position, it would seem that the magnitude of age differences in response times is reduced for the 'finding' task components. This is also consistent with the results of the series of stepwise multiple regression equations, in which age only remained in the equation for those task components concerned with the process of 'generating'. It would seem that whilst older subjects are at a response time disadvantage for all component elements of word processing this is particularly evident in the case of the generating task components.

A relatively uniform age difference in performance speed across all component tasks is in keeping with recent investigations into the effects of age upon spatial and logical reasoning tasks. As mentioned in Chapter 1, it has been proposed that fundamental age differences in processing speed underlie age differences in many more complex cognitive tasks (Salthouse, 1991b; Salthouse, 1992). This is also consistent with the findings of Davies et al. (1993) who report age differences in word processing as being related to average time per key press but not to the number of keystrokes performed. Their results may also be indicative of a fundamental age-related processing speed overhead. The significant age difference in molar task completion time is consistent with previous research (e.g. Egan and Gomez, 1985; Czaja, Joyce, and Hammond, 1989) and lends further weight to the existence of an overall age-related deficit in word processing performance speed.

With the exception of the scroll component task, there were no age differences in performance accuracy. Previous research has produced somewhat inconsistent results with respect to age differences in word processing error rates. As mentioned above, error type may be of importance in this respect. It is possible that the tasks used in the current experiment were not of a type conducive to the commission of age-related errors. It is interesting to note that error performance upon the scroll task was not only subject to age differences, but was also significantly correlated with spatial visualisation, spatial memory and logical reasoning, suggesting that accurate performance of this task requires specific cognitive processing abilities of a type which usually deteriorate with age (Salthouse and Mitchell, 1990; Horn and Hofer, 1992). Evidence for this decline is apparent in the significant age differences for each of these measures in the present sample. Why the scroll component task should be particularly sensitive to such age and ability differences is unclear and warrants further investigation.

2.4.2 Cognitive ability

Spatial memory was consistently significantly correlated with each of the component tasks. This is concordant with the results of Egan and Gomez (1985) which indicated that spatial memory was predictive of both the finding and generating task components. However, spatial visualisation and logical reasoning also followed a similar pattern. When considered in conjunction with the high intercorrelation between these factors (see Table 2.07), it would seem that a higher-order fluid intelligence factor may underlie these results (Horn and Hofer, 1992; Carroll, 1993). This is also supported by Egan and Gomez's (1985) fourth experiment in which logical reasoning was found to account for significant unique variance in performance times when using a display editor, and also by the significant correlation between logical reasoning and performance reported by Czaja, Joyce, and Hammond (1989). In the present experiment none of these variables remained in the stepwise multiple regression equation for molar task completion times. Whilst, for the component tasks, spatial memory remained in the equation for the fine search, decision, backspace, and delete tasks, logical reasoning remained in the equation for both of the coarse search tasks and the insert task. An examination of the component task correlation matrix reveals no obvious pattern of association which differentiates those components which are related to spatial memory and those components which are related to logical reasoning. There is no indication, therefore, of common task processing demands which were specifically associated with either of these cognitive abilities. However, it is possible that the relative strength of the correlation between each of these cognitive abilities and molar task performance will be influenced by the relative contribution of each of the component elements to the molar task.

The only cognitive ability which differentially predicted task component response times was verbal ability (Nelson-Denny vocabulary test). This measure was only significantly correlated with performance upon the coarse search tasks, and more modestly correlated with performance upon the decision task. It is interesting to note that there is a comparatively strong correlation between the scroll task and the decision task response times. It is possible that this is indicative of a common visual scanning component. Subjects in each instance are required to rapidly process a stream of lexical or symbolic information. It may be that the association with verbal ability is due to a speed advantage for high verbal ability individuals with respect to such processing (Hunt, 1978). A further indication that verbal ability is more strongly associated with the finding task components can be found in the relative magnitude of correlation with the finding and generating composites. Consistent with previous research (e.g. Sebrechts et al., 1984) verbal ability was also significantly correlated with molar task performance and accounted for unique variance, in combination with age, in the stepwise multiple regression equation for the molar task. The association between verbal ability and age will be elaborated below. It would appear that the variance in molar task performance which is accounted for by verbal ability is probably due to the coarse search demands of this task. This hypothesis is supported by the fact that Egan and Gomez (1985) found that verbal ability was the strongest predictor of performance when subjects were using a display editor, but accounted for very little variance when subjects were using a line editor in which the process of coarse search was not required.

The correlations between cognitive ability and performance accuracy were far weaker than those with performance speed. However, as mentioned above, the scroll task proved an exception in this respect. Significant but modest correlations were also obtained for the delete task with respect to spatial memory, vocabulary, computer literacy, and logical reasoning. All of these correlations were indicative of high ability scores being associated with more accurate performance, and do not therefore suggest an ability related speed accuracy trade-off. Similarly, spatial visualisation was significantly correlated with decision task errors. Whilst it would appear that a pattern exists with respect to the correlations between cognitive ability and error performance, with the scroll task and the delete task being the primary focus of this relationship, the mechanisms underlying this pattern are unclear. It is interesting to note that the correlations between error performance for the scroll task, delete task, and molar task are comparatively strong. It would appear that there may be an underlying similarity in the cognitive demands of these tasks. It is possible that, in the same way that age differences in error performance seem to be related to error type, similar differences occur in relation to cognitive ability. On the basis of the current findings it is difficult to speculate as to the exact nature of this relationship. However, this is an area which future research might usefully address.

The fact that significant correlations were found between cognitive ability and performance for expert word processor users indicates the importance of individual differences as an area for research. In support of Ackerman's (1988) theory of individual differences in skill acquisition there is some indication that the variance accounted for by cognitive ability is smaller than that obtained in comparable experiments which have used a novice sample. The fourth experiment of Egan and Gomez (1985) used very similar predictors to examine performance using a display editor. The regression equation for this experiment indicated that the predictor variables had accounted for 65% of the variance in response times, whereas the molar task regression in the current experiment accounted for only 35% of the variance. However, such differences may be due to task factors. The tasks used in this experiment were less complex and might therefore be predicted to be less strongly associated with cognitive ability. Alternatively, these differences may be attributable to random variation. Regardless of the reduction in variance accounted for, it would appear that even simple word processing tasks retain sufficient cognitive demand to ensure that cognitive ability remains predictive of the performance of moderately skilled users. An interesting area for future research would be to examine molar and molecular word processing tasks over repeated trials, with consideration being given to the relationship between task performance and measures of cognitive ability, processing speed, and psychomotor skill.

2.4.3 Age and cognitive ability

As discussed earlier, given that there is an age related decline in many cognitive abilities such as spatial ability and reasoning (Salthouse, Mitchell, Skovronek, and Babcock, 1989; Salthouse, 1991c), as is apparent for the present sample, one way in which older subjects could compensate for these deficits would be to rely more heavily upon other cognitive abilities (e.g. verbal ability, crystallised intelligence) for which the effects of age are less pronounced (Dixon, 1992). Upon this basis verbal ability should be more strongly associated with performance in the older group. This was apparent in the interactive effects of age and verbal ability with respect to response times for the coarse search tasks and the molar task. The regression lines shown in Figures 2.02 to 2.04 suggest that verbal ability is predictive of performance speed for the older group but not the younger group when performing coarse search task elements, and that it is this association which accounts for the interactive effects of age and verbal ability upon molar task performance. This position is further supported by the finding that the response times of the reduced molar equivalent sample were not significantly different for the coarse search task components or for scores on the vocabulary test. A similar result was obtained with respect to the computer literacy for the page task. Once again this is consistent with a process of compensation in which older subjects utilise age insensitive abilities, in this case semantic knowledge, to improve performance. The mechanisms which underlie this process, however, are not clear, and could usefully be the subject of future research.

It would appear that verbal ability is an important element in a process of age-related performance compensation, and that older subjects with greater verbal ability are able to engage performance strategies which improve performance. This finding is contrary to the results of Davies et al. (1993) who report a trend such that verbal ability is more strongly predictive of performance accuracy for the younger subject group. This apparent disparity may be due to differences in the relative experience of the experimental samples and indicate, as previously mentioned, that experience is required before compensation can occur.

A comparison of the performance of the younger and older groups in the molar equivalent sample indicates that age differences no longer exist for the finding component tasks or for the insert task. A process of compensation relating to the coarse search tasks has already been discussed. However, typing speed may form the basis of another compensatory mechanism, of the sort proposed by Charness and Bosman (1990). There is a non significant response time advantage for the molar equivalent older group on both the insert task and the typing test. It would appear these older subjects may compensate for slower performance upon some component elements of the molar task by virtue of increased typing speed (see Salthouse, 1984). This is supported by the non-significant trend for the older group to report lower self-report workload in both the full sample and the reduced sample, suggesting that the performance improvement of the reduced sample cannot be attributed to increased task resource allocation.

2.4.4 Conclusions

To summarise, pervasive effects of age upon response times were found which were greatest for task components associated with the process of command generation (Egan and Gomez, 1985), but which also extended to the finding component tasks. Spatial memory, logical reasoning, and, to a slightly lesser extent, spatial visualisation, were all consistently predictive of component task response times. The relative importance of spatial memory and logical reasoning appeared to be task dependent, although no pattern could be determined. The consistent association between these three variables and task performance may indicate that a higher order fluid intelligence factor (Horn and Hofer, 1992; Carroll, 1993) is of greater predictive importance.

Two possible patterns of age-related compensation were identified. The first of these was such that, upon the basis of an interaction between verbal ability and age for the coarse search and molar tasks, it was suggested that successful performance by older subjects was the result of an increased reliance upon processing related to verbal ability upon the part of high verbal older subjects. This is consistent with verbal ability being relatively insensitive to the effects of cognitive ageing (Davies, Taylor, and Dorn, 1992; Horn and Hofer, 1992). The second compensatory process appeared to be related to typing speed, with better older performers being marginally faster than molar equivalent younger subjects upon tests relating to typing speed, although still at a performance speed disadvantage (albeit reduced) for the other task component elements.

The fact that substantial individual differences in performance were apparent in a sample of experienced word processor users was taken to reflect the importance of this area of investigation.

Chapter 3

The 'finding' component: Text search

3.1 Introduction

This chapter presents an experimental investigation of individual differences in the location of target words within passages of text. Consideration is given to the cognitive demands associated with different command methods, window sizes, target distances, and target locations. This introductory section reviews previous research relating to each of these factors, with particular reference to a model of 'visual momentum' which may be applied to such tasks.

The location of target words within text is common to many computer applications. The tasks of information retrieval, word processing, and programming all contain such a requirement. The importance of spatial ability as a predictor of this task component has been demonstrated by Egan and Gomez (1985) in a series of word processing experiments, and by Vicente, Hayes and Williges (1987; Vicente and Williges, 1988) in studies of information retrieval within a textual database. However, comparatively little attention has been given to the mechanisms underlying this process. In another study of text editing, Sebrechts, Deck, Wagner, and Black (1984) failed to replicate the above mentioned results of Egan and Gomez (1985). They suggested that this may have been due to differences in the density of edits within the text which influenced the spatial processing requirements of the task. It would certainly appear that the spatial processing demands which are placed upon the user and the spatial information content which is presented within the interface may be of great importance when assessing the impact of individual differences upon text search. These factors are investigated further in this chapter.

One explanatory model which may be applied in these settings is that of visual momentum (Woods, 1984). According to Woods an important dimension of cognitive demand within the interface relates to the degree of continuity of movement which is achieved. Discontinuous movements within the task environment require the user to reorient their perspective and to locate newly acquired information within the broader information context. This situation of low visual momentum places increased cognitive demands upon the user. Attentional resources must be allocated to the process of orientation within the task environment. In contrast, continuous movements within the task environment are held to facilitate the assimilation of new information, and consequently provide a situation of high visual momentum which permits a greater degree of automatic processing. Vicente and Williges (1988) suggest that this model may account for spatial ability related differences in command selection during the process of information retrieval from a textual database. Vicente, Hayes, and Williges (1987) found that subjects who scored low on a test of spatial visualisation (VZ2 ; Ekstrom, French, and Harman, 1976) used the scroll command more frequently than high spatial ability subjects who used more powerful, but less continuous, search commands. They attributed this difference to a problem of disorientation within the text files and a need for high visual momentum on the part of the low spatial ability subjects. In an attempt to improve the relative performance of these subjects Vicente and Williges (1988) designed an interface which contained additional locational information. They hypothesised that this would reduce the spatial processing demands of the task. Whilst this modification did not reduce the variance associated with spatial ability there was a general performance improvement and a significant decrease in the use of the scroll command, suggesting that visual momentum had increased for the whole sample.

A series of experiments conducted by Elkerton, Williges, Pittman, and Roach (1982; Elkerton and Williges, 1985, 1987) also indicates that increased cognitive demand is associated with low visual momentum. These experiments compared the file search performance of novices and experts. Experts were found to perform significantly more quickly than novices, and to use fewer total operations (different commands). However, there was also a significant difference in the performance strategy adopted by experts. Experts tended to use more powerful, less continuous string search and boolean search commands, whilst novices adopted command strategies which gave greater continuity of file movement. Elkerton and Williges (1987) divided the novice group into good performers and poor performers on the basis of response times. They found that poor performers used the scroll command significantly more frequently than good performers or experts.

3.1.1 Page and search commands as a means of text search

This experiment examines individual differences in performance relating to the cognitive demands of two alternative text search commands. The first of these is the scroll command which, as discussed above, is thought to generate high visual momentum and reduce cognitive demand, particularly the requirement for spatial processing. The second is the page command, which is held to generate greater discontinuity of movement than the scroll command and therefore reduced visual momentum. These commands are perhaps the two most universally implemented text search commands. The level of discontinuity associated with the page command is a function of window size, something which will be further discussed below.

There are few direct comparisons of text search performance whilst using the page and scroll commands. Schwarz, Beldie, & Pastoor (1983) examined the relative performance of subjects upon word reading, line searching, and sorting tasks whilst using either paging or scrolling as a means of manipulating the contents of the computer screen. It should be noted that only two screen pages of information were used. Consequently, the number of commands required were few and orientation demands were low, particularly for the page task. There was no significant completion time difference for any of the tasks. However, in the sorting task the number of errors made was significantly less when using the page command, and there was a significant difference in the preference ratings given by subjects for the reading task, with paging being the preferred method. The results of Alvarez, Murray, and Hakkinen (1984) also indicated a performance superiority for the page command in a text search experiment which required subjects to locate target words on the basis of semantic or physical content. However, Swierenga (1990) found the scroll command to be a consistently more efficient means of search in an experiment which compared subject performance whilst accessing a small database using menuing, paging (full page), paging (half page) and scrolling techniques. Kolers, Duchnicky, and Ferguson (1981) demonstrated that the rate of scroll may be an important factor in determining the most efficient means of text manipulation. Reading comprehension was examined in relation to the use of either the page or scroll commands. Performance with each command was found to interact with the rate of scroll. There was an advantage for the page command when scrolling occurred at the preferred rate of the subject, but an advantage for the scroll command when the rate of scroll was increased to 20% above the preferred rate.

A related study by Card, Moran and Newell (1983) suggests that the process of skill acquisition may be another related factor. This experiment compared text selection performance whilst using scroll keys or 'text' keys. Text keys provided four commands: move the cursor to the beginning of the next paragraph, move the cursor down one line, move the cursor forward one word, and move the cursor forward one character. Issuing these commands in conjunction with an additional key allowed for backwards movement. Five distances to target were used: 1, 2, 4, 8, and 16 cms. Whilst the number of subjects was small (n=4) a large number of trials were completed. Performance was initially quicker when using the scroll keys, however, there was a steeper acquisition slope for the step keys and after 1200 trials there was a slight performance advantage for this command method. The step keys were more sensitive to target distance and the scroll command remained quicker when the distance to target was great (8 or 16 cms.)

In summary, the evidence is somewhat contradictory. It would appear that comparative page and scroll command performance may interact with the rate of scroll. However, it should be noted that the evidence supporting this position was obtained using a reading comprehension task for which preferred rates of scroll may be much slower than would be the case for text search. Given that the information processing constraints of the individual are not exceeded, the rate of scroll can be predicted to be directly related to the efficiency of this command. An inverse relationship between scrolling rate and task load is supported by findings of Alvarez, Murray, and Hakkinen (1984) such that scrolling rate was faster when feature search was being performed than when the more cognitively demanding semantic search was being performed. In the present experiment the scrolling rate was set at a speed comparable with most commercial word processing applications. The results of Card et al. (1983) suggest that expertise may also be an interacting factor when comparing performance using the page and scroll command. The present experiment examines the performance of both novice and expert users.

3.1.2 The effects of window size and target location

The effects of window size are of growing importance. Along with the advent of graphical user interfaces (GUIs) comes the use of windowing techniques which often result in the user performing many tasks in several small windows (Norman, Weldon, and Shneiderman, 1986). Additionally certain makes of laptop computer use smaller that average displays for reasons of portability (Hotson and Shackel, 1990). Given that the number of lines of text moved when using the page command is generally related to the size of the window in which the text is displayed, this factor can also be predicted to interact with comparative command efficiency. Windows of greater size will produce greater discontinuity, resulting in reduced visual momentum and increased cognitive demand. However, a further complicating factor is that large windows also provide more locational clues, such as paragraphs or line indentations, which may be used to facilitate the process of text search. This experiment examines the importance of these factors in relation to cognitive ability. Similarly, the location of target words within passages of text alter the level of spatial information which is provided to the user. If the target word is located within the body of a paragraph it can be predicted that there will be less spatial information available relating to the location of the target than if the word was located on the borders of a paragraph.

These variables were considered in the experiments mentioned above by Elkerton, Williges, Pittman and Roach (1982; Elkerton and Williges,1984a). File search performance was examined in text windows consisting of 1, 7, 13, and 19 lines. Subjects had five search procedures available, scroll, page, search (find), absolute line movement, and relative line movement. File length (45 lines or 200 lines), type of file (text or data) and type of target (embedded, non-embedded, or repeated) were also manipulated. Main effects of file length and window size were found for all dependent measures (operations per trial, number of different operations, time taken, and file movement distance). In addition embedded and repeated targets were significantly more difficult to locate than non-embedded targets. On a range of dependent measures, including completion rates, performance using the 1 line window was significantly worse than for the other three window sizes. However, performance in the 7 line window was significantly worse than the 13 and 19 line windows with respect to the number of operations per trial and the number of different operations. File movement, however, was found to increase with larger windows. It would appear therefore that larger windows promoted an increased efficiency with respect to the number and types of command generated but decreased the economy of file movement.

A lack of performance advantage associated with window sizes above seven lines is supported by an experiment conducted by Shneiderman (1987c). Performance upon an information retrieval task was compared for window sizes of 9, 18, and 34 lines (n=12). There was no significant effect upon task completion times although there was a trend in the direction of faster performance with increasing window size. There was, however, a significant effect with respect to subject preference with the majority of subjects preferring the 18 line display, and the remainder preferring the 9 line display. Similar results were also obtained in the study of Swierenga (1990), mentioned above. Performance was compared whilst using 12 line and 24 line windows. No significant effects of window size were found. Hotson and Shackel (1990) provide further evidence of the detrimental effect upon performance of window sizes of seven lines or less in a study of text editing performance. Three window sizes were compared (6 lines, 12 lines, and 18 lines). Performance time was significantly worse for the smallest screen size, but there was no significant difference between the 12 and 18 line screens. However, as Hotson and Shackel (1990) point out, screen related differences may be highly dependent upon the nature of the text being edited, e.g. the number of lines between edits, and the amount of spatial contextual information (number of paragraphs, etc.).

Contrary results were obtained by Neal and Darnell (1984), who examined text editing performance whilst using partial line, partial page (20 line), and full page (60 line) displays (see also Darnell and Neal, 1983). Four target distances were used, 0 lines, 1-3 lines, 7-12 lines, and 21-40 lines. No overall difference in target location time was found between the partial line and the partial page displays. There was a significant interaction, however, such that targets which were on the same line, or targets which were 7-12 lines away took longer in the partial line condition. The lack of significant effect in the longest target condition can be explained by the frequent use of a 'find' command and the infrequent use of cursoring commands. In a comparison of partial page and full page displays Neal and Darnell (1984) found a performance advantage for the full page display, and a target distance x display size interaction which indicated that this difference primarily occurred in the longer target distance conditions. This, however, may be attributable to the fact that each document was only one page long, in the full page condition, and therefore no verification was required with respect to whether the target was currently being displayed.

It would appear that reading tasks may impose very different demands, and are consequently less susceptible to the effects of window size. Duchnicky and Kolers (1983) examined the readability of text in windows of 1, 2, 3, 4, or 20 lines. Subjects (n=10) were required to adjust the rate at which text was scrolled within these five window sizes to provide a comfortable reading speed. There was a significant effect of window size, such that reading speed in the 4 line and 20 line windows was significantly faster than for the 1 and 2 line windows. A reading task was also used by Dillon, Richardson, and McKnight (1990) who compared performance using displays of 20 lines and 60 lines. They found no comprehension or performance time differences, but subjects using the small window changed direction significantly more frequently than those using the large window, suggesting a greater need for reorientation. There was evidence for a subjective preference for larger window sizes.

Cherry, Fischer, Fryer, & Steckham (1989) examined performance upon program editing tasks whilst using help screen presented in three different sizes of windows (full-screen, half-screen, and windowed). Unfortunately the exact size of these windows is not specified. Similarly, no details of sample size are given and it is therefore impossible to assess the weight that should be given to their findings. No significant relationship was found between the size of the help screen and task performance time, the number of help requests, or subject preference. Subjects spent a greater percentage of task time in the help system whilst using the split screen and the windowed help. However, it would appear that this may have been confounded with the process of data entry which was simultaneously available in these conditions.

Perhaps unsurprisingly, it would seem that the effects of window size vary with the nature of the task demand. It would appear that reading text from a CRT is far less sensitive to the effects of window size than search tasks or text editing tasks. Whilst reading comprehension shows no performance increase when comparing windows of four lines and greater, search tasks and text editing performance is asymptotic at window sizes of 12 lines and greater. There is some evidence to suggest that target location (embedded vs non-embedded) is an important factor in search task performance, with the reduced spatial information associated with embedded targets resulting in slower target acquisition. Similarly it would seem that increased distance to target accentuated the effects of window size.

3.1.3 Experimental aims

The experiments of Egan and Gomez (1985), as described in the previous chapter, indicate an assocation between spatial ability and the 'finding' component of word processing. This experiment was designed to examine some factors which may be important in predicting the magnitude of such individual differences in text search. Spatial ability, verbal ability, and expertise were examined in relation to the spatial demands and spatial content of the interface, which were manipulated in terms of command type (page vs scroll), window size (large vs small), target location (embedded vs non-embedded), and distance to target (long vs short). As discussed above, there is evidence to suggest that the page command is more spatially demanding than the scroll command by virtue of a reduction in visual momentum. This is thought to be related to the size of the window being used, with larger windows resulting in greater discontinuity of movement when using the page command. Similarly, spatial processing demands can be predicted to increase in line with the distance to the target. The further the distance to the target the greater the number of commands which are required. Spatial ability is held to be related to the availability of spatial processing resources (Wickens and Weingartner, 1985; Derrick, McCloy, Marshak, Seiler, and Reddick, 1986), and consequently to the level of spatial processing demand. Upon this basis, it can be predicted that low spatial ability individuals will be at a performance disadvantage when using the page command, and that this will be particularly apparent when the window size is large or the distance to target is great. The results of Elkerton et al. (1982; Elkerton and Williges, 1984a) indicate that the location of non-embedded targets is less cognitively demanding than the location of embedded targets. This may be attributable to additional spatial information which is associated with these targets. Similarly, larger windows can be predicted to provide increased spatial information, and therefore be less cognitively demanding than small windows. The relationship between spatial ability and increased spatial information content within the interface is unclear. In the study mentioned above, Vicente and Williges (1988) found that both high and low spatial ability subjects benefited from the provision of additional spatial information. Similarly, in a comparison of text editing performance whilst using a line editor and a display editor, Egan and Gomez (1985) found that the changes in spatial information did not interact with spatial memory. This experiment aims to clarify the association between individual differences in cognitive ability and the spatial information content of the interface during text search. It should be noted that the effects of window size may produce conflicting spatial demands when the page command is used. Whilst a larger window provides additional spatial information it also produces greater movement discontinuity. The relative importance of these effects can be assessed by comparing the effects of window size when using different command methods.

With respect to the relative performance of experts and novices it can be predicted that due to a greater reliance upon controlled processing (Schneider and Shiffrin, 1977) novices will find the task more cognitively demanding and that expertise will therefore interact with the resource requirements of the task. Consequently, it is hypothesised that the increased cognitive demand associated with low visual momentum will disproportionately disadvantage novice subjects. This is consistent with the results of Card et al. (1983) which indicated a comparative performance advantage for the scroll keys during the early stages of skill acquisition, which was reversed following substantial practice. In addition, due to this increased dependence upon controlled processing, it can be predicted that the association between cognitive ability and task performance will be stronger for novices than experts (Fleishman, 1972; Ackerman, 1988).

When considering individual differences in performance a distinction can be drawn between ability and strategy. Individual differences in the completion of spatially demanding tasks have been shown to be subject to substantial differences in performance strategy (Lohman and Kyllonen, 1983). Frequently, within the field of human-computer interaction a variety of command options are provided which accomplish the same task and allow for strategy selection. In the present experiment such alternatives can be seen in the form of the page and scroll commands. Whilst these commands are thought to differ in their associated cognitive demands, for the purposes of initial target location (as opposed to the fine positioning of the cursor required in certain applications) they provide an identical function. Given that individuals will attempt to minimise the level of cognitive workload associated with task performance (Zipf, 1949), and that spatial ability is related to the differing cognitive demands of the page and scroll command, it can be predicted that individuals will adopt a performance strategy which provides a best fit with their available resource capacities (Kyllonen, Lohman, and Woltz, 1984).

In order to examine these hypotheses separate task performance conditions were used in which subjects were restricted to either page or scroll commands for text movement, or they were given a choice of commands (the 'strategy' condition). Performance in the former conditions was assessed solely in relation to ability, and measures of response time and accuracy were recorded. Performance in the latter condition was assessed solely in relation to strategy and a measure of the frequency of selection of each command type was recorded.

3.2 Method

3.2.1 Subjects

64 subjects aged between 18 and 30 years (mean 19.92 years) were recruited from the student population of Aston University. Of this sample 32 were 'novices' (less than 20 hours computer experience) and 32 were 'experts' (more than 100 hours computer experience). Equal numbers of male and female subjects were recruited to all experimental conditions. All subjects were right handed, reported vision which was normal or corrected to normal, and spoke English as their first language.

3.2.2 Measures of individual differences

Subjects initially completed tests of spatial memory (MV2: Ekstrom et al., 1976), spatial visualisation (VZ2 : Ekstrom et al., 1976), verbal ability (Nelson-Denny vocabulary test, 1973), and computer literacy (CALIP : Poplin, Drew, and Gable, 1984).

3.2.3 The experimental task

This task was performed on a PCAT 80286 computer with a Hercules screen. Software was purpose written in C and Assembler with a timing resolution of ± 5 ms. The cursor speed was adjusted to be comparable with that of most commercial word processing packages. Text mode was used in which 25 X 80 characters were displayed on a full screen.

At the beginning of each experimental condition subjects were presented with 'hardcopy' of a passage of text placed on a table to the left of the computer. Within this text a number of target words had been circled in red ink. Exactly the same text was presented on the computer screen. Subjects were required to locate the target words within the text displayed on the computer screen, using the page and scroll commands to manipulate the contents of the screen, and to press the ENTER key once they had done so. It was necessary for target words to be located in sequence moving from the top to the bottom of each printed page. In order to equate the performance demands of the page and the scroll commands, subjects were not required to move the cursor to the exact position of the target word. As long as the target word was displayed on the screen a hit was recorded when the ENTER key was pressed. Successful target location resulted in the target word being highlighted, and the word 'HIT' and the elapsed time being displayed on a status bar at the bottom of the screen. Unsuccessful target location (pressing the ENTER key when the current target word was not displayed upon the screen) resulted in the word 'MISS' being displayed on the status bar.

Three command availability conditions were used. In one condition subjects could only use the page keys, this resulted in text being moved up or down by a number of lines equal to one less than the number currently being displayed on the screen. In another condition subjects were only able to use the scroll keys, this resulted in text being moved up or down by one line at a time. In the third condition subjects had both methods of text movement available and could select a preferred performance strategy. The page only and the scroll only conditions were presented in a counterbalanced sequence (ability conditions). However, the condition in which both commands were available (strategy condition) was always presented last in order that subjects should have equal amounts of practice with each method before being allowed to choose a performance strategy. At the end of each experimental condition subjects completed a measure of self-report workload (NASA-TLX; Vidulich and Tsang, 1986).

Two window sizes were used, one in which 5 lines of text were displayed, and one in which 13 lines of text were displayed. All three command availability conditions were performed for each window size, with window sizes being presented in a counterbalanced sequence. For each condition two pieces of text were presented. The first of these was used for practice purposes in order to allow subjects to familiarise themselves with the task. Target location performance for the second piece of text was recorded. Dependent measures for the 'page only' and 'scroll only' condition were response time per successful first-try target location and the number of unsuccessful first-tries. The dependent measure for the strategy condition (both commands available) was the number of lines of file movement accomplished with each command. This provided an equivalent measure for both the page and scroll commands regardless of window size.

The texts used for this experiment were drawn from the Aston University Library Business Magazine Database, and were matched as far as possible for reading ease using the Flesch formula (Flesch, 1949). Each passage of text used was between 700 and 720 lines long, with 75 characters per line, and contained 66 paragraphs. Target words were four letters long (in order to control for target size), contained one syllable, and were matched as far as possible for frequency of occurrence within the passage (cf. Swierenga, 1990). All of the target words used occur in normal language with a frequency greater than 50 times per million as given by Thorndike and Lorge (1972). The same word was never used twice within one passage of text. Each text passage was marked with 48 targets of which 24 were 'long' targets (between 13 and 36 lines from the previous target) and 24 were 'short' (between 1 and 12 lines from the previous target). This variable will subsequently be referred to as 'target distance'. Average distance to long targets was 22.25 lines, and average distance to short targets was 6.98 lines. For both long and short targets, 12 targets were designated as 'inner' targets (within the body of a paragraph) and 12 were designated as 'outer' targets. Of the outer targets, equal numbers were placed at the top, bottom, left and right of the paragraph. This variable will subsequently be referred to as 'target location'. Texts were allocated to experimental conditions in a random sequence.

3.3 Results

3.3.1 Distribution of scores

Error data did not generally conform to a normal distribution. Subjects made relatively few errors, and therefore the data was positively skewed. Attempts at data transformation did not improve this situation. However, the method of statistical analysis used, ANOVA, is considered to be relatively robust with respect to non-normality of distribution (Hinkle, Wiersma, and Jurs, 1979).

3.3.2 The effects of command type, window size, target distance, and target location.

The effects of command type (page or scroll), window size (long or short), target position (inner or outer), and distance to target (long or short) were examined for reaction time and proportion of errors using a 2x2x2x2 repeated measures analysis of variance.

3.3.2.1 Response times

Mean response times for all experimental cells are presented in Table 3.01. There was no significant difference in reaction time when comparing performance using the page command or the scroll command (see Table 3.02). However, subjects located targets faster in the large window than in the small window ($F(1,63)=172.68$; $p<.001$), located short targets faster than long targets ($F(1,63)=1224.10$; $p<.001$), and located outer targets faster than inner targets ($F(1, 63)=39.35$; $p<.001$).

Table 3.01 : Mean response times (secs) for all experimental cells

		Page		Scroll	
		Large	Small	Large	Small
Long	In	6.456	8.803	6.449	9.127
	Out	6.060	7.727	6.155	8.219
Short	In	3.217	4.394	3.081	4.321
	Out	2.966	3.964	2.895	3.979

Table 3.02 : Response times (secs) : Means and standard deviations for the main effects of command type, window size, target distance, and target location.

	Mean	SD
Page command	5.448	1.035
Scroll command	5.528	0.846
Large window	4.660	0.948
Small window	6.317	1.073
Long targets	7.374	1.210
Short targets	3.602	0.669
Inner targets	5.731	1.017
Outer targets	5.245	0.836

There was a significant interaction between window size and location of edit ($F(1, 63)=12.74$; $p=.001$) such that the performance advantage for outer targets over inner targets was comparatively greater when using the small window (see Table 3.03).

Table 3.03 : Interaction between window size and target location for response time (secs).		
	Large window	Small window
Inner targets	4.801	6.661
Outer targets	4.519	5.972

There was also a significant interaction between window size and distance to target ($F(1, 63)=35.35$; $p<.001$) such that the difference in performance between long and short targets was comparatively greater when using the small window (see Table 3.04).

Table 3.04 : Interaction between window size and distance to target for response time (secs).		
	Large window	Small window
Long targets	6.280	8.469
Short targets	3.040	4.165

The interaction between target location and distance to target was significant ($F(1, 63)=6.53$; $p<.05$) with cell means (see Table 3.05) indicating a greater comparative difference between inner and outer targets when locating long targets. It is possible that this interaction may have been in part attributable to the increased number of page commands required in the small window conditions. However, the data do not bear this out, with the difference being most pronounced in the scroll command conditions.

Table 3.05 : Interaction between target location and distance to target for response time (secs).		
	Long targets	Short targets
Inner targets	7.709	3.753
Outer targets	7.040	3.451

3.3.2.2 Proportion of errors

The mean proportion of errors for all experimental cells are presented in Table 3.06. There was a main effect of command type ($F(1, 63)=16.29$; $p<.001$). Subjects made fewer errors in the scroll condition than in the page condition (see Table 3.07). There was also a significant effect of distance to target ($F(1, 63)=19.74$; $p<.001$) such that subjects made more errors whilst locating long targets than when locating short

targets. The effects of target location (inner or outer) and window size were not significant.

Table 3.06 : Mean proportion of errors for all experimental cells					
		Page		Scroll	
		Large	Small	Large	Small
Long	In	.099	.074	.081	.053
	Out	.079	.073	.065	.049
Short	In	.035	.065	.044	.043
	Out	.074	.056	.035	.033

Table 3.07 : Proportion of errors: Means and standard deviations for the main effects of command type, window size, target distance, and target location.		
	Mean	SD
Page command	.069	.04
Scroll command	.050	.04
Large window	.064	.05
Small window	.056	.05
Long targets	.072	.05
Short targets	.048	.03
Inner targets	.062	.04
Outer targets	.058	.04

There was a significant interaction between window size and distance to edit (F(1, 63)=4.02; p<.05) such that for short targets there was very little effect of window size. However, when long targets were being located fewer errors were made in the small window than in the large window (see Table 3.08).

Table 3.08 : Interaction between window size and target distance for the proportion of errors made.		
	Large window	Small window
Long targets	.081	.062
Small targets	.047	.049

The three-way interaction between window size, target location and distance to edit was significant (F(1, 63)=5.30; p<.05) with short inner targets being located more

accurately in the large window whereas all other targets were located more accurately in the small window (see Table 3.09). Additionally, in the large window conditions, a greater proportion of errors was made when locating long inner targets than long outer targets, whereas this pattern was reversed with respect to short inner and short outer targets.

Table 3.09 : Interaction between window size, target distance, target location for the proportion of errors made.

Distance	Location	Large	Small
Long	Inner	.090	.064
	Outer	.072	.061
Short	Inner	.040	.054
	Outer	.055	.045

3.3.3 The effects of expertise

The effects of expertise were examined using a 2 x 2 x 2 x 2 x 2 (expertise x command x window size x target distance x target location) ANOVA.

3.3.3.1 Response times

There was no main effect of expertise upon response time, the only effects of expertise being manifested in a complex four-way interaction between expertise, command, window size, and target position ($F(1, 62)=15.69$; $p<.001$). Novices performed more quickly than experts when using the page command for inner targets in the large window and outer targets in the small window, and when using the scroll command to locate inner targets in the small window (see Table 3.10). Experts performed more quickly in all other conditions.

Table 3.10 : Interaction between expertise, command type, window size, and target location for response times (secs).

		Novice		Expert	
Window	Location	Page	Scroll	Page	Scroll
Large	Inner	4.776	4.840	4.896	4.690
	Outer	4.578	4.548	4.448	4.502
Small	Inner	6.694	6.606	6.503	6.842
	Outer	5.659	6.325	6.032	5.872

3.3.3.2 Proportion of errors

With respect to error rates, there was a main effect of expertise ($F(1, 63)=6.30$; $p<.05$) with novices generating a greater number of errors than experts (see Table 3.11).

Table 3.11 : Means and standard deviations for the proportion of errors for the novice and expert subject groups		
	Mean	SD
Novice	.072	.042
Expert	.048	.031

There was a significant interaction between expertise, target location, and distance to target ($F(1, 62)=4.85$); $p<.05$) such that the difference in error rates between novices and experts was comparatively greater for long outer targets and short inner targets (see Table 3.12). There was also a four-way interaction between expertise, command method, target location, and distance to target ($F(1,62)=4.88$; $p<.05$). This appears to be primarily due to novices being more error prone when locating long, outer targets than long, inner targets when using the page command. However, this pattern is reversed for experts. Novices also make more errors when locating short outer targets than short inner targets when using the page command. In all other conditions a greater proportion of errors are made when locating inner targets.

Table 3.12 : Mean proportion of errors for novice and expert subject groups for window size, distance to target, and target location conditions.							
		Novice			Expert		
Distance	Location	Page	Scroll	Mean	Page	Scroll	Mean
Long	Inner	.084	.088	.086	.090	.049	.070
	Outer	.106	.075	.091	.047	.041	.044
Short	Inner	.063	.051	.057	.038	.036	.037
	Outer	.069	.042	.056	.062	.026	.044

3.3.4 Computer literacy

The relationship between computer literacy and performance was examined using a similar design, with subjects being allocated post-hoc to high and low computer literacy groups upon the basis of a median split of CALIP (Poplin, Drew, and Gable, 1984) scores. Computer expertise was not included in these analyses because the strong association with computer literacy would have produced highly unequal cell sizes (see Table 3.13).

Table 3.13 : Mean scores and results of t-test for novice and expert groups upon computer literacy test.

	Novice	Expert	t	p
Computer literacy	23.84	29.56	-4.90	<.001

3.3.4.1 Response times

There was no main effect of computer literacy. There was a significant four-way interaction between computer literacy, command, window size, and target location for response time ($F(1, 62)=4.18$; $p<.05$), such that subjects high in computer literacy performed faster in all page conditions, and particularly so for outer targets in the large window, but performed more slowly than low computer literacy subjects for inner targets in the small window whilst using the scroll command (see Table 3.14). High computer literacy subjects located targets more quickly using the page command than the scroll command, whilst the reverse was true for low literacy subjects, with the exception of outer targets in the small window. There was also a five-way interaction between computer literacy, command method, window size, distance to target and target location ($F(1,62)=4.12$: $p<.05$), the complexities of which were beyond reasonable interpretation.

Table 3.14 : Interaction between computer literacy, command type, window size, and target location for response time (secs).

		Low LIT		High LIT	
Window	Location	Page	Scroll	Page	Scroll
Large	Inner	4.976	4.829	4.7004	4.705
	Outer	4.745	4.589	4.295	4.465
Small	Inner	6.728	6.622	6.476	6.820
	Outer	5.857	6.251	5.834	5.956

3.3.4.2 Proportion of errors

With respect to the proportion of errors, there was a significant interaction between computer literacy and window size ($F(1,62)=4.94$: $p<.05$), such that subjects low in computer literacy performed more accurately in the large window than the small window, and were marginally more accurate than subjects high in computer literacy for the large window. Subjects high in computer literacy, however, performed much more accurately in the small window (see Table 3.15).

Table 3.15 : Interaction between computer literacy and window size for the proportion of errors.		
Window	Low LIT	High LIT
Large	.060	.068
Small	.069	.044

There was also a four-way interaction between computer literacy, command, window size and target location ($F(1, 62)=5.67 : p<.05$) such that subjects high in computer literacy were less accurate when locating outer targets in the large window whilst using the page command as compared to the scroll command, but the reverse was true for inner targets in the large window (see Table 3.16).

Table 3.16 : Interaction between computer literacy, command type, window size and target location for the proportion of errors.

		Low LIT		High LIT	
Window	Location	Page	Scroll	Page	Scroll
Large	Inner	.069	.057	.066	.069
	Outer	.065	.052	.089	.049
Small	Inner	.079	.069	.061	.029
	Outer	.081	.047	.049	.036

3.3.5 Cognitive ability

A series of unrelated t-tests was used to examine differences between novice and expert subject groups upon each of the cognitive ability measures used. The results of these tests are presented in Table 3.17. As can be seen, whilst there is a performance advantage for the expert group on all measures, there are no significant differences. Table 3.18 presents the correlation matrix for these measures taken over the entire sample. Correlations are all positive, but are uniformly modest and non-significant.

Table 3.17 : Mean scores and results of t-tests for novice and expert groups upon cognitive ability tests.				
	Novice	Expert	t	p
Spatial visualisation	12.05	13.20	-1.53	ns
Spatial memory	18.94	19.69	-.74	ns
Nelson-Denny vocab.	69.94	76.97	-1.92	ns

Table 3.18 : Correlation matrix for cognitive ability measures		
	Spatial memory	Vocabulary
Spatial visualisation	.09 ns	.21 ns
Spatial memory		.13 ns

In order to examine the effects of cognitive ability in relation to interface and target location factors, subjects were allocated, post-hoc, to high and low ability groups for each measure based upon a median split of the relevant scores. 2 x 2 x 2 x 2 x 2 x 2 (cognitive ability x expertise x command x window size x distance to target x target location) ANOVAs were used to analyse performance for each cognitive ability and each dependent measure.

3.3.5.1 Spatial visualisation

The main effect of spatial visualisation with respect to response time was not significant. However, there was a significant interaction ($F(1,60)=4.76 : p<.05$) between spatial visualisation, window size, and target location (see Table 3.19). Subjects in the high spatial visualisation group performed more quickly for both inner and outer targets in the large window, with the difference being greater for outer targets. There was very little difference between the response times of high and low spatial ability groups when locating inner targets in the small window, but low spatial visualisation subjects performed more quickly when locating outer targets whilst using this window (see Table 3.19).

Table 3.19 : Interaction between spatial visualisation, window size, and target location for response time (secs).			
Window size	Location	Low SV	High SV
Large	Inner	4.872	4.756
	Outer	4.659	4.400
Small	Inner	6.676	6.658
	Outer	5.821	6.153

There was also a significant four-way interaction between spatial ability, window size, distance to target and target location (F(1,60)=4.04 : p<.05) which indicates that the above mentioned response time advantage for the low spatial group when locating outer targets using the small window was only apparent when the distance to target is long (see Table 3.20).

Table 3.20 : Interaction between spatial visualisation, window size, target distance, and target location for response time (secs).

		Low SV		High SV	
Window	Location	Long	Short	Long	Short
Large	Inner	6.543	3.200	6.388	3.123
	Outer	6.360	2.957	5.878	2.921
Small	Inner	8.947	4.404	8.977	4.339
	Outer	7.651	3.991	8.332	3.975

There were no significant main or interactive effects of spatial visualisation with respect to accuracy of performance.

3.3.5.2 Spatial memory

The main effect of spatial memory, with respect to response times, was not significant. However, there was a significant interaction between spatial memory and command for reaction time (F(1, 60)=4.12 : p<.05) such that subjects high in spatial memory performed more slowly when using the page command and more quickly when using the scroll command (see Table 3.21).

Table 3.21 : Interaction between spatial memory ability group and command type for response times (secs).

	Low SM		High SM	
Command	Mean	SD	Mean	SD
Page	5.304	1.210	5.584	0.834
Scroll	5.572	0.922	5.487	0.780

There was also an interaction between spatial memory and target location (F(1, 60)=4.17 : p<.05) such that subjects high in spatial memory performed more slowly for inner targets but marginally faster for outer targets (see Table 3.22). This interaction was further clarified in a three-way interaction between spatial memory, target location, and distance to target ($F(1, 60)=7.19$) : $p<.01$) such that subjects low in spatial memory located long, inner targets more quickly but located long, outer targets more slowly. They were also marginally faster for both inner and outer short targets. The five-way interaction between spatial memory, expertise, command method, window size, and target location was significant ($F(1,60)=4.82$: $p<.05$), however, no parsimonious interpretation was possible for the complex pattern of cell means. With respect to the proportion of errors, there was a main effect of spatial memory ($F(1, 60)=7.07$: $p=.01$) such that subjects high in spatial memory performed more accurately (see Table 3.23).

Table 3.22 : Response times (secs) for high and low spatial memory groups in both target distance and target location conditions.

Location	Distance	Low SM	High SM
Inner	Long	7.482	7.977
	Short	3.746	3.773
	Mean	5.614	5.875
Outer	Long	7.146	6.958
	Short	3.421	3.500
	Mean	5.284	5.229

Table 3.23 : Proportion of errors for subjects in high and low spatial memory groups.

	Mean	SD
Low spatial memory	.074	.044
High spatial memory	.047	.028

There was an interaction between spatial memory and distance to target ($F(1, 60)=11.87$: $p=.001$) indicating that high spatial memory subjects performed slightly more accurately than subjects low in spatial memory when locating short targets. However, low spatial memory subjects became considerably more inaccurate when locating long targets, whilst there was little change in the performance of high spatial memory subjects (see Table 3.24).

Table 3.24 : Interaction between spatial memory group and distance to target for proportion of errors

	Low SM	High SM
Long	.091	.051
Short	.051	.044

There was also a significant interaction between spatial memory, command type and target location (F(1, 60)=10.38 : p<.01) such that, whilst high spatial memory subjects were more accurate in all cells, this difference was greater for outer targets whilst using the page command but greater for inner targets whilst using the scroll command (see Table 3.25).

Table 3.25 : Interaction between spatial memory, command type, and target location for proportion of errors

Command	Location	Low SM	High SM
Page	Inner	.077	.060
	Outer	.087	.052
Scroll	Inner	.072	.037
	Outer	.048	.042

3.3.5.3 Verbal ability

The main effect of verbal ability with respect to response time was not significant. However, there was a significant interaction between verbal ability and target location (F(1, 60)=5.50 : p<.05) with high verbal ability subjects locating outer targets more quickly than low verbal ability subjects, but the reverse applying to the location of inner targets (see Table 3.26).

Table 3.26 : Interaction between verbal ability and target location for response time (secs).

	Low Verbal	High Verbal
Inner targets	5.689	5.772
Outer targets	5.383	5.107

There was also a significant three-way interaction between verbal ability, expertise, and window size (F(1,60)=4.21 : p<.05). For the novice group, high verbal ability subjects performed more quickly than low verbal ability subjects in the small window and more slowly in the large window. However, for the expert group, high verbal ability was associated with better performance in the large window and there was little effect of verbal ability for the small window (see Table 3.27). This position was further clarified by a significant four-way interaction between verbal ability, expertise, command method, and window size (F(1,60)=6.74 : p<.05) which indicated that the above mentioned relationship between verbal ability, expertise and window size was particularly evident when using the scroll command.

Table 3.27 : Response times (secs) for novice and expert subject groups, high and low in verbal ability, performing in each window size and command condition.

		Low Verbal		High Verbal	
Window	Command	Novice	Expert	Novice	Expert
Large	Page	4.650	4.715	4.712	4.639
	Scroll	4.433	4.763	5.029	4.467
	Mean	4.542	4.739	4.871	4.553
Small	Page	6.335	6.325	5.972	6.222
	Scroll	6.755	6.308	6.094	6.395
	Mean	6.545	6.317	6.033	6.309

There was a significant four-way interaction between verbal ability, expertise, command method, and target location ($F(1,60)=5.02 : p<.05$), in which low verbal novices performed more quickly than low verbal experts when locating outer targets using the page command, and inner targets using the scroll command. High verbal novices only performed more quickly than high verbal experts when locating inner targets using the page command (see Table 3.28).

Table 3.28 : Interaction between verbal ability, expertise, command type and target location for response time (secs).

		Low Verbal		High Verbal	
Command	Location	Novice	Expert	Novice	Expert
Page	Inner	5.820	5.550	5.626	5.815
	Outer	5.166	5.490	5.058	5.045
Scroll	Inner	5.611	5.762	5.867	5.770
	Outer	5.576	5.310	5.257	5.092

There were no significant main or interactive effects of verbal ability upon performance accuracy.

3.3.6 Differences in self-report workload

Differences in self-report workload relating to command type and window size were examined using a 2 x 2 repeated measures ANOVA. The main and interactive effects of command type were not significant. However there was a significant difference relating to window size ($F(1, 63)=17.28$: $p<.001$) with self-report workload being higher in the small window condition (see Table 3.29).

Table 3.29 : Self-report workload for both command type and window size conditions.

	Page	Scroll	Mean
Large	45.04	43.42	44.23
Small	48.69	49.95	49.32
Mean	46.87	46.69	46.78

Additional factors were added to this design, as for the earlier analyses, to examine individual differences in self-report workload in relation to expertise, computer literacy, spatial visualisation, spatial memory, and verbal ability.

There were no main or interactive effects of expertise, computer literacy, spatial visualisation, or verbal ability upon the level of self-report workload.

There was a significant interaction between spatial memory and window size ($F(1, 62)=7.31$: $p<.01$) such that the level of self-report workload was very similar for high and low spatial memory groups when using the large window. However, high spatial memory subjects reported comparatively greater workload when using the small window (see Table 3.30).

Table 3.30 : Interaction between spatial memory groups and window size for self-report workload.

	Low SM	High SM
Large Window	44.64	44.61
Small Window	46.48	52.82

3.3.7 Strategy conditions

File movement distance was examined for the strategy conditions using a 2 x 2 x 2 x 2 (command type x window size x distance to target x target location) repeated measures ANOVA. Cell means are shown in Table 3.31.

Table 3.31 : File movement distance using page and scroll commands in the strategy condition

		Page		Scroll	
		Large	Small	Large	Small
Long	In	159.56	194.13	145.31	108.34
	Out	150.19	189.00	134.52	106.08
Short	In	48.94	79.69	46.78	43.30
	Out	53.81	61.75	45.69	39.70

File movement distance was greater for inner targets than for outer targets ($F(1,63)=5.26$: $p<.05$) and was also greater when locating long targets than short targets ($F(1,63)=1141$: $p<.001$). The main effects of command type and window size were not significant.

Initially, in order to examine individual differences in command selection strategy, 2 x 2 (cognitive ability x expertise) ANOVAs were used to determine differences in file movement efficiency. None of the main or interactive effects for any of the included variables were significant. Having established this, performance strategy was considered by using the ratio between the number of lines moved with a particular command and the total file movement as a dependent measure. This gave an index of strategy which was relatively independent of ability. 2 x 2 x 2 x 2 x 2 x 2 (cognitive ability x expertise x command x window size x target distance x target location) ANOVAs were then used to examine strategy, with the only effects of interest being those which included the command type factor.

The only significant interaction involving the command factor was a five-way interaction between spatial visualisation, expertise, command type, window size, and target location ($F(1,60)=4.71$: $p<.05$). The cell means for this interaction are shown in Table 3.32.

Table 3.32 : Interaction between spatial visualisation, expertise, command type, window size, and target location for the proportion of lines moved with each command						
			Low SV		High SV	
			Novice	Expert	Novice	Expert
Page	Large	Inner	.121	.118	.163	.115
		Outer	.127	.117	.150	.111
	Small	Inner	.166	.158	.135	.199
		Outer	.151	.161	.134	.175
Scroll	Large	Inner	.123	.134	.095	.133
		Outer	.117	.130	.088	.119
	Small	Inner	.098	.096	.121	.073
		Outer	.097	.086	.114	.075

3.4 Discussion

There was no significant difference in response time or self-report workload between the page and scroll command conditions. However, performance using the scroll command was significantly more accurate than performance using the page command. This is consistent with the predicted effects of visual momentum indicating reduced cognitive demand associated with the scroll command.

The effects of window size, distance to target and target location were all broadly consistent with the experimental hypotheses. There were significant response time differences for each of these variables, with performance being superior in the large window as opposed to the small window, for short targets as opposed to long targets, and for non-embedded targets as opposed to embedded targets. Accuracy data also generally followed this pattern, and the interactions between variables for both speed and accuracy suggested that the effects of interface factors were additive. However, there was an interaction between window size and distance to target such that performance accuracy was reduced in the large window for long targets. This may indicate a complex speed accuracy trade off such that long targets are located more quickly but less accurately in the large window.

The significant effect of window size supports the findings of Elkerton and Williges (1984a) and Hotson and Shackel (1990) and indicates that a reduction in window size to 5 lines will detrimentally influence the performance of text search tasks. There was also a significant effect on self-report workload with greater workload experienced when using the small window. This not only illustrates the potential viability of such a measure as a means of assessing interface design variables, but also indicates the stress inducing possibilities of small window sizes. This would appear to apply particularly to users with high spatial memory ability, as discussed below. Whilst the effects of window size upon file movement efficiency were non-significant, there was a tendency for file movement to be greater for the small window than the large window. This is contrary to the findings of Elkerton et al. (1982). However it supports the experimental hypothesis that smaller windows provide less spatial information and are therefore subject to greater orientation difficulties. The significant effect of target location is consistent with the findings of Elkerton and Williges (1984a), with embedded targets being more difficult to locate. This indicates the importance of layout in the presentation of textual material, and emphasises the need for frequent spatial landmarks.

Overall, these results suggest that the experimental manipulation of command type, window size, target location, and distance to target were successful and produced the predicted changes in cognitive demand. The following sections will now consider these manipulations in relation to individual differences in performance.

3.4.1 Expertise

Whilst there was no significant difference between the novice and the expert groups for response times or self-report workload, the expert group were significantly more accurate in locating targets. This suggests that the experimental manipulation of expertise was successful, and that the expert group were indeed more skilled than the novice group. Whilst the two-way interaction between expertise and command type was not significant for any of the dependent measures, there was a significant four-way interaction for response time which included expertise, command type, window size, and target location. Experts were found to perform quicker when using the page command than the scroll command to locate outer targets in the large window, and were also faster than novices for this condition. In the small window, however, they were faster when using the scroll command than the page command to locate outer targets, and they were faster than novices in this condition. A performance advantage for experts when using the page command in the large window is consistent with the hypothesised interaction between expertise and cognitive demand. There was also a significant three-way interaction between expertise, target location and distance to target for the proportion of errors which indicated that experts were more accurate than novices when locating long, outer targets. The four-way interaction which also included command type indicated that this was particularly true when using the page commands. Again, this is consistent with an interactive effect between expertise and the hypothesised high cognitive demand associated with using the page command in the large window.

With respect to computer literacy there was a two-way interaction between literacy and window size for the proportion of errors. Whilst high literacy subjects were more accurate than low literacy subjects in the large window, low literacy subjects were marginally more accurate in the small window. Cell means for response times in these conditions suggest that the high literacy group may be engaging in a speed accuracy trade-off when using the large window. The four-way interaction, including command type, window size, and target location for the proportion of errors was such that high literacy subjects were more accurate than low literacy subjects except when using the scroll command to locate inner targets in the large window and when using the page command to locate outer targets in the large window. This latter difference was the more substantial and suggests that high literacy subjects trade speed for accuracy when cognitive demand is high.

Whilst these results do not indicate strong interactive effects they broadly support the predicted association between expertise and interface factors, with novices being at a performance disadvantage in conditions of high cognitive demand.

3.4.2 Spatial ability

Contrary to the experimental prediction and the results of Vicente et al. (1987; Vicente and Williges, 1988) there was no main effect of spatial visualisation for response time, accuracy or workload data. However, there was a significant three-way interaction between spatial visualisation, window size and target location, and a four-way interaction which also included distance to target. The pattern of cell means indicated that whilst high spatials perform more quickly than low spatials in the large window, they perform markedly more slowly for outer targets in the small window. The four-way interaction indicated that this pattern was the result of performance for long targets. The superior performance of high spatials when using the large window would suggest that these subjects are more able to use the increased spatial contextual information provided in this condition. The fact that the magnitude of this effect was similar for the page and scroll commands indicates that it cannot be attributed to the effects of greater discontinuity, i.e. reduced visual momentum when using the page command for the large window. These results appear to indicate that, whilst the use of the page command is associated with increased cognitive demand, this does not extend to a spatial processing requirement. The finding that high spatial ability subjects are slower to locate outer targets in the small window may indicate the importance of spatial context to these subjects. Whilst outer targets provide a spatial landmark, in the smaller window size there are few other spatial reference points. It may be that high spatial subjects continuously use landmark information as a basis for their attempts at target location, even when other factors prove detrimental to this performance strategy. The only indication of command strategy differences for subjects of differing spatial visualisation ability was in an extremely complex five-way interaction for file movement distance involving command type, expertise, window size and target location. However, on the basis that the expert group will have had greater opportunity for strategy development, it is of interest to compare the cell means for this group alone. Whilst the file movement distance for high and low spatials is similar in the large window, in the small window low spatials prefer to use the scroll command whilst high spatials prefer to use the page command. This latter strategy selection is particularly apparent and suggests that these subjects were not relying upon a 'default method' (Card, Moran, and Newell, 1983) or applying a 'compensation schema' (Young and MacLean, 1988) irrespective of interface variables, but were making an active strategy decision. It is surprising that command strategy differences should be apparent for the small window and not the large window, when the difference between the page and scroll commands is greater. However, ability and strategy are not always well matched in the performance of spatially demanding tasks (Cooper and Mumaw, 1985).

Consistent with the experimental hypotheses, there was a main effect of spatial memory with respect to performance accuracy such that subjects in the high spatial memory group made fewer errors. This appeared to be primarily attributable to the location of long targets. There was also a significant interaction between spatial memory and command type for response time, with the high spatial memory group performing faster than the low spatial memory group in the scroll command conditions, but slower when using the page command. This is contrary to the experimental hypothesis which predicted that the decreased visual momentum associated with the page command would favour subjects with high spatial ability. There are no differences in accuracy or self-report workload data which clarify the situation, and no evidence that these relative command response time differences translated into strategy differences when subjects were given the choice of command type. The significant interaction between spatial memory and target location was such that high spatial memory subjects were slightly faster for outer targets but performed slower than low spatial memory subjects for inner targets. This is consistent with increased spatial information within the interface either being utilised more efficiently by high spatial memory subjects, or being of greater importance in their performance strategy. However, it also indicates that subjects in the high spatial memory group are less skilled at locating targets when this information was not present. A relative disadvantage for high spatial memory subjects in low spatial information conditions is also supported by the significant interaction between spatial memory and window size with respect to self-report workload. Both groups reported similar workload in the large window condition, but the high spatial memory group reported comparatively greater workload in the small window condition.

Considered together, these results provide no evidence to support the hypothesised association between file movement discontinuity and spatial ability. However, it would appear that high spatial ability subjects are at a performance advantage when the spatial information content of the interface is high, although, when spatial information content is low this position may be reversed due to performance strategies associated with the use of landmark information. This may result in high spatial ability subjects experiencing comparatively greater workload.

3.4.3 Verbal ability

There was a significant interaction between verbal ability and target location such that high verbal subjects were quicker at locating outer targets. It is unclear why verbal ability should be related to landmark information in this way. A strong association between verbal ability and spatial ability would provide one explanation, however this was not the case. There was also a significant interaction between verbal ability, expertise and window size such that, for novices, high verbal ability was associated with slower performance in the large window but markedly quicker performance in the small window, whilst for the expert group high verbals were slightly faster in the large window and performance was very similar in the small window. These effects were more pronounced for the scroll command. There would appear to be no convincing rationale for these results. If speed of lexical access (Hunt, 1978) were to be a factor it would seem probable that high verbals would demonstrate an advantage for inner targets where lexical content may be presumed to play a more important role in target location. These results may be clarified by future experiments using a varied match between the format of the hardcopy and the format of the text displayed on the CRT. If text is reformatted on the computer to use wider or narrower margins the demand for verbal processing in the form of text scanning could be predicted to increase whilst spatial processing demands would decrease. The fact that verbal ability was more strongly related to performance for novices than for experts is consistent with the model of individual differences in skill acquisition proposed by Ackerman (1988).

3.4.4 Conclusions

In summary, the hypotheses relating to the effects of command type, window size, distance to target and target location were broadly supported with performance benefiting in conditions of high spatial information content, and when interface discontinuity was minimised. Detrimental effects upon performance were apparent when window size was reduced to five lines. Self-report workload was found to be sensitive to such interface changes and to the interactive effects of individual differences. Whilst the effects were not strong, an association was apparent between expertise and cognitive demand, with novices being at a disadvantage when demand was high. No support was obtained for the hypothesised association between spatial ability and visual momentum. This may be attributable to insufficient discontinuity of movement in the present experiment. Future experiments might compare keyword text search to the scroll command in order to achieve a wider disparity of visual momentum. Spatial ability was found to interact with the spatial information content of the interface. High spatial ability subjects performed better than low spatial ability subjects when there was high spatial information content within the interface. However, when spatial information was low this pattern was reversed with high spatials being at a performance disadvantage and experiencing greater self-report workload. This may be attributable to performance strategies which focus upon the use of landmark information, and which become inefficient when it is not available.

Chapter 4

The 'Finding' Component: Menu and Network Navigation

4.1 Introduction

The navigation of networks and menus is an important feature of many computer-based tasks. Whilst such demands are particularly common to the process of information retrieval, other application packages such as word processors and spreadsheets also contain similar performance components. Parallels can also be drawn with the mental representation of program structures and the navigation of program facilities (cf. Canter, Rivers, and Storr, 1985). This chapter presents five experiments which examine individual differences in the navigation of menus and networks. The first two of these experiments are concerned with the relative contribution of spatial and semantic information to this activity. The interactive effects of spatial and verbal ability are considered, and overall conclusions are drawn. The final three experiments were pilot studies which investigated age and working memory as predictors of navigation performance. They have methodological similarities with the first two experiments, and may be of some interest to the reader. However, each is presented in isolation and only as a brief report.

This initial introduction presents an overview of literature pertinent to such navigational tasks, and includes discussions upon the mental representation of spatial and verbal information, attention resource demands, and 'real world' navigation.

4.1.1 The mental representation of spatial and verbal information

The nature of mental representation has been the subject of an ongoing and seemingly unresolvable debate, with a pivotal point of contention being the existence of a spatial processing code. Proponents of one side of this argument hold that a propositional, abstract representation is the only fundamental processing code, and it is from this foundation that analogue representation is subsequently generated (Pylyshyn, 1973). Proponents of the other side of the argument hold the view that imagery is the product of a fundamental analogue processing code (Kosslyn, 1980; Paivio, 1991). A comparatively recent entry into this arena is the concept of the mental model (Johnson-Laird, 1983). The mental model, as described by Johnson-Laird, may be based upon propositional, rule-based representations, but it is an analogical, determinate representation of a world state. This model may incorporate imagery although this is not a fundamental requirement. However, it is essentially spatial in nature. Coexistantly, a propositional or rule-based representation is "the representation of a function from states of affairs to truth values" (Johnson-Laird, 1985, p. 89). It is true or false, and "Reasoning in this case proceeds without a mapping of premises into examples." (Galotti, Baron, and Sabini, 1986, p.17). Whilst subject to some variation of emphasis (cf. Eysenck and Keane, 1990) such processing can be held to be essentially verbal in nature.

The concept of the mental model has been widely used within the human-computer interaction (HCI) /human factors (HF) literature, with attention being given to questions of mental model ownership (Norman, 1986), rules which map action to response (Payne and Green, 1986), the effects of particular training regimes (Carroll, 1984; Borgman, 1986), and so on. However, within this context little attention has been given to the underlying cognitive dimensions of the phenomenon. The assumption frequently appears to be that the mental model is a common denominator which may be used intuitively, and which is adequately understood. Further examination of the literature, however, suggests that this concept is being used in a variety of context dependent ways, and that there is a consequent divergence between the HCI/HF and the cognitive psychological approaches. An interesting comparison can be seen between the perspectives of Rouse and Morris (1986) who approach the topic from a HF standpoint, and Johnson-Laird (1983) who approaches the topic from a cognitive psychological standpoint (see also Rutherford and Wilson, 1991). Rouse and Morris acknowledge the danger of the mental model being synonymous with the knowledge-base of the individual. However, from a cognitive psychological perspective tougher constraints are placed upon the concept. For example, Johnson-Laird regards computability as being an essential ingredient, whilst Rouse and Morris recognise no such necessity. An associated contrast concerns what Johnson-Laird regards as the fundamental analogue nature of mental models or what may be termed the principle of 'constructivism' (1983, p. 398). Rouse and Morris impose no such limitation and include abstract representation within their definition. This suggests that propositional representation may be legitimately included within the HCI/HF definition, whereas from a cognitive psychological standpoint such rule-based representations are excluded. Unsurprisingly the HCI/HF approach to the mental model tends to be systems oriented. Carroll and Olson (1988, p. 51), for example, suggest that a mental model is "a rich and elaborate structure, reflecting the user's understanding of what the system contains, how it works, and why it works that way". Rasmussen (1990, p. 42) proposes that the concept of the mental model "is used to characterise features of the resident knowledge base, representing properties of the task environment which can serve the planning of activities and the control of acts when instantiated and activated by observation of the actual state of affairs". Many applied examples of the use of the concept in this way can be found in the HF literature (cf. Gentner and Stevens, 1983).

It should be noted that in order to avoid possible confusion arising from the different uses of the term mental model, as described above, the term mental representation will be taken to include both spatial and verbal processing. The term mental model will only be used from the cognitive psychological standpoint.

The nature of individual differences in the process of mental representation has received comparatively little attention. There is, however, evidence to support the view that stable individual differences do exist. Egan and Grimes-Farrow (1982) found individual differences in the performance of series reasoning problems, such that subjects could be classified as abstract or concrete in their mental representations of problems. Mani and Johnson-Laird (1982) concluded that two sorts of processing occurred in the mental representation of spatial descriptions, one which was propositional and linguistic in nature, and one which was a mental model, primarily spatial in nature. Similarly, Galotti, Baron, and Sabini (1986) found individual differences in the use of model or rule-based methods of solving syllogistic reasoning problems. Such differences in representation have also been related to individual differences in cognitive ability. Hunt (1978; McLeod, Hunt, and Matthews, 1978) presents evidence which suggests that representational strategy is related to spatial and verbal ability. This experiment used a sentence comprehension task in which subjects were given a verbal description of the spatial relationship between two objects. Based upon a model of verbal processing proposed by Carpenter and Just (1975) response times to pictures of these items indicated that verbal ability, and particularly spatial ability, were related to the form in which the problem was mentally represented, with subjects of high spatial ability tending to use analogue representation and subjects of high verbal ability tending to use propositional representation. Sternberg and Weil (1980) investigated individual differences in the representation of linear syllogisms. Their results indicated that individual differences in spatial and verbal ability may be more strongly related to the efficiency with which spatial or verbal strategies are used than to the process of strategy selection. This is consistent with a small study (n=24) conducted by Cooper and Mumaw (1985) in which spatial ability was examined in relation to the performance of a task requiring the resolution of different isometric drawings. Two strategies were identified, a 'constructive' or model-based strategy, and an 'analytic' or feature comparison strategy. Ability did not appear to be strongly related to strategy selection. However, an ability related strategy was found to be of particular importance to the performance of low spatial ability subjects, who were more effective with an 'analytic' strategy. High spatial ability subjects performed well regardless of strategy. It should be noted, however, that these results were not statistically analysed and were based upon very low cell numbers.

4.1.2 Attentional resources for spatial and verbal processing

The distinction between analogue and propositional mental representation can be related to models of attentional resource availability (Oakhill and Johnson-Laird, 1984; Wickens and Weingartner, 1985). Capacity models of attention hold that a quantifiable pool of attentional resources may be flexibly allocated to various competing task demands (Moray, 1967; Kahneman, 1973). Performance becomes 'resource limited' (Norman and Bobrow, 1975) when attentional resources are being maximally used and no further resource allocation can be made to cope with additional task demands. A variety of methods have been used to assess this model, one of the most common involves the use of a secondary task to 'mop up' attentional resources not allocated to primary task performance (cf. O'Donnell and Eggemeier, 1986). There are variations in procedure, but essentially a measure of resource availability is derived from a comparison of independent and simultaneous primary and secondary task performance. A number of studies utilising these methods indicated shortcomings in a model of undifferentiated processing resources. Certain combinations of primary and secondary tasks were found not to produce the expected additive resource demands (e.g. Wickens and Liu, 1988). Consequently various multiple resource models have been proposed (Navon and Gopher, 1979; Wickens, 1984a, b). Whilst some variation is evident in the nature of the resources identified, there is a good deal of consistency in the view that spatial and verbal processing draw from separate resource pools (cf. Wickens, 1984a, b; Baddeley, 1986). It has been proposed that there may be a physiological basis for such a distinction, involving hemispheric laterality (Wickens, 1984b).

Oakhill and Johnson-Laird (1984) used secondary task methodology in an attempt to establish the relative spatial and verbal processing resource demands associated with the formation of analogue and propositional mental representations of a spatial descriptions task. The requirement for verbal processing resources was found to increase when propositional representation was used, however, contrary to prediction, spatial processing resources were not found to increase when mental models were formed. Oakhill and Johnson-Laird attributed these results to an attempt on the part of subjects to form a mental model in every experimental condition, even though this was not always possible. Morra (1989) found that children were less likely to form mental models than adults. He suggested that this may be due to the comparatively more rapid development of the articulatory loop (cf. Baddeley, 1986) and consequent resource superiority for verbal processing leading to propositional representation. Obviously this result may equally be attributed to insufficient spatial processing resources for mental model formation. Wickens and Weingartner (1985) used a dual task experiment in order to examine individual differences in the mental representation of a process control monitoring task. Subjects were recruited to 'high spatial / low verbal' and 'high verbal / low spatial' ability groups. Two secondary tasks were used, one of which primarily demanded spatial processing and the other which primarily demanded verbal processing. The spatial secondary task was subject to a greater 'cost of concurrence' (cf. Navon and Gopher, 1979) for both subject groups, indicating the importance of spatial processing to primary task performance regardless of ability group. This is consistent with the findings of Oakhill and Johnson-Laird (1984) and suggests that individuals will use analogue representation when possible. Wickens and Weingartner (1985) also generated two models which related individual differences in cognitive ability to attentional resource availability. One model proposed that differences in cognitive ability do not relate to the availability of spatial or verbal attentional resources but rather to the efficiency with which they are deployed. The second model proposed that high spatial ability is associated with increased spatial resource availability, and similarly high verbal ability is associated with increased verbal resource availability. Results supported the latter model, as did a study conducted by Derrick, McCloy, Marshak, Seiler, and Reddick (1986) in which psychometric tests of spatial ability were found to be related to increased spatial processing resource availability as measured using secondary tasks.

4.1.3 The process of navigation

The discussion so far has considered the nature of mental representation, the associated spatial and verbal processing resource demands, and the nature of individual differences in mental representation and resource availability, The cognitive demands of navigation will now be examined within that context. Whilst the current focus is upon the navigation of computer task environments, there is also pertinent evidence which relates to the problems of navigation in 'real world' settings (cf. Dillon, McKnight, and Richardson, 1990). The successful performance of such tasks relies upon an efficient mental representation of the area to be navigated. In a widely cited paper, which primarily investigates the navigation of mazes by rats, Tolman (1948) used the term 'cognitive map' to refer to such a representation. Both analogue and propositional representation can be used for such navigational information (Johnson-Laird, 1983; O'Malley and Draper, 1992) with implications for the relative spatial and verbal processing resource demands. Thorndyke (1981) reports experimental results which support the existence of three qualitatively different types of navigational knowledge. 'Landmark' knowledge refers to a situation where prominent landmarks are recognised but the relationship between them is not necessarily clear. 'Route' knowledge refers to the ability to navigate from one point to another from an ego-centric perspective. Finally, 'survey' knowledge, refers to an understanding of relative spatial position, and may be gained from map study or from repeated route navigation. Route knowledge does not provide the flexibility of navigation that survey knowledge provides. If an individual is in an unfamiliar location, survey knowledge is more likely to facilitate reorientation. Route knowledge and survey knowledge can be seen to vary with respect to their verbal and spatial processing demands, with route knowledge requiring greater verbal processing resources and survey knowledge requiring greater spatial processing resources (Wickens, 1992). This processing resource distinction is supported by two experiments conducted by Wetherell (1979) in which subjects were required to navigate an unfamiliar area. A performance advantage for subjects using route information as compared to subjects using graphically presented survey information appeared to be attributable to the competing spatial demands of driving the route. Similar results were produced by Streeter, Vitello, and Wonsiewicz (1985) in a comparison of driving navigation using taped or map route information.

The concepts of landmark, route, and survey navigational information can be related to navigation within a computer task environment, and similar processing demands may apply. Evidence pertaining to this can be found with respect to many of the mechanisms employed to facilitate navigation. Icons have been used within information retrieval settings as a means of providing landmark information (Lansdale, 1988) although performance advantages have not always been apparent (see Chapter 6). An example of the provision of route information can be seen in the use of history mechanisms within hypertext which provide the user with sequential route information (Nielsen, 1990) and also in the semantic links between the system elements (O'Malley and Draper, 1992). Survey information, can be seen in the provision of system maps. The relative importance of different navigation information formats was examined by Billingsley (1982) who provided subjects with two aids to hierarchical menu search. One group was given lists of routes to their target, following the semantic links between menus (route information). A second group was given a map of the hierarchy (survey information). A final group acted as a control and was given no assistance. Performance accuracy was better in the assisted conditions than in the control condition, and when new semantic information was introduced into the hierarchy for later trials, performance in the map condition was better than that for the route condition. Parton, Huffman, Pridgen, Norman, and Shneiderman (1985) also investigated navigational assistance for menu hierarchy search. During an initial training period subjects studied either lists of menu contents with no linkage information, route information (command sequences) to targets, or a map of the system. A fourth training condition allowed subjects to explore the hierarchy using trial and error. Performance tended to be best in the map condition followed by the trial and error condition. The importance of both survey and route information was demonstrated by O'Malley and Draper (1992) who report experiments examining menu navigation and menu item recall in which subjects were found to rely upon both the semantic context of the menus and the spatial organisation of items. From these experiments it would appear that both survey and route knowledge are important to the navigation of computer information systems. Whilst the provision of assistance in either form has been shown to benefit performance, the advantage for user support would appear to lie with survey information.

Unsurprisingly, individuals with high spatial ability generally perform navigational tasks more efficiently (Thorndyke and Stasz, 1980; Vicente, Hayes and Williges, 1987; Campagnoni and Ehrlich, 1989). However, the relationship between cognitive ability and different representations of navigational information is less clear. In a study of 'real world' navigation Streeter and Vitello (1986) found that low spatial ability subjects reported placing greater reliance upon landmarks and upon verbal rather than spatial directional information. This is consistent with Kerr (1990) who found that, when asked to give a pen and paper impression of a database they had been using, good performers tended to draw spatial representations whilst poor performers did not. Given the strong association between field dependence and spatial ability (see Chapter 1) the importance of landmark information to low spatial ability subjects is also indicated by the results of Ambardar (1988), who found that field dependent subjects performed better in an information retrieval task when they were returned to the top of the menu hierarchy between searches. Jennings, Benyon and Murray (1991) examined the effects of spatial ability in relation to the performance of an information retrieval task using five different interfaces ('button', command line, iconic, menu, and question). Significant correlations between spatial ability and response times were only found for two of these interfaces (command and question). Jennings et al. (1991) suggested that this may have been because the navigational demands of these interfaces was greater. Unfortunately, there are a number of other dimensions upon which the five interfaces differ which makes it difficult to draw firm conclusions. However, there are a number of experiments which have found no interaction between spatial ability and the spatial content of the interface. In the menu navigation study mentioned earlier, Billingsley (1982) found that spatial memory was consistently predictive of performance across three conditions of varying navigational information (survey information, route information, and control). Similarly, Vicente and Williges (1988) hypothesised that the provision of a map indicating current database position would favour subjects low in spatial ability. On the basis that such individuals would have difficulty developing and using survey information this would seem a reasonable supposition. Experimental results did not support this hypothesis and both high and low spatial ability groups were found to benefit equally from these interface changes. Seagull and Walker (1992) examined the relationship between spatial ability and manipulations of menu depth and breadth. No significant relationship was found.

In summary, whilst there is evidence to suggest that navigational performance is related to spatial ability, there is only limited support for an association between spatial ability and manipulations of spatial, navigational information. It is possible, however, that in the experiments of Billingsley (1982) and Vicente and Williges (1988) high and low spatial ability subjects benefited from the provision of survey information for different reasons. For high spatial ability individuals the provision of survey information could be seen as a 'capitalisation strategy' (Cronbach and Snow, 1981) in that information is being provided in a format which takes maximum advantage of the abilities of the individual, perhaps freeing spatial processing resources for other task relevant purposes. For low spatial ability individuals the provision of survey information may be viewed as a 'compensation strategy' (Cronbach and Snow, 1981) in that it provides information in a form which the user has not been able to acquire previously and which supports the users' cognitive ability. Consequently, these studies do not necessarily reflect the relative processing strategies of high and low spatial ability individuals with respect to survey and route information.

4.2 Experiment 1: The interactive effects of spatial interface content and cognitive ability

The first experiment was designed to investigate this issue, by determining the relative performance of subjects of differing cognitive ability when survey information could no longer be applied as part of a successful navigational strategy. Subjects were recruited to groups of 'high spatial/low verbal' or 'high verbal/low spatial' ability. These groups were required to perform a search task using a menu hierarchy. It was presumed that, over a number of trials, survey knowledge of the hierarchy would be developed, and that this would be more complete for subjects with high spatial ability who would rely more heavily upon spatial processing when locating targets, whilst subjects with high verbal ability were presumed to rely more heavily upon procedural, route information in the form of the semantic links between nodes. Having performed a number of trials and developed a mental representation of the hierarchy, the spatial information content of the interface was disrupted by changing the positional constancy of items within each of the menus, whilst maintaining the semantic links between nodes. This technique has been shown be significantly disruptive of performance times (Teitelbaum and Granda, 1983). It was predicted that, given a greater reliance upon survey knowledge, the performance of high spatials would be significantly more disrupted than the performance of high verbals whose performance was more reliant upon semantic route information.

4.2.1 Method

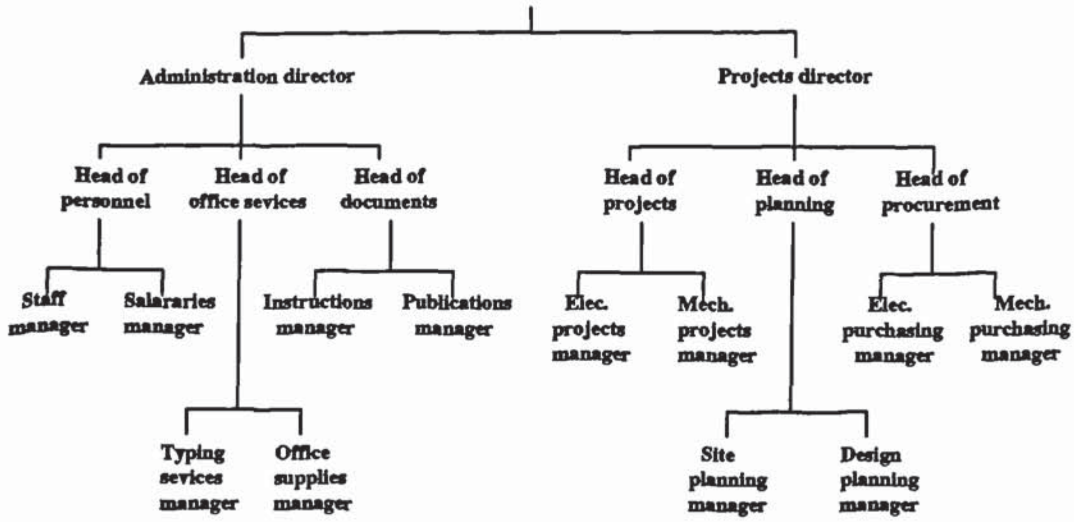
4.2.1.1 Subjects

100 computer naive (less than 20 hours interactive computer experience) Aston University students were tested for spatial (VZ2; Ekstrom, French and Harman, 1976) and verbal ability (Nelson-Denny vocabulary test, 1973). From this initial pool 16 subjects were designated as being of comparatively high verbal and low spatial ability, and a further 16 were designated as being of comparatively high spatial ability and low verbal ability. The cut off points for these divisions were approximately median splits of each performance measure from the original sample (a score of 11.75 on VZ2, and a score of 65 on the Nelson-Denny vocabulary test). This gave an eventual sample of 32 subjects who participated in the computer-based section of the experiment (mean age 19.97 years). Equal numbers of subjects from each ability grouping were randomly allocated to each of the experimental conditions described below.

4.2.1.2 Design

A hypothetical company structure was used to create a menu hierarchy, in which three levels of staffing from two company divisions were displayed (see Figure 4.01).

Fig. 4.01 : Menu hierarchy



This menu gave 2 choices at the top level (Directors), 2 x 3 choices at the second level (Heads of Department), and 6 x 2 choices at the third level (Managers). Subjects always started their search from the top level, and their target was always one of the managers from the bottom level of the hierarchy. Targets were given in a random sequence in blocks of twelve until all the managers had been used. Targets were constantly displayed immediately below the menu. Only one menu could be seen on the screen at any given time. In order to examine individual differences in the use of survey information two experimental conditions were used. Initially subjects performed five blocks of trials during which time they were able to establish a mental representation of the menu structure. For the final two blocks of the experimental condition the spatial arrangement of menus was randomly changed, whilst the semantic links between menu items was retained. In other words, the chains of responsibility within the company hierarchy remained constant, but the position of each job title upon its menu was subject to alteration. Subjects had no prior warning of this experimental condition. Menu arrangements were only altered once per trial in order that exhaustive search strategies would not be penalised. A control group was also tested in which seven performance blocks were completed with no spatial interference. This enabled the effects of spatial interference to be considered independently from the effects of fatigue which might also account for performance reductions in later blocks. It was predicted that the performance of subjects who used survey knowledge of the menu hierarchy would be disrupted by the uncertainty of spatial location in the final blocks in the experimental condition, whereas the performance of subjects who used route knowledge, and who were more concerned with the semantic links within the menu, would be comparatively unaffected. The first block of trials for both experimental and control conditions was used for demonstration and practice and was therefore not included in the analysis. The NASA TLX measure of self-report workload was administered after block five (prior to spatial disruption), and again after the final performance block.

4.2.2 Results

4.2.2.1 The sample

Table 4.01 gives the breakdown of males and females recruited to the experimental sample as either high spatials or high verbals.

Table 4.01 : Breakdown of male and female subjects in each ability group

	High Spatials	High Verbals
Male	7	5
Female	9	11

Table 4.02 presents the means and standard deviations for spatial visualisation and vocabulary scores for each ability grouping. Taken across the sample as a whole there was a highly significant negative correlation between spatial and verbal ability ($r = -.71 : p<.001$).

Table 4.02 : Means and standard deviations for spatial and verbal ability scores for high spatial and high verbal experimental groups

	High Spatials		High Verbals	
	Mean	SD	Mean	SD
Spatial Visualisation	14.86	2.21	7.63	3.12
Nelson-Denny vocab.	51.25	12.35	75.44	6.59

4.2.2.2 Model acquisition blocks

Performance during the initial model acquisition trials were analysed using a 2 x 4 (ability x block) ANOVA for both response time and accuracy (number of links traversed) data with repeated measures for the block factor. It should be noted that there will inevitably be an association between these two dependent measures.

Fig. 4.02 : Response times for blocks 1-4 for high spatials and high verbals averaged across experimental and control conditions

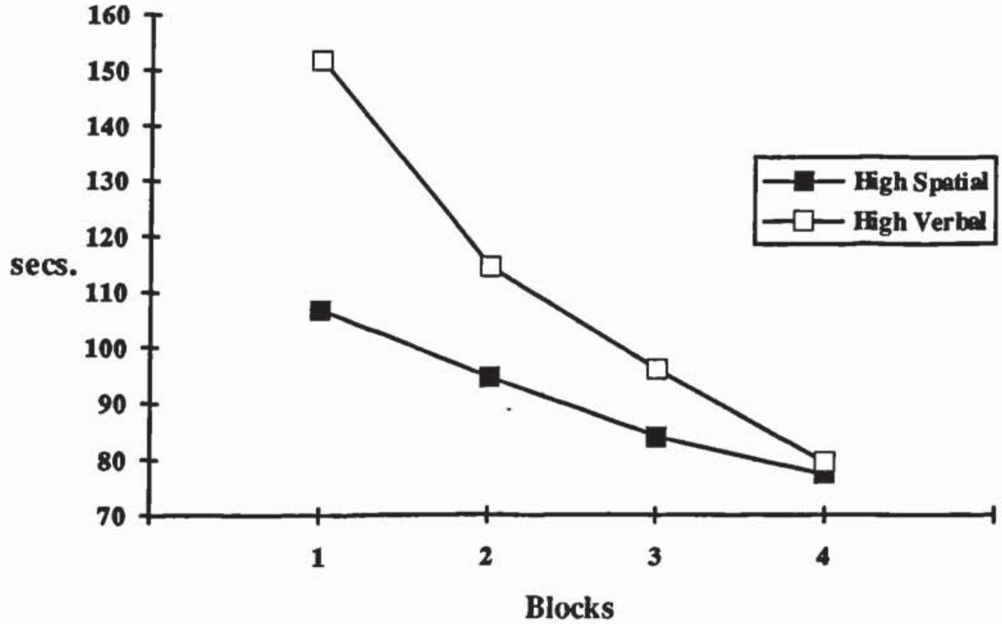
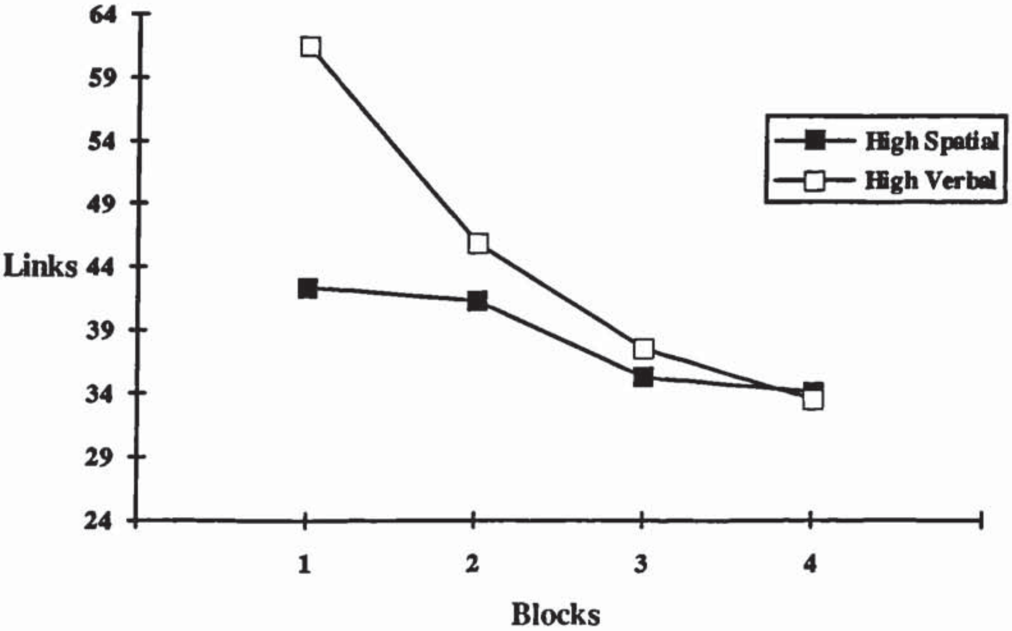


Fig. 4.03 : Accuracy of performance for blocks 1-4 for high spatials and high verbals averaged across experimental and control conditions



With respect to response times (see Figure 4.02), there was a main effect of ability group ($F(1,30)=5.89 : p<.05$), with high spatials performing more quickly than high verbals. There was also a main effect of block ($F(3,90)=37.43 : p<.001$) with performance improving over blocks, and a significant interaction between ability group and block ($F(3,90)=6.54 : p<.001$), such that high verbals were slower in the initial blocks but showed a steeper slope of acquisition to the point where speed of performance was very similar by block 4.

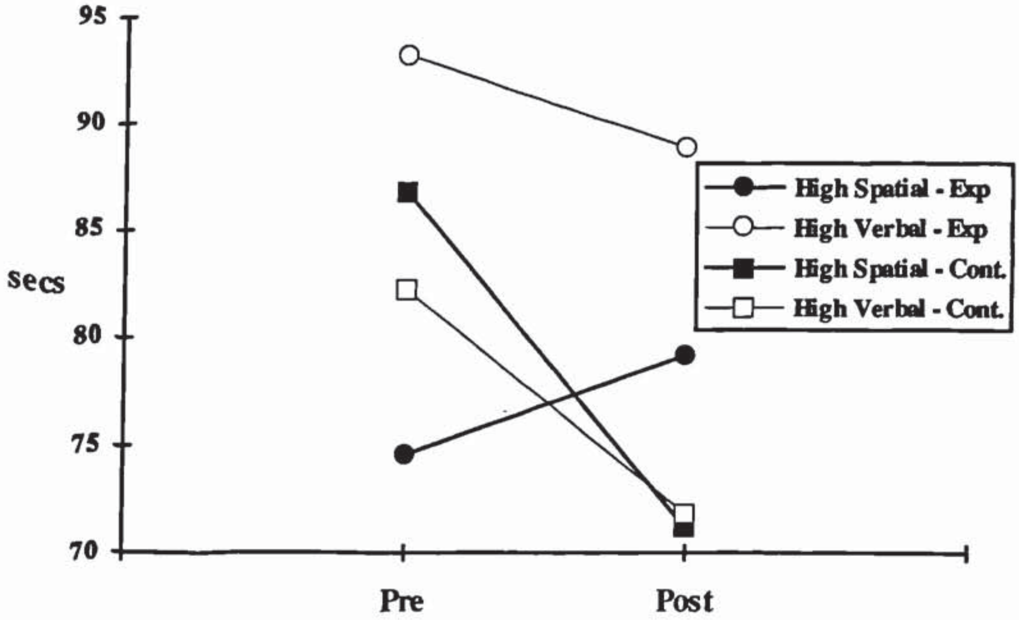
With respect to performance accuracy (see Figure 4.03), there was a main effect of block ($F(3,90)=23.59 : p<.001$) and an interaction between ability group and block ($F(3,90)=6.85 : p<.001$) such that the performance of high verbals was less accurate in the early blocks but was of similar accuracy in the later blocks. The main effect of ability group did not reach significance ($F(1,30)=3.26 : p=.081$).

4.2.2.3 Pre vs. post disruption performance

Pre vs post disruption performance was analysed using a 2 x 2 x 2 ANOVA (pre vs post x ability group x disruption condition) for response times, accuracy data, and self-report workload scores. Pre-disruption response times and accuracy data were taken as the mean of scores for blocks three and four. This provided an equal number of data points as the mean of blocks five and six, which were used as a measure post-disruption performance.

As can be seen from Figure 4.04, the only group whose response times slowed from pre to post disruption conditions were the high spatial experimental group as predicted. However the interaction between ability group, experimental vs. control group, and pre vs. post disruption failed to reach significance. Even when the experimental group alone was examined this interactive effect was not significant. There was a main effect of pre vs. post disruption condition ($F(1,28)=4.74 : p<.05$) such that performance was faster in the post disruption condition. There was also a significant interaction between experimental vs. control grouping and pre vs. post disruption condition ($F(1,28)=5.02 : p<.05$) such that there was a greater pre to post improvement in the control subject groups.

Fig. 4.04 : Response times for high spatials and high verbals in experimental and control conditions pre and post disruption



With respect to performance accuracy, as can be seen from Figure 4.05, the only group to become less accurate from pre to post disruption conditions was the high verbal group in the experimental condition. Whilst the magnitude of this effect is small, it is nevertheless contrary to the experimental hypothesis. The three way interaction between ability group, experimental vs. control group, and pre vs. post condition was not significant. However, the two way interaction between ability group and experimental vs. control grouping was significant ($F(1,28)=6.36 : p<.05$), with high verbals performing more accurately in the control condition but less accurately in the experimental condition, whilst the reverse was true for high spatials. The main effect of pre vs. post disruption condition approached significance ($F(1,28)=4.10 : p=.053$) with performance tending to be comparatively better in the post condition. When the data from the experimental group were examined in isolation then there was a main effect of ability group ($F(1,14)=7.25 : p<.018$) with high spatials performing more accurately than high verbals. There were no significant interactive effects.

Fig. 4.05 : Accuracy for high spatials and high verbals in experimental and control groups pre and post disruption

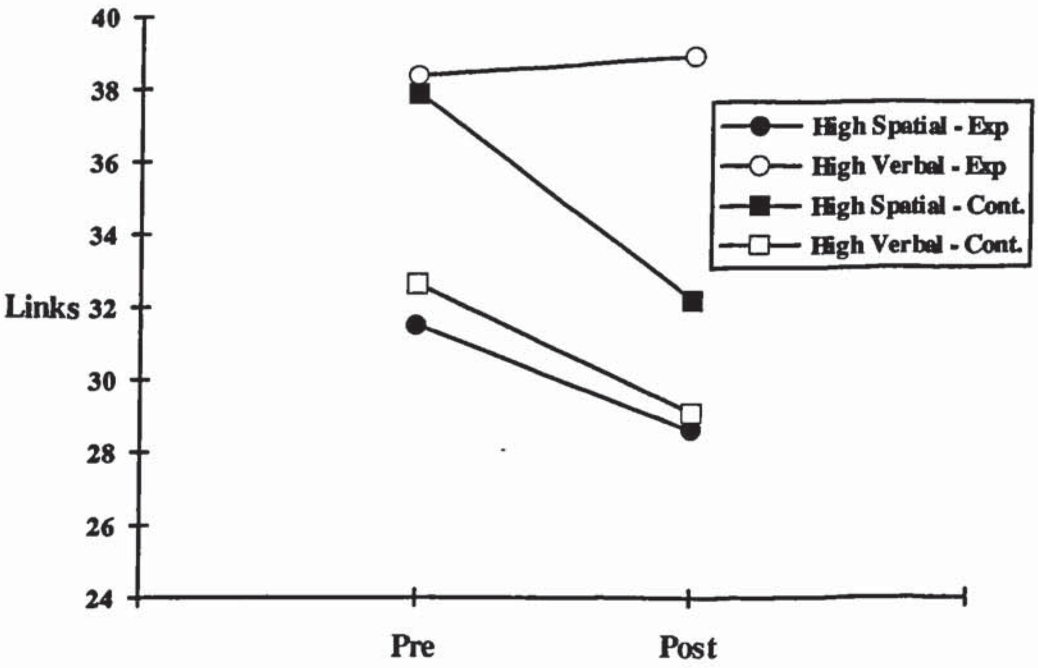
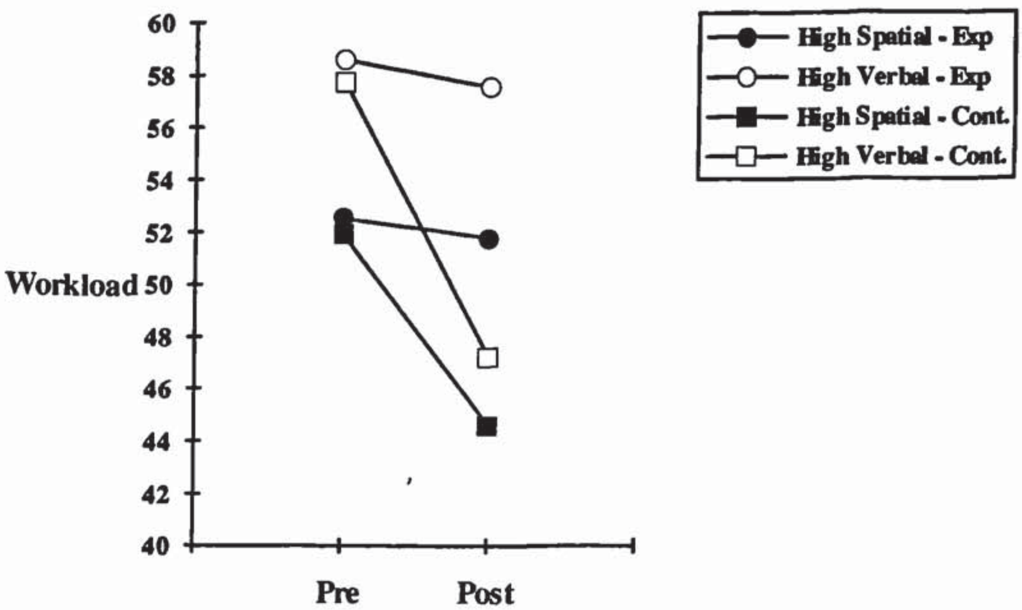


Fig. 4.06 : Self-report workload for high spatials and high verbals in experimental and control groups pre and post disruption



As can be seen from Figure 4.06 there was a greater pre to post disruption decrease in self-reported workload in the control condition. This was true for both high spatial and high verbal groups. However this interactive effect was not significant. There was a main effect of pre vs. post condition ($F(1,28)=4.59$: $p<.05$) with self-report workload being less in the post condition. All other main and interactive effects relating to workload were non-significant.

4.2.3 Discussion

The significant interaction between ability group and performance block during the model acquisition phase of this experiment was due to an initial performance superiority for the high spatial group. It seems reasonable to suppose that the locus of this effect is not concerned with the generation of a detailed spatial representation of the position of specific nodes within the menu hierarchy, as the performance advantage occurs in the early blocks when both subject groups are unfamiliar with the task environment. A more plausible explanation might relate to individual differences in mentally representing the general hierarchical structure, given the limited on-screen information. Whilst the instructions to subjects explicitly dealt with the structure of the hierarchy, there was no immediate indication when viewing a menu to which level of the hierarchy it belonged. It is possible that the performance advantage for the high spatial subject group in the early blocks was due to a greater ability to conceptualise the three level hierarchy. This would be consistent with Jennings et al.'s (1991) interpretation of their findings, as mentioned earlier.

By the fourth block, the performance of high spatials and high verbals was comparable. It would seem that by this stage high verbals have achieved a mental representation of the hierarchy of similar efficiency to that of high spatials. It may be that this representation incorporates the same level of spatial, 'survey' detail as that of the high spatial subjects, but that high verbals are slower to acquire this information. Alternatively, it may be that this representation, whilst allowing similar levels of performance efficiency, is based upon greater use of semantic, 'route' information. The comparison of pre vs post disruption scores for each ability group indicates that the former alternative best fits the data. Whilst high spatial subjects in the experimental condition were the only group whose response times deteriorated when the spatial constancy of menu items was removed, this effect was not significant. In addition, there was a tendency for the performance accuracy of the high verbal experimental group to deteriorate post disruption, whereas this was not the case for the high spatials. As expected there was a significant deterioration in overall response times as a result of spatial disruption, which was also supported by the trend for the workload scores. This indicates that the experimental manipulation of menu structure was effective and that spatial, survey information was important to successful navigation. Consequently this leads to the conclusion that survey information was an important factor in the performance of both high spatial and high verbal subjects.

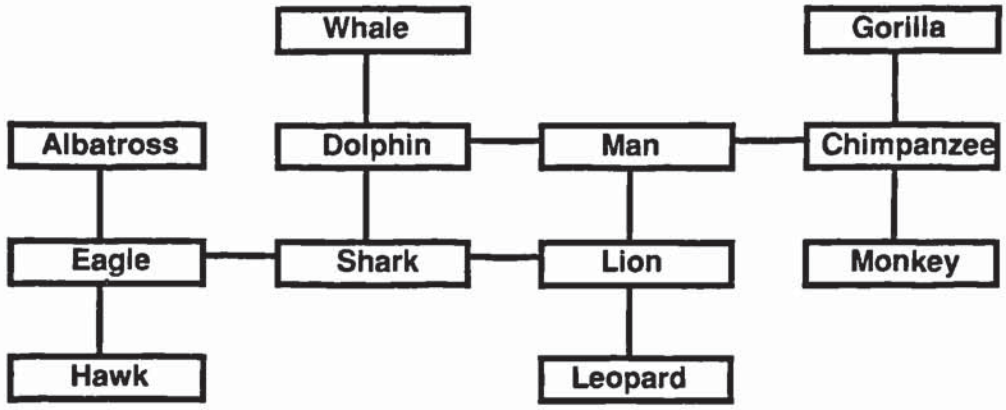
The decrease in performance accuracy post disruption for the high verbal experimental group but not for the high spatial experimental group was a surprising result. One possible explanation would be that the high verbal group were less flexible in their performance strategy, and as a result they were less able to adjust to changed task demands. This would be consistent with the findings of Cooper and Mumaw (1985) as mentioned above. On the basis of their results and those of Kyllonen, Woltz, and Lohman (1981), they suggested that high spatial ability may be associated with increased flexibility of strategy use and more efficient spatial task performance regardless of strategy.

The more pronounced pre to post disruption decrease in self-report workload for the control groups supports the effectiveness of the manipulation of spatial content, whilst the lack of interaction with spatial ability is also consistent with the performance data. It is interesting to note that there was a non-significant trend for high verbal subjects to experience greater workload during all comparative conditions of task performance than high spatials.

4.3 Experiment 2 : The interactive effects of semantic interface content and cognitive ability

Whilst the variance associated with spatial ability has not been found to be attenuated by manipulations of the spatial information content of the interface, there is some related evidence to suggest that semantic interface content may be of more importance. In a small study (n=10) examining the use of a mental model as a training aid for a multi-purpose office system, Van der Veer (1989a) found a tendency for subjects designated as 'imagers' (see Chapter 1) to have a better understanding of the relationship between the model and the system, and also of system functions. Similarly, Van der Veer and Wijk (1988) found that subjects with high spatial ability had greater knowledge of the functions of a spreadsheet, and Van der Veer (1989b) found high spatials to have a more complete mental representation of an electronic mail system. Sein and Bostrom (1989) examined the effects of spatial ability in relation to training with analogical or abstract models of an electronic mail filing task. In addition to a main effect of spatial ability, the interaction with model type approached significance for more complex tasks, with spatial ability predicting performance when using the abstract model but not for the analogical model. Using a similar task, Sein, Bostrom, and Olfman (1987) obtained a significant interaction between spatial ability and model type. Given that semantic distance (cf. Hutchins, Hollan, and Norman, 1986) is greater when using the abstract model, this suggests that the effects of spatial ability are reduced when semantic distance is low. Further support is provided by the results of Davis and Bostrom (1992), who investigated the relationship between spatial ability and the performance of a file manipulation task using a command line interface and a direct manipulation interface (DMI). The effects of spatial ability were apparent only for the command line condition. Similar results were also apparent in the information retrieval study of Jennings et al. (1991), mentioned earlier, in which spatial ability was correlated with performance when using a command line interface but not with variations of DMIs. This evidence is all consistent with the hypothesis that spatial ability is related to semantic distance, with the magnitude of the effect being greatest when semantic distance is great. This experiment was designed to test this hypothesis within a navigational context. A network navigation task was selected because of its relevance to a range of computer-based applications. Examples of similar task demands can be seen in the use of linked structures in databases, movement between the cells of a spreadsheet, and the location of text within documents. Subjects were required to navigate to targets in a small network of spatially related nodes (see Figure 4.07). They were provided with information as to their current position within the network, and nodes which were one link from this position. The degree of semantic information contained within the network was manipulated by varying the labels of the nodes as a between subjects factor. It was presumed that as the task progressed subjects would form a mental representation of the network which would enable them to navigate more efficiently to targets. At the end of this task, subjects were presented with a test of survey and route knowledge of the network. It was predicted that the effects of spatial ability would be most pronounced when semantic content was low. This was also predicted to interact with time on task, in that efficient mental representation of the network would require several trials, particularly when semantic content was low. In addition, on the basis of the results of the previous experiment, it was predicted that spatial ability would not be differentially predictive of levels of survey knowledge.

Fig. 4.07: Diagram of the network used in the semantic task condition



4.3.1 Method

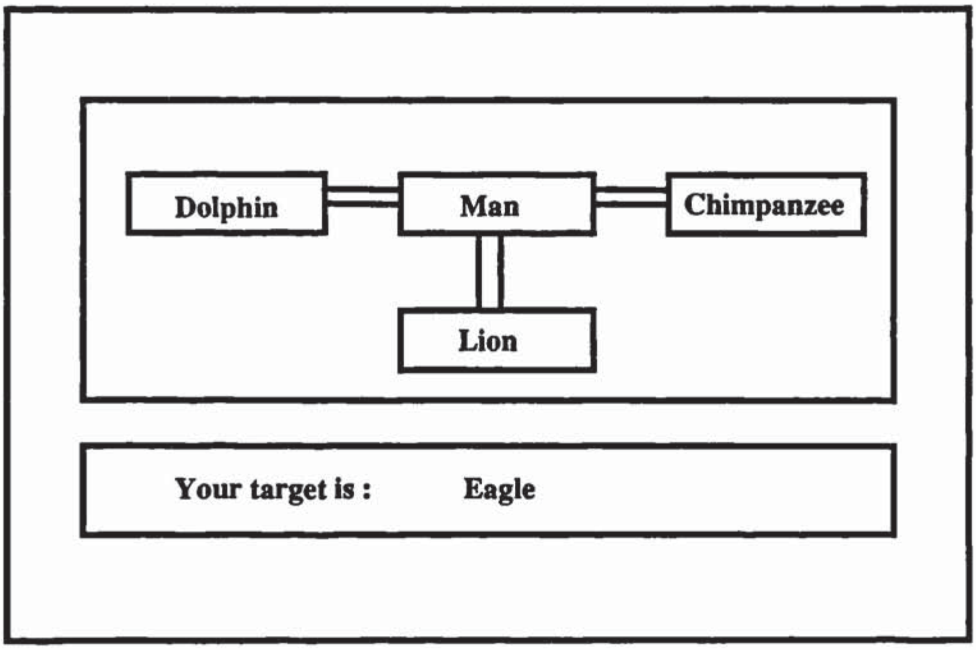
4.3.1.1 Subjects

101 computer naive (less than 20 hours interactive computer experience) Aston University students were tested for spatial (VZ2; Ekstrom et al., 1976) and verbal ability (Nelson-Denny vocabulary test). From this initial pool 24 subjects were designated as being of comparatively high verbal and low spatial ability, and a further 24 were designated as being of comparatively high spatial ability and low verbal ability. The cut off points for these divisions were approximately median splits of each performance measure from the original sample (a score of 12 on VZ2, and a score of 65 on the Nelson-Denny vocabulary test). This gave an eventual sample of 48 subjects who participated in the computer-based section of the experiment (mean age 19.27 years). None of these subjects had participated in the previous experiment. Equal numbers of subjects from each ability grouping were randomly allocated to each of the experimental conditions described below.

4.3.1.2 Design

The experimental task required subjects to navigate a network of twelve linked nodes (see Figure 4.07). Only a small portion of the network was visible at any given time (see Figure 4.08), with the current position, and one node in each direction, being displayed in a window on the computer screen. The current position was highlighted (displayed in reverse video) and positioned centrally within the window. Four nodes were initially visible within a window displayed on the computer screen. The remaining eight nodes were used as targets, and were presented in a random target sequence with each one occurring only once in each set of eight trials. The name of the current target node was displayed immediately below the network window.

Fig. 4.08 : Screen layout of network task

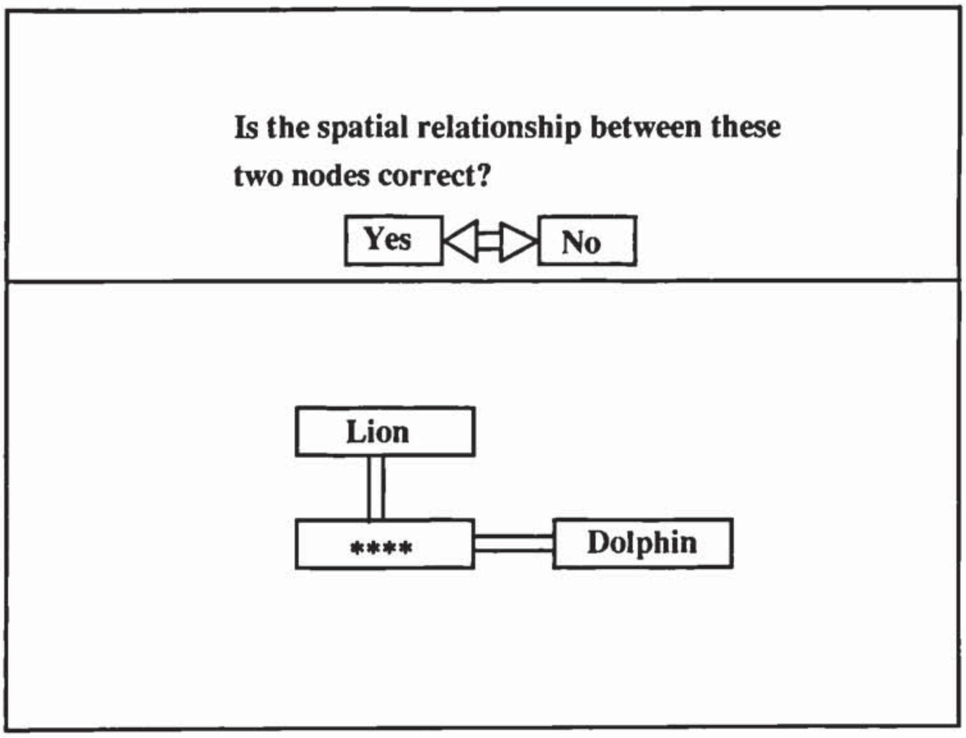


The starting position of each trial was always the same in order that the random sequence would not influence the total distance required to be navigated over a set of trials. Once a target was successfully located, therefore, the current position in the network, as indicated by a highlighted bar, was returned to the starting position. Subjects moved around the network by means of the left, right, up, and down cursor keys. Links between nodes were displayed as double lines. At any given time the network window only displayed the current position and all nodes which were only one link away.

The labels used at each of the network nodes were consistent for each condition across all trials, however, the semantic content of these labels was manipulated as a between subjects factor. In the 'high semantic' condition labels consisted of the names of animals, and nodes were linked on the basis of semantic relatedness upon a number of dimensions. So, for example, one concept used in constructing the network was 'intelligence', and for this reason the node 'man' was connected to the nodes 'chimpanzee' and 'dolphin'. In the experimental instructions subjects were made aware of these semantic links and the suggestion was given that there may be further concepts which they could develop for themselves, which would prove useful when navigating the network. In the 'low semantic' condition the node labels consisted of the numbers 1-12. These numbers were placed in pseudo-random positions within the network, although, again, this was consistent across all trials. Subjects were instructed that there was no logical sequence used in the positioning of node labels in this condition. Whilst the level of semantic content was varied between conditions, in all other respects the networks were identical.

Subjects were initially presented with one set of trials (eight targets) for demonstration / practice. This was followed by three performance blocks each of which comprised two complete sets of network targets (16 targets per block), during which measures of speed and navigational efficiency were recorded. Navigational efficiency was assessed as the ratio between the shortest possible distance to the target node and the actual distance travelled. This gave a measure of navigational efficiency in which a score of '1' represented the shortest route being taken, whilst scores closer to '0' indicated less accurate navigation.

Fig. 4.09 : Screen layout for questions designed to tap knowledge of spatial relationships



Immediately following the performance of the network task subjects were required to complete a computer-based test relating to the extent and nature of their knowledge of the network. The questions in this test were designed to tap either survey or route knowledge of the network. Survey knowledge questions took the form of a diagram of two nodes which were non-adjacent in the network (see Figure 4.09). An intervening node was represented by a row of asterisks. Subjects were required to decide whether the spatial relationship between the two nodes was correct. Route knowledge questions presented subjects with the names of two nodes and asked whether in order to travel from one of these nodes to the other it would be necessary to pass through a third named node. The two initial nodes were always separated by only one intervening node. Subjects were presented with sixteen 'survey' and sixteen 'route' questions in a random sequence. Equal numbers of 'survey' and 'route' questions required true and false responses. Responses were made using the 'Z' key if the response was 'TRUE' and the '?' key if the response was 'FALSE'. Subjects were instructed to respond as quickly and as accurately as they could, with both speed and accuracy being recorded.

4.3.2 Results

4.3.2.1 The sample

Table 4.03 gives the breakdown of males and females recruited to the experimental sample as either high spatials or high verbals.

Table 4.03 : Breakdown of male and female subjects in each ability group

	High Spatials	High Verbals
Male	11	16
Female	13	8

Table 4.04 presents the means and standard deviations for spatial visualisation and vocabulary scores for each ability grouping. Taken across the sample as a whole, there was a highly significant negative correlation between spatial and verbal ability ($r = -.71 : p<.001$).

Table 4.04 : Means and standard deviations for spatial and verbal ability scores for high spatial and high verbal experimental groups

	High Spatials		High Verbals	
	Mean	SD	Mean	SD
Spatial Visualisation	15.14	2.10	9.74	1.80
Nelson-Denny vocab.	52.54	6.96	77.79	9.26

4.3.2.2 Distribution of scores

A Kolmogorov-Smirnov test indicated that the response time data for each performance block deviated significantly from a normal distribution. When the frequency distribution for total response time for all subjects was examined, two potential outliers were apparent. However, when response time distributions were considered in relation to ability and task condition groupings these scores could no longer be considered extreme and were therefore not excluded from the analysis. A square root transformation was performed upon response time data for each of the performance blocks, and subsequent Kolmogorov-Smirnov tests no longer indicated significantly non-normal distributions.

4.3.2.3 Network navigation

Results from the network navigation task were analysed for response time and accuracy measures using a 3 x 2 x 2 ANOVA (block x ability x semantic content) with block being the only repeated measure. As with the previous experiment, it should be noted that there will be an association between the two dependent measures.

With respect to response time (see Figure 4.10), there was a significant main effect of task condition ($F(1,44)=25.59 : p<.001$), such that performance was faster for subjects in the high semantic network condition. There was also a main effect of block ($F(2,88)=22.07 : p<.001$), with performance improving in successive blocks. There was no main effect of ability. The interaction between ability and task condition was also not significant, but there was a significant ability by block interaction ($F(2,88)=4.04 : p<.05$) and ability by block by task condition ($F(2,88)=4.66 : p<.05$). These interactions appear to be primarily the result of a steeper acquisition slope for high spatials when performing in the low semantic condition. Whilst both groups performed at very similar levels in the first two blocks, the acquisition rate of high spatials was consistent whereas that of high verbals slightly decreased. It should be noted that high verbals performed marginally more quickly than high spatials in the high semantic condition. There was also a significant task condition by block interaction ($F(2,88)=6.14 : p<.005$) with a steeper acquisition slope occurring in the low semantic task condition.

Fig. 4.10 : Response times for high spatials and high verbals across three performance blocks in both spatial and semantic task conditions

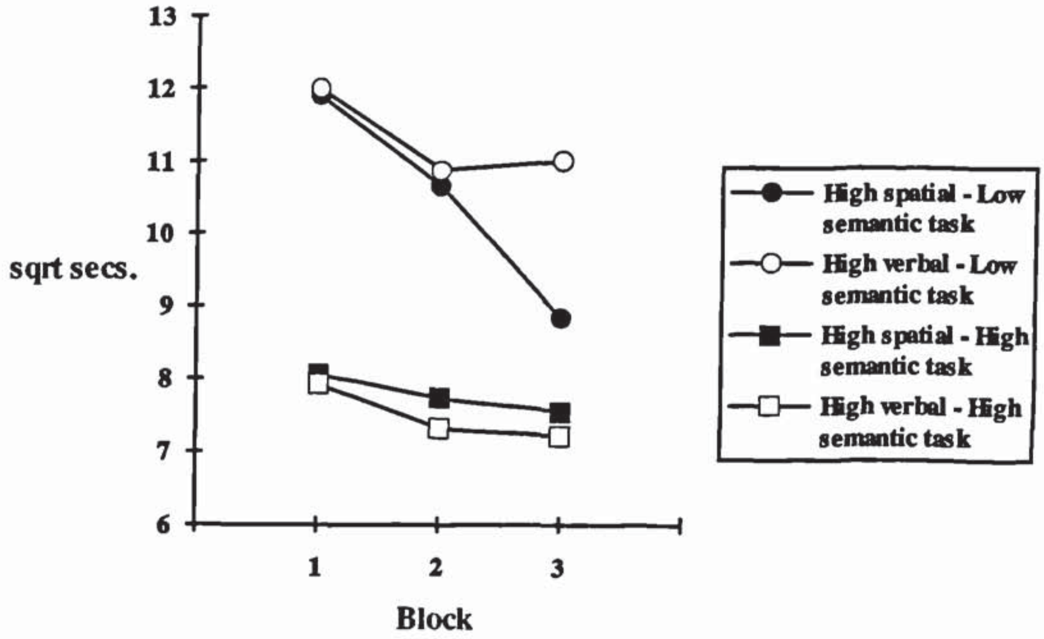
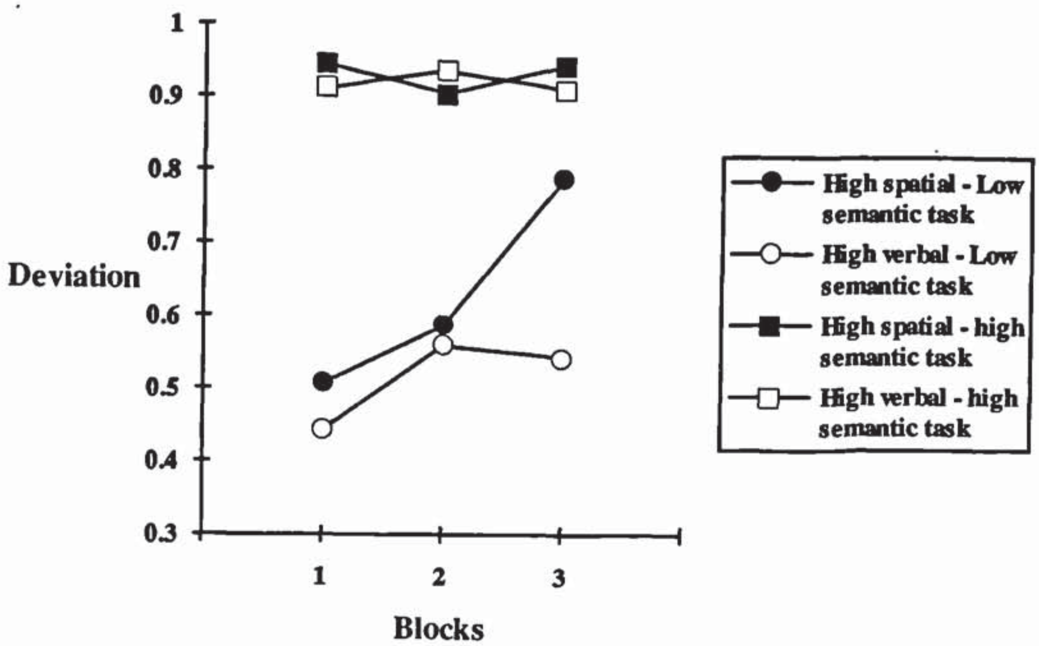


Fig. 4.11 : Accuracy for high spatials and high verbals across three performance blocks for both spatial and semantic task conditions



With respect to performance accuracy (see Figure 4.11), there was a main effect of task condition (F(1,44)=90.33 : p<.001) with accuracy being greater in the high semantic task condition. The effects of block were also significant (F(2,88)=9.95 : p<.005) with performance accuracy increasing over time. There was a significant interaction between task condition and performance block (F(2,88)=11.02 : p<.001) with greater increases in accuracy occurring in the low semantic task condition. There was a significant ability by block interaction (F(2,88)=5.90 : p<.005), which again seems to be attributable to a greater rate of performance improvement over time in the case of the high spatial group in the low semantic network condition. However, the three-way interaction between ability, task condition and performance block just failed to reach significance (F(2,88)=2.77 : p=.068).

4.3.2.4 Network recollection test

Responses to the network recollection test were initially analysed using 2 x 2 x 2 (survey vs. route questions x semantic content x ability) ANOVA for both response times and accuracy.

With respect to response times, there was a main effect of ability (F(1,44)=7.21 : p<.05) such that high verbal subjects responded to questions more quickly than high spatial subjects (5.349 secs. vs. 6.173 secs.). There was no main effect of network condition and no significant interaction between subject ability and network condition. Similarly, the interaction between subject ability and question type (survey or route) failed to reach significance.

There was no main effect of ability upon response accuracy, however there was a main effect of network condition (F(1,44)=7.27 : p=.01) such that the proportion of correct responses was greater in the high semantic task condition (.82 vs. .74). There was also a main effect of survey vs route question type (F(1,44)=24.90 : p<.001) with responses being more accurate to 'route' questions (.83 vs. .73).

In order to examine the possibility that important differences in test performance for either spatial or semantic versions of the network were being masked by very similar performance in the other network condition, a series of 2 x 2 ANOVAs was conducted in which the main effects of question type and ability group were analysed independently for the low semantic and high semantic network conditions. These analyses were conducted for both response times and accuracy. Significant ability related differences were only found following task performance on the low semantic network. There was a main effect of ability for response time (F(1,22)=7.74 : p=.011) such that high spatial subjects performed more slowly (7.118 secs. vs 5.394 secs.). However, the main effect of ability for accuracy of performance approached significance ($F(1,22)=4.17$: $p=.053$), with high spatials responding more accurately than high verbals in both question conditions (.79 vs .69). There were no significant interactive effects. These results suggest that a speed accuracy trade-off may have been occurring.

Tables 4.05 and 4.06 present the correlation matrices for the speed and accuracy of response to 'survey' and 'route' questions following performance on the 'low semantic' and 'high semantic' networks. The significant positive correlations between speed and accuracy following performance on the 'low semantic' network suggest a speed accuracy trade-off.

Table 4.05 : Correlation matrix for speed and accuracy of response to network recollection test following performance on the 'low semantic' network

	Route RT	Survey acc.	Route acc.
Survey RT	.63 ***	.35	.59 **
Route RT		.39	.59 **
Survey accuracy			.50 *

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

Table 4.06 : Correlation matrix for speed and accuracy of response to network recollection test following performance on the 'high semantic' network

	Route RT	Survey acc.	Route acc.
Survey RT	.87 ***	-.02	.16
Route RT		-.01	.21
Survey accuracy			.36

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

4.3.3 Discussion

The main effect of network task condition suggests that the manipulation of semantic content was successful, with response times and navigational efficiency deteriorating when semantic content was low. Performance upon the network task supported the experimental hypotheses. When the semantic content of the network was reduced the variance associated with spatial ability was more pronounced. In the 'high semantic' condition very little difference in performance was evident, although there was a slight performance advantage for high verbal subjects. In the 'low semantic' condition a performance advantage for high spatials became apparent in the third block of trials. This is consistent with subjects developing a mental representation of the network over time. Whilst the speed and accuracy of the performance of high verbal subjects appears to reach asymptote by the second block of trials, high spatials were able to consistently increase their performance efficiency. This suggests that high verbal subjects formed a less efficient mental representation of the network in the 'low semantic' condition, but were equally capable of representing and navigating the 'high semantic' network. This may have been attributable to a 'ceiling' effect in the high semantic condition. However, it should be noted that there is some improvement over trials in this condition, albeit a small change. In addition, performance accuracy is less than perfect, varying between 90% and 95%. Further to this, as mentioned earlier, previous experiments which have attempted to produce a 'ceiling effect' by increasing the level of spatial information, and thereby reduce the magnitude of the association between spatial ability and performance, have not been successful (cf. Vicente and Williges, 1988).

The lack of significant interaction between ability and question type for the network recollection test supports the results of the previous experiment. It would appear that recollection of route and survey information is similar for both ability groups. The response time advantage for high verbals when responding to questions may be attributable to faster reading speeds. However, this may also have been associated with response uncertainty following navigational performance in the 'low semantic' network condition in that a speed-accuracy trade-off across ability groups was apparent. High verbals responded more quickly but less accurately, perhaps indicating greater uncertainty and quick, guessed responses. The more accurate responses of high spatial subjects further supports the navigational performance data and indicates that high spatials were able to form a more efficient mental representation of the 'low semantic' network. However, it would appear that both spatial and semantic information were important in this respect.

4.4 Overall Conclusions

Taken together, these two experiments suggest that the spatial content of the interface has little influence upon the performance variance associated with spatial ability. Whilst the provision of additional spatial information, e.g. maps of data structures, may benefit performance, it will not differentially affect individuals of high and low spatial ability. The semantic content of the interface, however, appears to be of potentially greater importance in this respect. This is consistent with the results of Sein, Bostrom, and Olfman (1987; Sein and Bostrom, 1989) and Van der Veer (1989a, b) as described above. It would appear that high spatial individuals are better able to mentally represent semantically complex situations. The results of Oakhill and Johnson-Laird (1984) and Wickens and Weingartner (1985) suggest that individuals will attempt to form an analogue representation of task situations regardless of resource cost. This process appears primarily to demand spatial processing resources. The previous experiments suggest that the reduction of semantic distance facilitates mental representation and decreases the associated spatial processing resource requirement. As a consequence, the performance difference between high and low spatial ability individuals is reduced.

Following from this, there are a number of interface design issues which may usefully be addressed. In the previous experiment semantic distance was rather crudely manipulated, using just two fairly extreme positions. It is possible that there are a number of important dimensions, within this context, which may interact differentially with spatial ability. For example, Paivio (1991) distinguished between concrete and abstract concepts in relation to the ease with which imagery can be used. Klix (1986) identified property-related and event-related semantic concepts. A fine grained analysis of the association between spatial ability and semantic distance would enable a taxonomy to be developed which might be applied to the design of on-screen information presentation and the generation of system metaphors (cf. Carroll, Mack, and Kellog, 1988). This investigation should also be concerned with the comparative level of spatial and verbal processing demands associated with the process of interaction. Efficiency gains may be possible using multi-modal interaction (cf. Romary, Carbonell, and Pierrel, 1991) which is tailored to the spatial and semantic content of the interface and the relative spatial and verbal processing resource availability of the individual. Kunkel and Strothotte (1988) found that a reduction in spatial processing demand (presumed to be an overloaded system), by means of automatic speech recognition input, resulted in improved performance. Brown, Newsome, and Glinert (1988) investigated the use of auditory cues as an aid to file location. It may be that techniques such as these would prove valuable in matching the cognitive demands associated with the interface to the cognitive resources available to the individual.

4.5 Brief Reports

The following three brief reports present pilot studies which were conducted for puposes not wholly related to the present research project. They are included, however, because they are concerned with individual differences in menu and network navigation, and share some methodolgical similarities with the previous experiments. Consequently, it was thought that they may be of some interest to the reader. However, each is presented in isolation, and none are presented in great detail.

4.6 Brief Report 1 : The interactive effects of age and spatial interface content.

4.6.1 Introduction

The 'classic' pattern of ageing is held to be such that fluid intelligence exhibits a marked age-related decline, whilst crystallised intelligence is maintained at a relatively constant level or may even increase, albeit at a reduced rate. Whilst the cognitive mechanisms underlying this pattern have not been reliably established (Salthouse, 1992), it has been proposed that much of this variance can be accounted for by divergent changes in spatial and verbal processing (cf. Horn and Hofer, 1992). Spatial ability has been found to decline with age independently of levels of spatial experience or practice (Salthouse and Mitchell, 1990). This decline has been related to a reduction in the capacity of spatial working memory (Adamowicz, 1978; Hertzog and Rypma, 1991) and to a reduction in the efficiency with which spatial information is processed (Salthouse, Babcock, Skovronek, Mitchell, and Palmon, 1990). There is a corresponding age-related decline in the use of imagery (Treat and Reese, 1976), although it should be noted that older adults may be able to utilise imagery as a control process if prompted / trained (Fullerton, 1983). In contrast, facets of verbal ability have been found to be maintained or to increase into old age. Tests using the WAIS typically show that scores on the verbal sub-scale follow such a pattern (Davies, Taylor and Dorn, 1992). A number of studies have found semantic and procedural memory to be unaffected by age, although episodic memory generally declines (Mitchell, 1989). Mitrushina and Satz (1989) also present evidence which indicates that the retrieval of information which is not consistent with existing patterns of organisation within semantic memory is subject to impairment. This may have important implications for the organisation of data structures within information retrieval systems. (The matching of user and computer data models is discussed further in Chapter 7.) Measures of verbal fluency have also been found to show no age related decline (Bolla, Lindgren, Bonnacorsy and Bleeker, 1990).

Given this pattern of cognitive ageing, it can be predicted that older individuals will rely more heavily upon semantic, as opposed to spatial, processing strategies when performing computer-based tasks. However, a contrary position is suggested by Cohen and Faulkner (1983). Based upon the results of an experiment using a mental rotation task and a spatial description task similar to that used by Hunt (1978), they concluded that representational strategy was not subject to age differences, although greater performance inefficiency was apparent for older subjects (65-70 years: younger group 18-30 years) when a non-optimal strategy was selected. Cohen and Faulkner (1983) related this to an interaction between age and processing load. This experiment was designed to examine these alternate hypotheses. Subjects were recruited to younger and older age groups and performed the menu task as described in Experiment 1, above. If older subjects rely more heavily upon semantic processing strategies it can be predicted that spatial disruption of the menu organisation will have a less disadvantageous effect upon this subject group. However, if performance strategy is unrelated to age then spatial disruption should either affect both age groups equally or disadvantage the older group due to the greater processing load resulting from what has become a non-optimal strategy.

4.6.2 Method

39 subjects were recruited from local job clubs and training agencies to younger (18-25 years : mean 21.48 years : 11 male : 10 female) and older (45 + years : mean 50.28 years : 7 male : 11 female) groups. They were then allocated in relatively equal proportions to experimental (spatial disruption) and control (no spatial disruption) conditions (19 experimental : 20 controls). The task was as described for Experiment 1 reported above.

4.6.3 Results

Square root transformations were applied to both speed and accuracy data due to non-normality of distribution. Analyses were as described for Experiment 1, with age group as a between subjects factor instead of cognitive ability.

4.6.3.1 Model acquisition blocks

With respect to response time (see Figure 4.12), there was a main effect of age ($F(1,37)=6.96 : p<.05$) with the younger subject group performing more quickly than the older subject group. There was also a main effect of block ($F(3,111)=11.56 : p<.001$) with speed of performance improving over time. There was no significant age x block interaction.

Fig. 4.12 : Response times for younger and older subjects for blocks 1-4

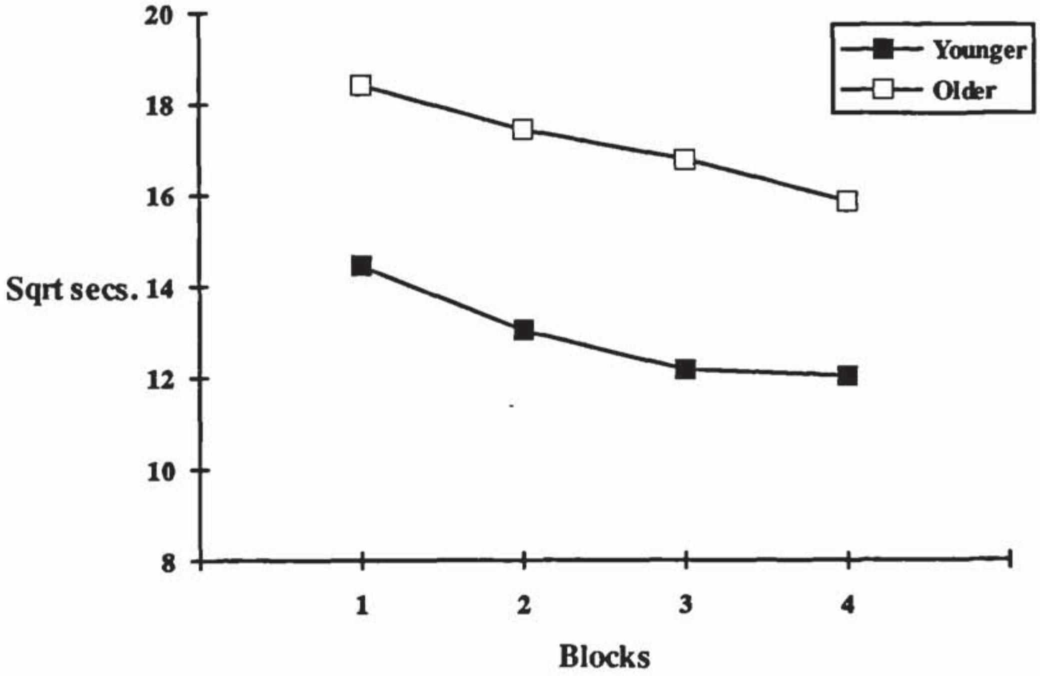
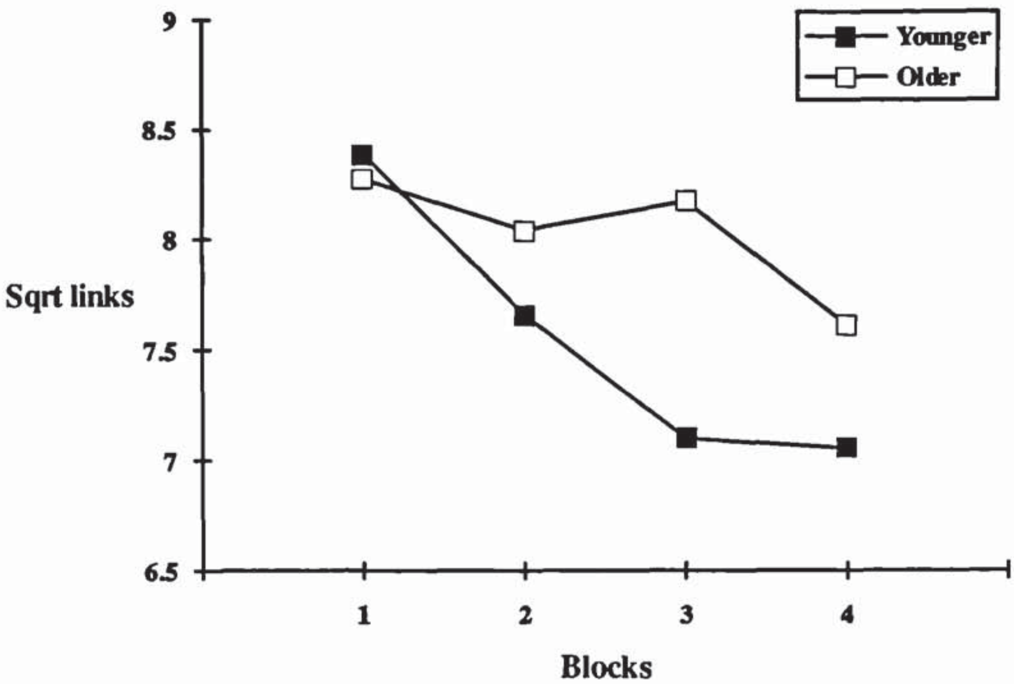


Fig. 4.13 : Accuracy for younger and older subjects in blocks 1-4

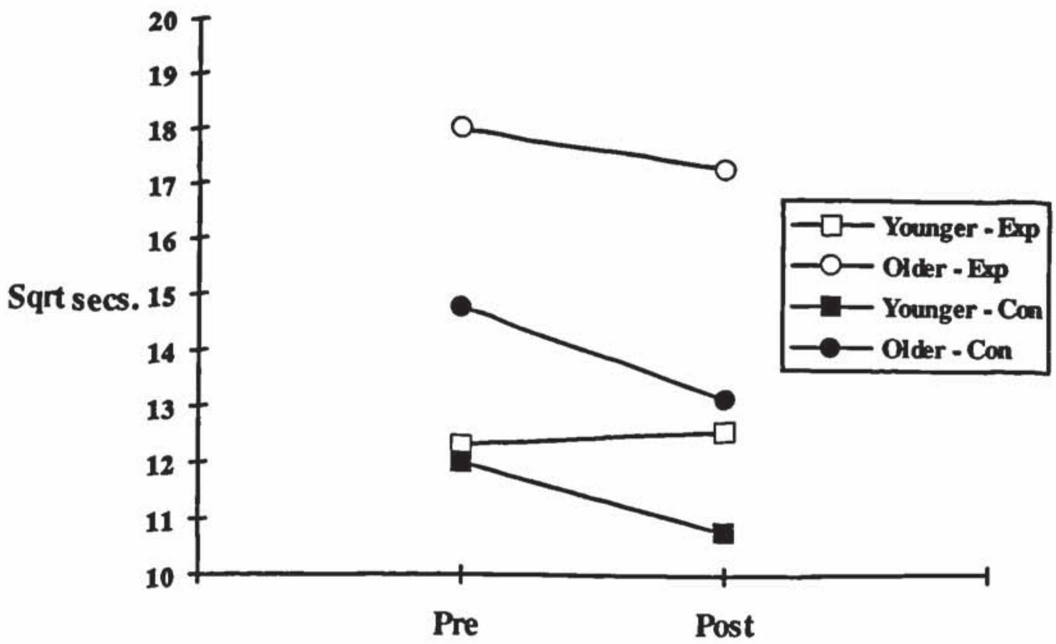


There was a significant main effect of block (F(3,111)=5.43 : p<.05) upon performance accuracy, although there was no main or interactive effects of age. However, as can be seen from figure 4.13 the performance trend suggests that, from a similar starting point, younger subjects have acquired a more efficient mental representation of the task domain.

4.6.3.2 The effects of spatial disruption

These analyses examined the pre (blocks three and four) and post (blocks five and six) disruption performance. With respect to response times there was a significant main effect of age group (F(1, 35)=8.51 : p<.01) with younger subjects performing more quickly than older subjects. There was also a significant effect of 'pre vs. post', with subjects performing more quickly in the 'post' condition (F(1,35)=8.33 : p<.01). This was primarily due to the continuous improvement of the control group as is evidenced by the significant interaction between 'experimental vs. control' and 'pre vs. post' conditions (F(1,35)=4.99 : p<.05), such that the improvement in 'pre vs. post' scores was small in the experimental condition but pronounced in the control condition (see Figure 4.14).

Fig. 4.14 : Response times for younger and older subjects in experimental and control groups in pre and post disruption conditions



Whilst the three-way interaction between 'experimental vs. control', 'pre vs. post', and age failed to reach significance, the younger group performing in the experimental condition was the only group to respond more slowly in the post disruption condition. This effect also failed to reach significance in a separate ANOVA which included only the experimental groups.

Similar analyses were conducted with respect to performance accuracy. The only significant effect was that of 'pre vs post' ($F(1,35)=9.16$: $p=.005$), in which post disruption performance was more accurate than pre disruption performance. Unlike the response time data, all experimental groups performed more accurately in the 'post' condition.

4.6.4 Discussion

During the initial model development phase of this experiment there was a consistent response time advantage for the younger subject group. This might be explained by an age related processing overhead (Salthouse and Babcock, 1991). However, the increasing performance advantage over time with respect to navigational accuracy suggests an age difference in the efficiency with which the menu structure was mentally represented. The effects of spatial disruption were such that there was a tendency for the response times of the younger experimental group to increase whilst all other groups decreased. This is consistent with the hypothesis that, due to the differential effects of cognitive ageing, younger subjects will rely more heavily upon spatial processing whilst older subjects will rely more heavily upon semantic processing. Conversely, this is not the pattern of results which would be predicted if younger and older subjects only differed along a single representational efficiency dimension. However, the magnitude of this effect was such that no firm conclusions can be drawn.

4.7 Brief Report 2 : Age differences in the spatial and verbal processing resource requirements associated with mental representation.

4.7.1 Introduction

This experiment adopts a different approach to the examination of age differences in the use of spatial and semantic information during the performance of a computer-based navigation task. The results of Oakhill and Johnson-Laird (1984) suggest that individuals will consistently attempt to form an analogue representation of task environments and that, as a result, spatial processing resource demand remains at a relatively constant level regardless of the success or otherwise of such a strategy. However, propositional representation was found to increase demand for verbal processing resources. Following from this, it can be predicted that if older subjects rely more heavily upon semantic information, as suggested in the introduction to the previous experiment, they will be subject to comparatively greater verbal resource depletion. In order to examine this hypothesis, subjects performed a network navigation task similar to that described in Experiment 2 above, with the addition of spatial and verbal secondary tasks, in order to determine the relative resource demands of navigation. Subjects were recruited to younger and older age groups. If both subject groups used similar navigational strategies it was predicted that there would be no age difference in the cost of concurrence for spatial and verbal secondary tasks. However, if semantic information was of greater importance to the older group it was predicted that there would be an increased cost of concurrence associated with the verbal task when compared to the spatial task, which would not be apparent for the younger group.

It should be noted that there is divergence of opinion as to the effects of age upon the ability to divide attention, with some suggesting that it is age sensitive (Horn and Hofer, 1992), and other suggesting it remains unaffected (McDowd and Craik, 1988).

4.7.2 Method

Eight subjects were recruited to both the younger (18-25 years) and older (45 + years) groups. Unfortunately, due to a misunderstanding of the secondary task and difficulties with the automatic speech recognition (ASR) equipment, two of the older subjects had to be excluded from analysis. The mean ages for the groups after these exclusions were 22.12 years and 52.50 years.

Only the 'high semantic' task condition (see Experiment 2) was used for this experiment as pilot work revealed that some older subjects were unable to perform the more difficult 'low semantic' network condition within a reasonable time span. Secondary tasks were designed to demand either spatial or verbal working memory resources. The verbal task used letters as stimuli whilst the spatial task used shapes (ASCII characters). Each secondary task involved the presentation of a set of three characters every five seconds, centrally displayed upon the screen. Subjects were required to respond manually, indicating whether the current presentation matched the previous presentation. Network navigation commands were issued using ASR in order to minimise response competition.

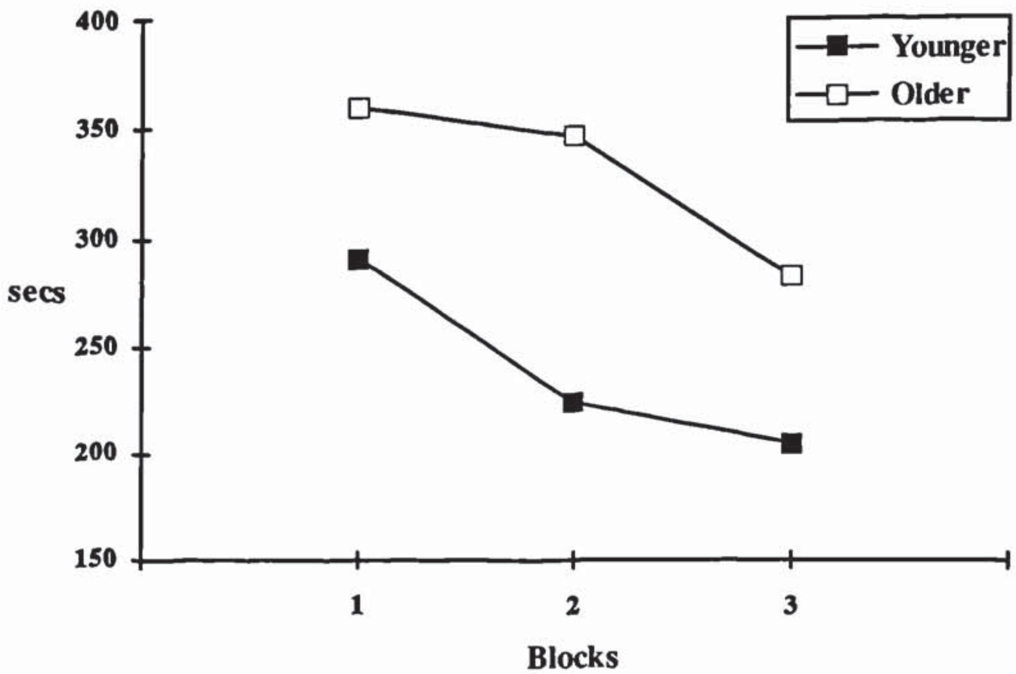
Subjects performed an initial model development phase in which the primary task was performed alone. This was followed by the performance of each of the secondary tasks independently. Finally a counterbalanced sequence of primary task alone and primary task with each secondary task was performed.

4.7.3 Results

4.7.3.1 Model development phase

3 x 2 (block x age group) ANOVAs were used to examine speed and accuracy of performance. There was no significant difference in the response times of the two age groups, although as can be seen from Figure 4.15 the young subject group tended to perform more quickly than the older subject group. The main effect of block approached significance ($F(2, 24)=3.26$: $p=.056$) as did the linear term for this variable ($F(1,12)=4.15$: $p=.064$), with response time decreasing with practice. The interaction between age and block was non-significant.

Fig. 4.15 : Response times for younger and older subjects for model development blocks

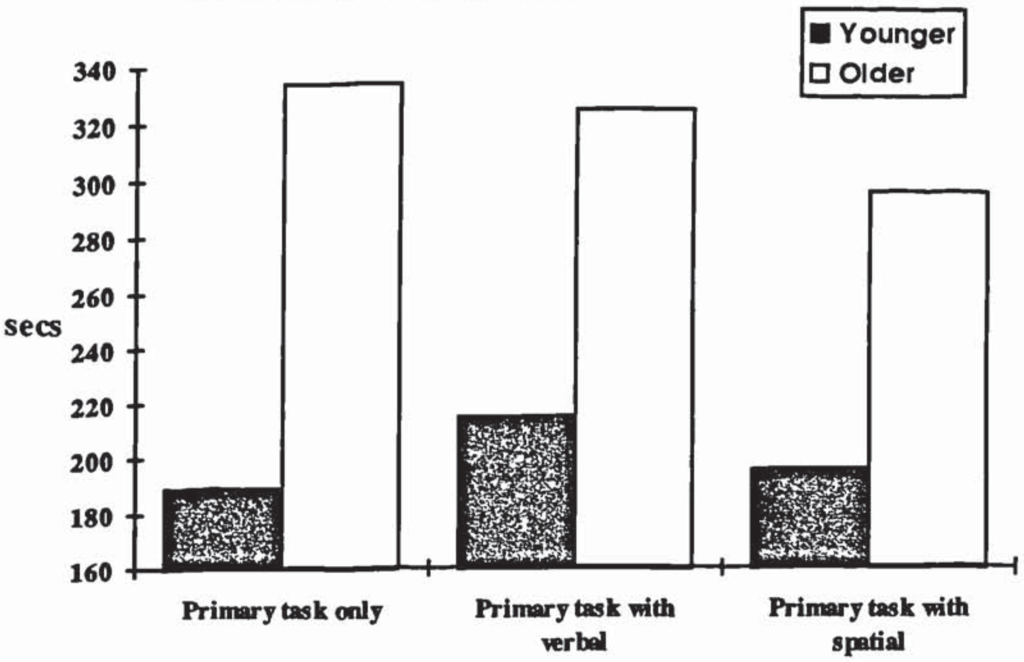


With respect to performance accuracy, there were no significant main or interactive effects of age or block for the 'model development blocks'.

4.7.3.2 Primary task performance

Primary task performance was analysed for three experimental conditions, these being primary task alone, primary task when paired with the verbal secondary task, and primary task when paired with the spatial secondary task. Performance in each condition was based upon three blocks of trials. Response time and accuracy were examined using separate 3 x 2 (primary task condition x age) ANOVAs.

Fig. 4.16 : Response times per block for younger and older subjects in primary task conditions



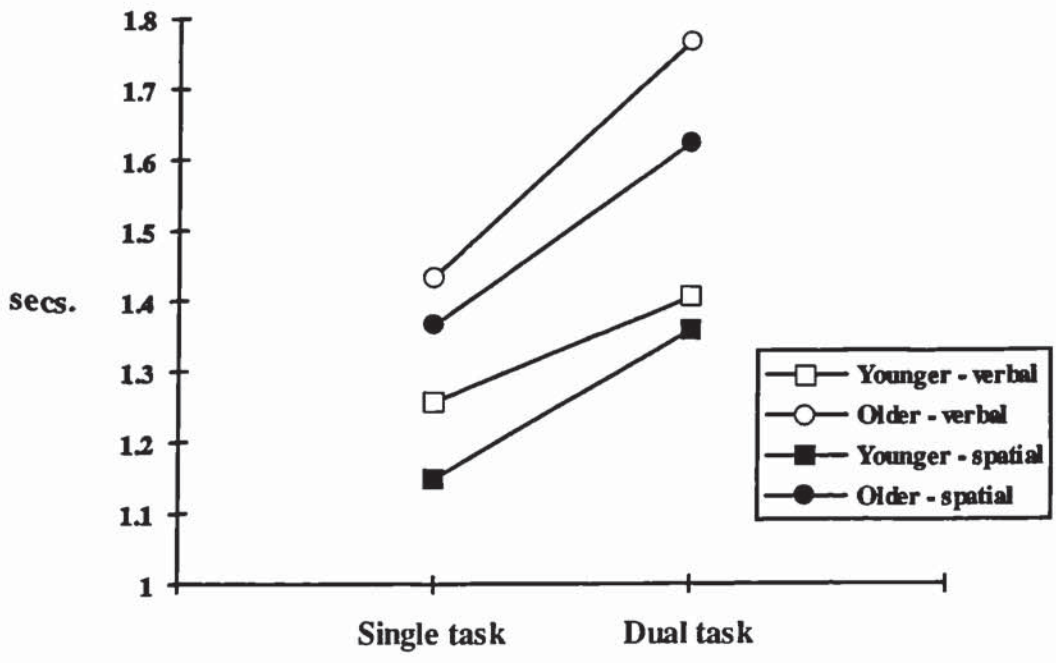
Although subjects were instructed to maintain primary task performance at prior levels, a strong effect of dual task performance is evident. The main effect of age for response time (see Figure 4.16) approached significance ($F(1, 12)=4.42$: $p=.057$) with the young age group responding more quickly than the older age group. There was a significant effect of condition ($F(2,24)=13.28$: $p<.001$) such that performance in the primary task only condition was quicker than in either of the dual task conditions, and performance upon the primary task when paired with the spatial task was quicker than when it was paired with the verbal task. There was also a significant interaction between task condition and age ($F(2,24)=6.52$: $p=.005$) such that the age difference in response speed upon the primary task was increased in the dual task conditions, and this was particularly evident when the primary task was paired with the verbal secondary task. Age differences in performance in the primary task only condition were not significant. However age differences for the primary task with verbal secondary ($F(1,12)=5.45 : p<.05$) and primary task paired with spatial secondary ($F(1,12)=5.76 : p<.05$) conditions both reached significance.

There were no significant main or interactive effects with respect to performance accuracy for the primary task.

4.7.3.3 Secondary task performance

Secondary task performance was examined using 2 x 2 x 2 ANOVAs (spatial vs. verbal, single vs. dual, and age) for response time and accuracy. Spatial vs. verbal and single vs. dual were within subjects conditions.

Fig. 4.17 : Response times for younger and older subjects in secondary task conditions



With respect to response time (see Figure 4.17), the effects of age group upon secondary task response times were not significant. However there was a main effect of spatial vs. verbal task ($F(1,12)=6.11 : p<.05$) such that response times were quicker for the spatial secondary task (1.374 secs. vs 1.466 secs.). There was also a main effect of single vs. dual task performance ($F(1,12)=23.08 : p<.001$) with response times being shorter in the single task conditions (1.288 secs. vs 1.516 secs.). There were no significant interactions.

There was a main effect of single to dual task performance for the proportion of secondary task targets hit ($F(1,12)=61.35$: $p<.001$) with the proportion of targets hit decreasing in dual task conditions (.79 vs .54). There were no significant main or interactive effects of age.

4.7.4 Discussion

The tendency for the older subject group to perform more slowly across all trials in the model acquisition phase supports the findings of the previous experiment. The fact that this effect was not significant may be attributable to the small sample size.

The cost of concurrently performing primary and secondary tasks was evident in the data for both tasks. With respect to the secondary tasks, a comparison of single and dual task performance indicates very similar effects of concurrence for both age groups and both secondary task conditions. There was a significant interaction between age group and task condition for primary task response times which was due to an increased cost of concurrence for the older age group and which applied to both verbal and spatial secondary tasks. The verbal secondary task was more disruptive of primary task performance than the spatial task for both age groups. Whilst the primary task performance decrement when paired with the verbal secondary task was greater for the older group, when compared to the relative costs of concurrent spatial secondary task performance substantial differences were not apparent. Plotting z-scores for primary task performance against z-scores for secondary task performance to derive performance operating characteristics (cf. Wickens, 1984; Kantowitz and Welden, 1985) confirmed this, with the distance between spatial and verbal tasks being very similar for both groups. It is interesting to note that for the younger age group there was no effect upon primary task response times of concurrently performing the spatial secondary task. Whilst this may suggest increased spatial processing resource availability for the younger group, it is not consistent with an increased use of survey information. In general these results support age differences in represenational efficiency (Cohen and Faulkner, 1983) stemming from an age-related decline in attentional resource availability (Salthouse, 1992). They do not indicate an age-related difference in the use of spatial and verbal processing during network navigation.

4.8 Brief Report 3 : The interactive effects of semantic distance and working memory

4.8.1 Introduction

It has been proposed that individual differences in working memory exist which cannot be attributed to encoding strategy (cf. Hunt, 1978, pp. 116-117). Measures of such differences have been found to be predictive of a number of cognitively demanding tasks, such as computer programming (Shute, 1991) and reading comprehension (Turner and Engle, 1989). This is further supported by Kyllonen and Christal (1990) who concluded from their experiments that working memory capacity accounts for nearly all the variance in reasoning ability. This may be attributed to the need to retain a number of variables before their inter-relationship can be determined (Hunt, 1978). Woltz (1988) proposed a model which related working memory efficiency and activation levels to the process of cognitive skill acquisition, in which efficiency was held to be the stronger predictor of initial controlled processing whilst activation levels were held to be the stronger predictor of later automatic processing. This experiment examined the importance of measures of working memory efficiency and activation as predictors of network navigation. The experimental task was as performed in Experiment 2, reported above. It was predicted that working memory efficiency and, to a lesser extent activation, would be related to the efficiency with which subjects were able to mentally represent the network to be navigated.

4.8.2 Method

36 subjects, aged between 18 and 25 years (mean 21.06 years) were recruited from the student population of Aston University. 20 were male and 16 were female.

All subjects completed tests of working memory efficiency and activation. These tests were based upon two of the tests used by Woltz (1988) and are described below. Subjects then performed the network navigation task as described in Experiment 2. All conditions were the same, with the exception that two levels of target difficulty were recorded. 'Easy' targets were those nodes one link away from the starting position in the network, whilst 'difficult' nodes were further than one link away.

4.8.2.1 Working memory efficiency test (WME).

Subjects were presented with three randomly selected letters on the CRT for three seconds. This was followed by letter transposition instructions. These took the form of an expression in the range -2 to +2 (not including zero). Subjects were required to memorise the initial letter presentation and then apply the transformation by counting forwards or backwards in the alphabet by the required number of places. Once this had been accomplished subjects pressed the spacebar at which point eight alternative sets of three letters were presented, one of which represented the correct solution. Subjects were required to enter a digit from one to eight in order to register their response. 28 trials were presented of which the first four were not included for analysis. Speed and accuracy of performance were recorded.

4.8.2.2 Working memory activation test (WMA).

Subjects were presented with two words, one above the other, on the CRT. The upper word represented a category, whilst the lower word represented an exemplar. Subjects were required to indicate (keyboard response) whether the exemplar was a member of the displayed category. Categories and exemplars were selected from the Battig and Montegue (1969) list of category norms. In each block of 24 trials, four trials were repeats of previous presentations. A measure of working memory activation was derived from the response time savings when comparing repeat presentations to initial presentations. Following 12 practice trials subjects performed six blocks of trials. Once again speed and accuracy of performance were recorded.

4.8.3 Results

Analyses were as performed for Experiment 2, with target difficulty as an additional within subjects factor. Subjects were allocated to high and low groups for each working memory measure on the basis of a median split.

There was a main effect of semantic content for response times ($F(1,34)=18.17$: $p<.001$) and accuracy ($F(1,34)=144.22$: $p<.001$) with performance being superior when using the high semantic content network. There was a main effect of difficulty for response time ($F(1,34)=8.37$: $p<.01$) but not for accuracy, with subjects taking longer to locate difficult targets. There was a main effect of block for response time ($F(2,68)=10.23$: $p<.001$) and accuracy ($F(2,68)=5.34$: $p<.01$) with performance improving over time. There was also a significant interaction between semantic content and block for response time ($F(2,68)=4.75$: $p<.05$) with the greater acquisition rate apparent in the low semantic content condition. There was a significant interaction between target difficulty and block for response time ($F(2,68)=8.55$: $p<.001$) and accuracy ($F(2,68)=4.96$: $p=.01$) with subjects taking longer to acquire an efficient cognitive map of difficult targets. There was also a significant three-way interaction between semantic content, difficulty and block for response time ($F(2,68)=5.33$: $p<.01$) such that performance was consistent across conditions when using the 'high semantic' network, but there was a steeper acquisition slope for difficult target when using the 'low semantic' network.

There were no main or interactive effects for WMA. However, there was a significant interaction between WME, target difficulty, and block for response time ($F(2,62)=4.87$: $p<.05$) and a significant four-way interaction which also included semantic content ($F(2,62)=4.09$: $p<.05$). Whilst there was little difference between high and low WME groups for the high semantic content network, subjects with low WME performed more slowly initially when using the low semantic content network, but showed a steeper acquisition rate, and performed marginally better than the high WME group for blocks two and three for easy targets, and for block three for difficult targets (see Figures 4.18 and 4.19).

Fig. 4.18 : Response times for subjects in high and low working memory efficiency groups for easy targets over three performance blocks

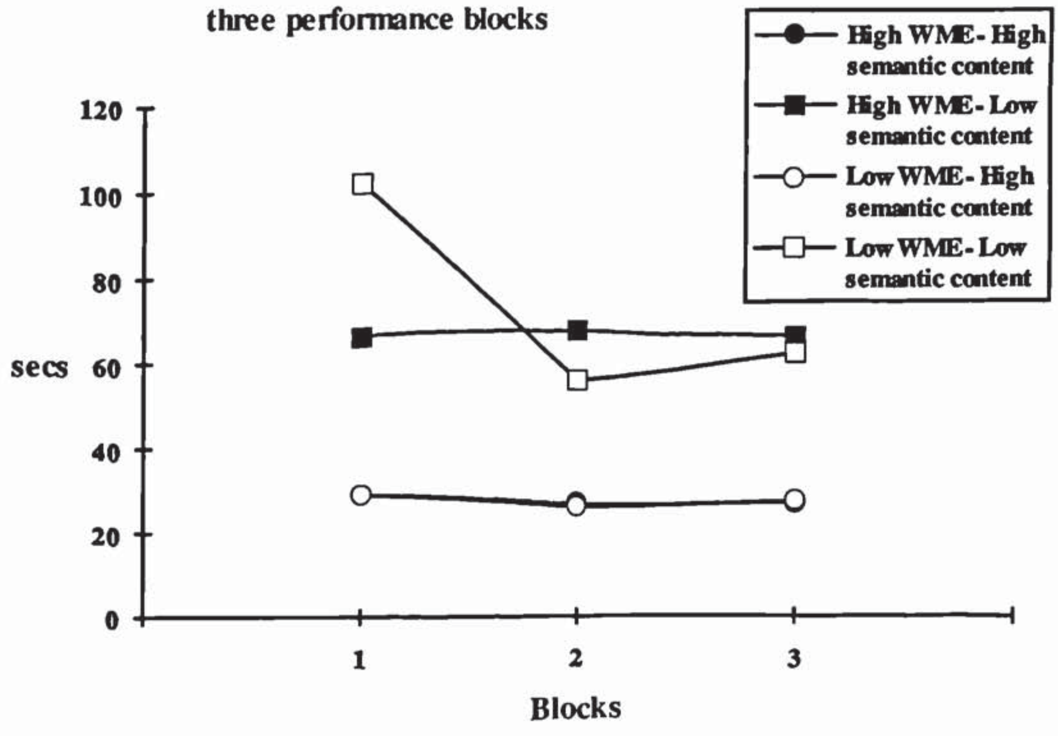
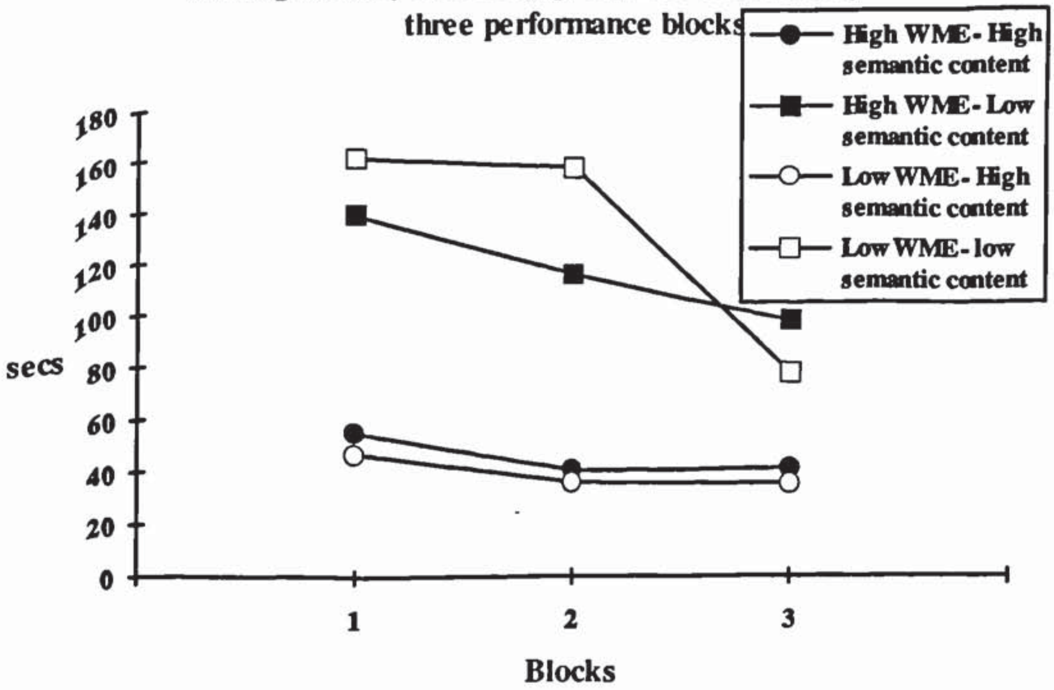


Fig. 4.19 : Response times for subjects in high and low working memory efficiency groups for difficult targets over three performance blocks



4.8.4 Discussion

The interactive effects of WME were only apparent for performance in the 'low semantic' network condition. Whilst both low and high WME groups were performing at similar levels by the third block of trials for both easy and difficult targets, the low WME group were slower to achieve this performance level. This also varied according to target difficulty with high and low WME groups having similar performance for easy targets by the second block of trials, but this was not the case until the third block of trials for difficult targets. It is interesting to note that the performance of the high WME group is relatively stable across block with respect to easy targets. It would appear that this group have efficiently modelled these targets during the practice trials, and no further improvement occurs.

WMA was not predictive of performance. If Woltz (1988) is correct and WMA is primarily associated with automatic processing then it is possible that an increased number of trials may have produced different results. The converging performance of high and low WME groups, however, might be interpreted to support the association between this measure and controlled processing. Generally these results support the hypothesised association between measures of working memory and the efficiency with which navigational information is represented, and indicate that further investigation may prove useful.

Chapter 5

The 'Generating' Component: Menu vs Command line

5.1 Introduction

This chapter presents an experimental investigation of the process of command generation. The comparative cognitive demands imposed by the menu and command line interfaces are related to individual differences in expertise and cognitive ability.

Whilst the process of command generation is universal to all computer applications, the means by which this interaction takes place are varied. Much recent interest has focused upon the use of direct manipulation interfaces (DMIs: Hutchins, Hollan, and Norman, 1986). However, the present experiment is concerned with two command generation methods of proved longevity and frequent application, the menu interface and the command line interface.

The menu interface is difficult to define with precision. There are considerable areas of overlap, particularly with DMIs. For the purposes of the current study menus will be held to provide users with a list of alternative commands or parameters from which they may select. This interface is thought to reduce the cognitive demands associated with the process of command generation. The availability of command prompts alters the user memory requirement from a process of command recall to one of command recognition (Paap and Roske-Hofstrand, 1988). This may be of particular importance when systems are frequently used by novice or occasional users, e.g. library catalogues (Baecker and Buxton, 1987), and may encourage users to employ a wider variety of commands than would otherwise have been the case (Davies, Lambert and Findlay, 1989). Menu interfaces are also thought to reduce the opportunity for syntactic error. The structure of a menu system allows a process of insulation to take place such that illegal options are not available to the user (Kantorowitz and Sudarsky, 1989). Further to this, there is an increasing tendency for menus to contain icons, or interactive display elements which would be difficult to convey parsimoniously in a command language (Thimbleby, 1990).

The command line interface typically requires the user to type commands and parameters into the computer maintaining the correct semantic content and syntactic form. In contrast to the menu interface it is thought to bring a greater flexibility to the process of command generation (Norman, 1983). The user is less restricted with respect to combinations of commands and parameters which may be linked together. The command line interface may well be a more efficient means of generating complex commands, allowing short-cuts to users who are familiar with the requirements of the application. Equivalent command generation using a menu system might entail a lengthy navigation of a complex hierarchical structure. A further potential advantage of the command line interface is a reduced requirement for screen space (Paap and Roske-Hofstrand, 1988), which is generally a valuable asset.

As mentioned, these interface related differences in cognitive demand are thought to interact with the expertise of the user (Norman, 1983; Shneiderman, 1987a; Paap and Roske-Hofstrand, 1988). The cognitive demand requirements associated with the menu interface are thought to be well suited to the novice user who is unfamiliar with the options and syntactic constructs of a particular application. In contrast the command line interface is thought to provide a more powerful and flexible means of command generation for the experienced computer user (Meister, 1989).

5.1.1 Experimental comparisons of menu and command line interfaces

A number of experimental studies in which the merits of command line and menu interfaces have been compared suggest that there is either little difference or that the performance advantage lies with the command line.

Hauptmann and Green (1983) compared command, menu and natural language modes of interaction for the performance of a graph drawing task. Subject experience varied widely, although no one had used the experimental application previously. No significant differences were found between command modes. However, the authors note that there was a problem with the menu interface such that error messages were obscured by a redisplay of the first menu.

Using a very different experimental methodology Whiteside, Jones, Levy and Wixon (1985) investigated performance differences between subjects of varying experience whilst using command line, menu, and iconic interfaces. A 'holistic' approach was adopted, comparing time to complete a file manipulation task using several different application packages (all but one of these were commercially available). They found that the performance of inexperienced users was better when using the command and menu systems (when considered together) as opposed to the icon systems. However, this appeared to be primarily due to performance when using a version of the command line interface with an improved help system. Completion times using this application were significantly faster than either the menu or the original command interface. The situation is further complicated by the fact that the help system provided many of the attributes normally associated with a menu interface, in that command syntax could be displayed in a window whilst subjects simultaneously engaged in the process of command entry. For transfer users (experienced, but not with the current system) performance with menu systems was worse than performance on the command systems, and performance on the icon systems was worse than on all the other systems considered together. No interaction between interface type and expertise was found. Experts tended to perform well on the same systems as the novices. User preferences for interface style were mixed and led to the conclusion that design quality was of greater importance than interface style.

Support for a command line advantage is also provided by Antin (1988) who compared the performance of six novice (experienced computer-users with no experience of the particular software package) and six experienced users (experienced in using the software package) upon a computer-aided design task using either menu only, command entry only, or combined menu and command modes. Completion times were found to be fastest in the command entry condition, with no interactive effect of subject expertise. Surprisingly, experts used the menu more frequently than novices in the combined condition, although the significance level of this effect is not reported. However, it should be noted that an unusual method of menu selection was employed. Subjects were required to press the control key along with either the U, D, L, or R keys in order to move up, down, left or right within the menu. In addition to issues of keystroke complexity, this system also lacks spatial congruence (Baecker and Buxton, 1987). Consequently the finding of a performance superiority for the command mode must be treated with some caution. Users did express a preference, however, for the combined mode.

A related study which demonstrated a command line performance advantage over a DMI (which included the use of pull down menus) is reported by Eberleh, Korfmacher, and Streitz (1992). Secondary task methodology was used to examine the comparative attentional resource demands involved in the use of a drawing application. Performance was faster using the command line, and it would appear that this advantage was at no additional cost in terms of workload. A second experiment was conducted in which subjects were allowed to select either method at any time during task performance. The command line method was selected most frequently, which is perhaps not surprising considering the speed advantage associated with this interface method. However, secondary task performance suggested that subjects were adjusting their command strategy to minimise workload.

There is also evidence to suggest that an increased depth of processing (cf. Craik and Lockhart, 1972) associated with use of command line interfaces improves command recollection. Davies, Lambert, and Findlay (1989) conducted an experiment which examined the effects of menu availability upon the acquisition of word processing skills. Subjects (mainly word processing novices) were recruited to four between-subject experimental conditions relating to the presence of menus during tutorial and post-tutorial performance. The availability of a permanently visible menu was not found to improve post-tutorial performance. In fact best performance times were achieved by those subjects who did not have menus present during either practice or performance sessions. The performance of subjects who had the menu visible during training suffered when the menu was not present during the performance session, suggesting an over-reliance upon the memory prop afforded by the menu.

The above studies all suggest that there is either little difference between interface modes or that the advantage lies with the command interface. There is limited empirical support for the reverse position. Streitz (1987; Streitz, Spijkers, and Van Duren, 1987) found that subjects performing a text editing task in a condition which provided menus formed a better overview of system commands (i.e. they were more aware of the system capabilities) than subjects performing in conditions without menus. Further to this, an experiment reported by Streitz (1987) suggests that the system metaphor may interact with the comparative effectiveness of these modes of interaction. Subjects performed file manipulation and text editing tasks using command line and menu interfaces. In addition the model of the system was also manipulated with one group of subjects being given a concrete 'desktop / office' metaphor and another being given an abstract 'computer' metaphor. No difference was found between interfaces when the computer metaphor was used, but performance was superior with the menu interface when the 'office' metaphor was used.

To summarise, whilst the system metaphor may be of importance in determining the relative effectiveness of the command line and menu interfaces, the weight of empirical evidence suggests that a performance advantage lies with the command line. It would appear that the speed with which commands are generated and the ease with which commands are recalled both favour this interface, although the menu interface may give a better overview of the system functions. There are, however, shortcomings in some of these experiments which may have led to an unduly pessimistic view of performance when using menus. As mentioned above, there was a problem in the experiment conducted by Hauptmann and Green (1983) relating to the display of error messages when using the menu interface, whilst Antin (1988) used a complex system for menu navigation which is not typical of the general application of menu interfaces. The study conducted by Whiteside et al. (1985) used a 'holistic' approach which was justified on the grounds that interface differences are "due to an inextricably complex interaction of causes" (p. 185). It was their view that the variables contributing to interface differences were so numerous and subject to such complex inter-relationships that the isolation of particular factors was not meaningful. However, as a result of this experimental design it is possible that other variables such as screen format or size of command set were of comparatively greater importance in determining relative interface performance. Consequently, there are grounds for believing that the empirical support for the command line interface may, in part, be attributable to methodological flaws.

5.1.2 The interaction between expertise and interface style

The above experiments provide little support for the hypothesised interaction between expertise and command method. Whiteside at al. (1985) found that experts performed well on the same systems as did novices. Antin (1988) also found no interactive effects of expertise and interface mode, although experts preferred the menu mode. A similar result can be seen in an experiment conducted by Benbasat, Dexter, and Masulis (1981). During the performance of a computer-based management simulation task, dialogue was either user-guided (command line), or system-guided (arguably similar in cognitive demand to a menu interface). There was no significant interaction between expertise and dialogue mode.

However, some evidence of the benefits of supported dialogue methods for novice users is provided by Macaulay and Norman (1984). During the performance of a task requiring the records of fictitious companies to be accessed, dialogue mode was either linear (complete command required) or substructured (commands given in stages in response to prompts). Novice users were found to perform better in the latter condition.

In experiments concerned with the development of an adaptive interface, Martin and Fuerst (1987, 1988; Trumbly, Arnett, and Martin, 1993) examined the relationship between computer semantic knowledge (as assessed by a recognition test of computer-related terminology) and interface design. Subjects were required to perform a jury simulation task (Martin and Fuerst, 1987, 1988) or a manufacturing simulation task (Trumbly et al., 1993) using one of two interfaces. The first was designated the 'human factors' interface and used menus, input verification, and automatic help functions, and was predicted to favour novices. The second interface was regarded as a control condition, it provided none of the 'human factors' elements, and was thought to favour experts. Results broadly supported these hypotheses with low semantic knowledge subjects performing better than high semantic knowledge subjects in the 'human factors' condition and worse in the control condition.

In summary, the evidence relating to an interaction between expertise and interface mode is mixed. Benbasat, Dexter, and Masulis (1981), Whiteside et al. (1985), and Antin (1988) found no evidence to suggest an interaction. However, as mentioned above, factors associated with the implementation of the menu interface may have clouded the situation. An interactive effect was apparent in the experiments conducted by Martin and Fuerst (1987, 1988) and Trumbly et al. (1993). However, in these experiments menu availability was not manipulated independently, and other elements of the 'human factors' interface, such as input verification or help functions, may have been more crucial in determining the relative performance of novices and experts.

5.1.3 Cognitive ability and command generation

The importance of cognitive ability as a predictor of command generation was demonstrated by Egan and Gomez (1985) in their investigation of individual differences in text editing. In an examination of the component elements of the task, both spatial memory and age were found to be associated with the process of command generation. These findings were supported by the experiment reported in Chapter 2 in which logical reasoning and spatial visualisation were also found to be consistently predictive of performance. However, empirical evidence relating individual differences in cognitive ability to differences in the method of command generation is limited.

Jennings, Benyon, and Murray (1991) examined spatial ability, verbal ability, field dependence, short term memory, and the Myers-Briggs Type Inventory as predictors of an information retrieval task in which subject performance was compared whilst using five different interface methods. These included command line and menu interfaces. There was a significant negative correlation between spatial ability and completion time for the command line interface and the 'question interface' (subjects typed responses to on-screen prompts). However, the correlation between spatial ability and performance using the menu interface was not significant. The only other significant association was between verbal ability and completion time when using the 'question' interface. Further support for a stronger association between spatial ability and performance when using a command line interface as opposed to a menu interface is provided by Greene, Gomez, and Devlin (1986). A pencil and paper exercise was used to examine individual differences in the generation of database queries. Four interface-related experimental conditions were used, of which two included the use of a table interface from which subjects selected search attributes, one required subjects to select from four alternative structured query language (SQL) commands, and one required subjects to generate SQL commands. These conditions can obviously be related to different demands associated with menu and command line interfaces, with the table conditions most closely resembling a menu interface, the SQL generation condition most closely resembling a command line interface, with the SQL choice occupying an intermediate position. In addition four within-subjects query operator conditions were used (the conditional expressions "AND", "NEGATION", "OR", and "AND + OR"). Green et al. (1986) used multiple regression analyses to investigate individual differences in performance for each operator condition, and also included 'indicator variables' to differentiate interface conditions. Age, verbal ability (Nelson-Denny, 1973), spatial memory (MV2, Ekstrom, French, and Harman, 1976), logical reasoning (RL2; Ekstrom, et al., 1976), and integrative processing (IP1; Ekstrom, et al., 1976) were all predictive of performance, but only as part of a complex pattern across the different operator conditions. No details were given as to the nature of the coding used for the interface indicator variables, however, it would seem that there was a tendency for reasoning ability and spatial memory to be more strongly predictive of performance in interface conditions which more closely resembled a command line. A later report of this experiment (Greene, Devlin, Cannata, and Gomez, 1990), which considered just the SQL generation and one of the table interface conditions, broadly supported this position. Given the strong association between measures of field dependence and cognitive ability discussed in Chapter 1, further support of this position is provided by a study conducted by Fowler, Macaulay, and Fowler (1985). Field independent subjects were found to prefer an unstructured method of command input, whilst field dependent subjects were found to prefer a sub-structured method of command generation which might be compared to the support given by a menu interface.

Davis and Bostrom (1992) investigated the relationship between spatial ability (VZ2 : Ekstrom, et al., 1976), Learning Style (Kolb, 1971) and the use of a command line interface and a DMI. Subjects were novice users who, after a period of training were required to perform a number of file manipulation tasks. Tasks were designated as being either near transfer (of equal complexity to those tasks for which they had been trained), or far transfer (requiring them to string together a number of near transfer tasks in order to achieve an end result, in ways for which they had not been trained). Far transfer tasks inevitably required a greater understanding of the nature of the system with which they were working. Dependent measures included judges' assessments of errors made, a test of system comprehension, and a questionnaire concerning satisfaction with the interface (Davis, 1989). Response times were not included as such as a measure of performance, although Davis and Bostrom reported that task completion time was limited to 30 minutes. All DMI subjects finished within the allocated time, but approximately 25% of those using the command interface failed to do so. Performance using the DMI was found to be superior to that using the command line interface. This difference was apparent for both near transfer and far transfer tasks, and for system comprehension. No significant effects of learning styles were found. However, there was a main effect of spatial ability with high spatials committing fewer errors. This was primarily due to a performance advantage in the far transfer condition. Individual differences in system comprehension were also apparent with high spatials exhibiting a greater understanding in the command line condition, although the difference was non-significant for the DMI condition. Whilst interface type had no effect upon the level of reported satisfaction, there was a significant difference in the levels of reported ease of use associated with the system by high and low spatials, with high spatials reporting greater ease of use. However, the study by Eberleh et al. (1992), mentioned above, found no significant differences in the spatial, verbal, or perceptual speed demands of command line or direct manipulation interfaces when assessed by means of secondary tasks.

In summary, these studies provide some evidence to support the existence of interactive effects between cognitive ability and the method of command generation. The experiments of Fowler, Macaulay, and Fowler (1985), Jennings et al. (1991) and Greene et al. (1986) suggest that the association between cognitive ability and interface type may be stronger when using a command line interface than when using a menu. In addition, the study by Davis and Bostrom (1992) indicates that spatial ability interacts with command complexity.

5.1.4 Experimental aims

As mentioned above, the varying cognitive demands associated with the command line and menu interfaces provide grounds for believing that they will differentially benefit expert and novice users. Whilst there is only limited support for such an interactive effect, the mechanics of the menu interface used in some previous studies may have disproportionately disadvantaged novice users. Consequently, one of the areas with which this experiment is concerned is a re-examination of this relationship. In particular the effects of syntax complexity were considered. Subjects were recruited to novice and expert groups on the basis of previous computer experience. Experts were therefore comparable with the 'high semantic knowledge' subjects of Martin and Fuerst (1987; 1988) or the 'transfer' users of Whiteside et al. (1985). They were required to perform a file manipulation task, using either menu or command line interfaces, in which the syntactic complexity of commands was manipulated. Expert subjects were thought to be more skilled in the use of command syntax, to have a more efficient mental representation of the task (Hanisch, Kramer, Hulin, and Schumaker, 1988), and better recall of complex sequences of command syntax (cf. Ericsson and Smith, 1991). Upon this basis, it was hypothesised that the reduced cognitive demand associated with the menu interface (Streitz, Spijkers, and Van Duren, 1987) would lead to a reduction in expertise related performance differences, whilst the increased demands associated with the command line interface would produce the opposite result. It was further predicted that the magnitude of this effect would increase in line with syntax complexity.

There is limited evidence relating individual differences in cognitive ability to differences in the method of command generation. There are, however, grounds to predict that cognitive ability will be more strongly associated with performance in the command line interface condition. Egan and Gomez (1985) have proposed that spatial memory is related to the processes of command symbol recall and organisation. As discussed earlier, the requirement for such processing is greater when using a command line interface, which suggests an interaction between spatial memory and command method. This would be consistent with the evidence of Jennings et al. (1991) and Greene et al. (1986; Greene et al., 1990), as presented above. The study by Greene et al. (1990) suggests that reasoning ability may also differentially predict performance for each interface condition. Further to this, in keeping with the results of Davis and Bostrom (1992), it can be predicted that cognitive ability will interact with command complexity, such that the effects of cognitive ability will be greatest in conditions of high command complexity.

Ackerman's (1988) theory relating individual differences in cognitive ability to the process of skill acquisition was outlined in Chapter 1. Upon this basis it was predicted that the effects of cognitive ability would interact with those of expertise. Given that novices would rely more heavily upon controlled processing (Schneider and Shiffrin, 1977) when generating commands, and that cognitive ability is more strongly related to controlled processing, it was hypothesised that the effects of cognitive ability would be greatest for novice users, and that this would be particularly apparent in experimental conditions with the greatest cognitive demand.

Finally, a distinction is drawn between command generation ability and command generation strategy. Separate experimental conditions were used in which subjects were either restricted to the command line or menu interfaces, or were given free access to both methods of command generation. This enabled an independent examination of the relationship between cognitive ability and individual differences in performance strategy, which may be a fundamental factor in the performance of many cognitively demanding tasks (Lohman and Kyllonen, 1983). It was predicted that subjects with comparatively low cognitive ability would choose to generate commands using the menu interface, whilst those with comparatively high cognitive ability would select the command line generation method. This was also predicted to interact with command complexity such that the relationship between cognitive ability and strategy would be strongest for complex command conditions. The relationship between expertise and strategy is clouded by the finding of Antin (1988) of an expert user preference for the menu method of command generation. However, on the basis of the cognitive demands associated with each interface it was predicted that novice subjects would use the menu interface more frequently than the command interface, and the reverse would be true of experts.

5.2 Method

5.2.1 Subjects

64 subjects aged between 18 and 30 years (mean 20.53) were recruited from the student population of Aston University. Of this sample, 32 were computer novices (having fewer than 20 hours interactive computer use) and 32 were experts (having over 100 hours interactive computer use). Equal numbers of male and female subjects were recruited to all experimental conditions. All subjects were right handed, reported vision which was normal or corrected to normal, spoke English as their first language, and had not taken part in any of the previous experiments.

5.2.2 Measures of individual differences

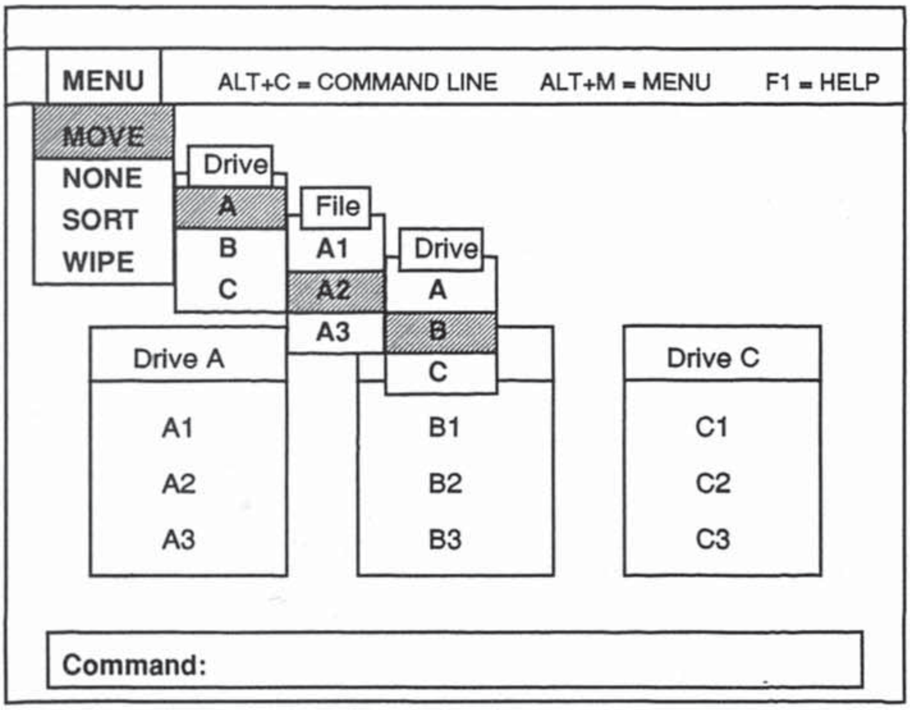
Subjects initially completed tests of verbal ability (Nelson-Denny vocabulary test, 1973), spatial memory (MV2; Ekstom, French, and Harman, 1976), spatial visualisation (VZ2; Ekstrom et al., 1976), logical reasoning (LR2; Ekstrom et al., 1976), associative memory (MA2; Ekstrom, et al., 1976), and computer literacy (CALIP; Poplin, Gable, and Drew, 1984).

5.2.3 Experimental task overview

This task was performed on a PCAT 80286 computer with a Hercules screen. The software was purpose written in C and Assembler with a timing resolution of ± 5ms. Text mode was used in which 25 x 80 characters were displayed on a full screen.

Subjects were required to maintain the contents of three computer 'disk drives' designated 'Drive A', 'Drive B', and 'Drive C' (see Figure 5.01). The contents of these disk drives were repeatedly displayed on the computer screen in three adjacent boxes. Each drive contained three files, the names of which consisted of one letter and one number. If the files were correctly organised (in the 'target state') the letter in each filename corresponded with the letter designating the drive and the files were in ascending numeric sequence. The contents of the drive could be displayed in four possible states. Firstly, all the files could be in their correct position. Secondly, all drives could contain files with the appropriate letters but with an incorrect numeric sequence on one of the drives. Thirdly, a file could be present on one of the drives which did not belong on any of the drives. Finally, a file could be present on one of the drives which belonged on one of the other drives.

Fig. 5.01: Screen layout including pulldown menus



In order to achieve the target state the following commands were available: NONE - This command indicates that the contents of the disk drives are all in the correct sequence; SORT [drive] - This command sorts the contents of the specified drive with respect to the number sequence of the files on the drive; WIPE [drive] [filename] - This command deletes the specified file from the specified drive; and MOVE [source drive] [filename] [destination drive] - This command transfers the specified file from the specified source drive to the specified destination drive.

These commands were selected in order that: (i) there should be a range of complexity within the command set. (This was manipulated by varying the number of parameters required to be included for each command), and (ii) each command name should be four letters long, containing only one syllable.

5.2.4 The task sequence

5.2.4.1 Initial testing

Following completion of the psychometric tests, subjects completed a short test of typing speed. This followed the same format as the typing test used in Chapter 2. However, each of the sentences presented for typing contained one of the command words used in the main experimental task.

5.2.4.2 Command recognition task

A potentially confounding variable in this investigation relates to the complexity of the stimuli which prompt commands. The process of interest in the current experiment is that of command generation, and is not related to stimulus recognition differences. In other words, the present concern is not with the ability of the subject to recognise the task conditions which determine the need for a particular command, but with the process of command generation once these task requirements have been established. In order to minimise or eliminate the effects of such unwanted variance a command recognition task was performed. This served two purposes. Firstly, it ensured that subjects were well practised in the processes involved in recognising the various command requirements (a similar procedure has been used in research by Xerox to reduce mouse performance variance; Baecker and Buxton, 1987). Ideally command stimuli (the state of the disk drives) would be processed automatically (Schneider and Shiffrin, 1977). Secondly, this task enabled performance upon the forthcoming command generation task to be adjusted for individual differences in the process of command recognition.

Subjects were repeatedly presented with the contents of the 'disk drives' and required to monitor the drives for <u>one</u> of the possible drive states outlined above. Each of the possible drive states was used as a target in a counterbalanced sequence. If this state was true then they were required to press the spacebar as quickly as possible. The contents of the drives were presented in two consecutive blocks of 36 trials for each command. In each block nine trials were randomly selected as target presentations. Response time and accuracy measures were only recorded for the second presentation block. The contents of the drives were displayed for a duration of two seconds, followed by an offset of one second.

5.2.4.3 Command generation

The final phase of the experiment required subjects to maintain the contents of the disk drives in the target state by issuing one of the four commands. Subjects were instructed in each of the commands and each of the command generation methods. They were then given a single familiarisation trial with each command for each command generation method.

Three conditions were used with respect to the method of command generation. In the first two of these conditions commands were generated by either command line arguments, or by a menu system, but only one or the other was available. The order of presentation of these conditions was counterbalanced. In the third condition both methods of command generation were available to subjects. This enabled subject strategy to be examined. In all conditions a process of mode selection was required before the command could be generated. This required the subject to press the 'ALT' key along with either the 'C' key for the command line or the 'M' key for the menu interface modes. The selection overhead for each mode was therefore the same. In the command line condition subjects were required to type the command, leaving one space between the command word and each of the following parameters, and then press the ENTER key. In the menu condition commands were organised in alphabetic sequence on the first menu. Subjects were required to use the up and down cursor keys to move a highlighted bar to the required command and then to press the ENTER key. If parameters were required additional menus would appear from which subjects could make further selections in a similar manner. In both conditions if an error was committed an error message was displayed in the centre of the screen until the subject pressed the spacebar, at which point the message and the previous command were cleared. A help screen was also available to subjects at any time by pressing the F1 key. This displayed the correct syntax for each command as described above. However, invoking the help screen cancelled the current command, and it was not possible to simultaneously display the help screen and issue commands.

In each condition 12 practice trails were given, followed by 36 commands for which performance was recorded. Equal numbers of each command type were presented. During the recorded trials the position of the files was controlled so that the file which prompted the command occurred randomly in every possible position. For the conditions in which performance ability was of prime concern, dependent measures included time per successful command, and proportion of errors. For the condition in which performance strategy was examined, the number of commands generated with each interface type was the dependent measure. Following each experimental condition subjects completed the NASA TLX measure of self-report workload.

5.3 Results

5.3.1 Command recognition tasks

Performance upon the four command recognition tasks was analysed using a one-way ANOVA for each dependent measure. Mean response times, the proportion of correct responses, and the proportion of false positives are shown in Table 5.01. Command means were significantly different for response times ($F(3,186)=180.33$: $p<.001$), proportion of misses ($F(3,186)=18.44$: $p<.001$) and proportion of false positives ($F(3,186)=13.11$: $p<.001$) with performance being best in each case for the 'WIPE' command, and worst for the 'NONE' command. In order to account for differences in command recognition, which were not of interest for the purposes of this investigation, response times for the command recognition task were subtracted from the appropriate command generation times recorded during the performance of the command generation experimental tasks.

Table 5.01 : Speed and accuracy of performance for command recognition tasks.

	None	Sort	Wipe	Move
Response time (secs.)	1.45	1.29	0.87	1.15
Proportion of misses	0.25	0.17	0.03	0.09
Proportion of false positives	0.15	0.14	0.03	0.08

5.3.2 Typing speed, computer literacy, and cognitive ability

Table 5.02 : Correlation matrix for cognitive ability tests

	SM	ND	LR	MA
Spatial visualisation	.20	.41 ***	.40 ***	.09
Spatial memory		.20	.21	.12
Nelson-Denny vocab.			.41 ***	.06
Logical reasoning				-.02

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

Table 5.02 shows the correlation matrix for the psychometric test scores. Spatial visualisation, vocabulary, and logical reasoning show significant intercorrelations, whilst spatial memory and particularly associative memory show no strong associations with other measures of cognitive ability.

Table 5.03 shows the results of correlations between each of the cognitive ability measures and computer literacy. As can be seen, there is a significant positive correlation between spatial visualisation and computer literacy.

Table 5.03 : Correlations between computer literacy and cognitive ability

	Computer Literacy
Spatial visualisation	.40 ***
Spatial memory	-.08
Nelson-Denny vocab.	.14
Logical reasoning	.15
Associative memory	.10

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

Table 5.04 : Correlations between typing speed and cognitive ability

	Typing speed
SV	-.43 ***
SM	-.09
ND	-.21
LR	-.17
MA	-.08

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

Table 5.04 presents the results of correlations between the typing speed test and the measures of cognitive ability. As with computer literacy, the only significant correlation was with spatial visualisation, with high spatials performing comparatively more quickly.

A series of unrelated t-tests was used to examine differences between the novice and expert groups upon tests of typing speed, computer literacy, and cognitive ability. The results of these tests are shown in Table 5.05. As can be seen, experts scored significantly better on the tests of typing speed, computer literacy, spatial visualisation, verbal ability, and logical reasoning. However, there was a small, non-significant performance advantage for the novice group with respect to spatial memory.

For the purposes of the analysis of individual differences described below, subjects were allocated, post-hoc, to high and low groups for computer literacy and each cognitive ability measure based upon a median split of scores.

Table 5.05 : Mean scores and t-tests for novice and expert groups for typing, computer literacy, and cognitive ability tests.

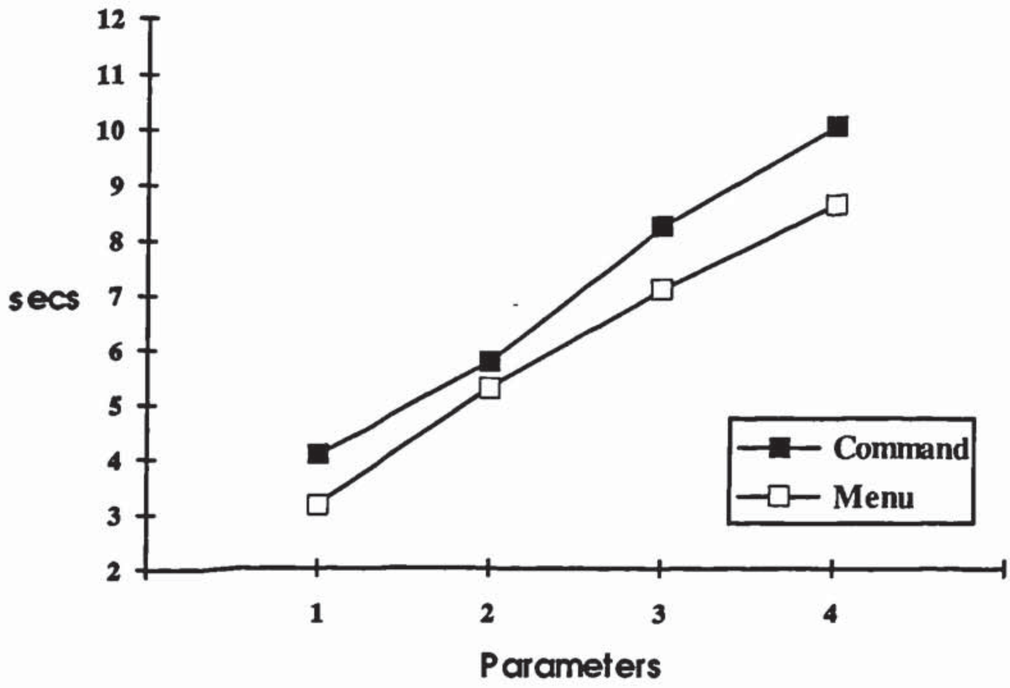
	Novice	Expert	t	p
Typing test (secs.)	4.92	3.34	3.74	=.001
Computer literacy	24.06	28.25	-3.22	<.01
Spatial visualisation	11.52	13.69	-2.38	<.05
Spatial memory	19.39	18.66	.57	ns
Nelson-Denny vocab.	73.34	80.00	-2.11	<.05
Logical reasoning	20.20	23.48	-2.41	<.05
Associative memory	9.92	9.38	-.14	ns

5.3.3 Command generation tasks

Only three subjects accessed the help screen, each on only one occasion. These data have therefore not been included for analysis. The effects of command generation method and number of required parameters were examined using 2 x 4 (command x parameter) ANOVAs for response times and proportion of errors. As mentioned above, response times have been adjusted to take account of individual differences in the process of command recognition.

With respect to response times, there was a main effect of command type ($F(1, 63)=25.53$: $p<.001$), with performance whilst using the pull down menus being superior to performance whilst using the command line for all levels of parameter complexity (see Figure 5.02). There was a significant effect of the number of command parameters, with performance times increasing in line with the number of parameters ($F(1, 63)=371.33$: $p<.001$). There was also a significant interaction between the command method and the number of parameters ($F(3, 189)=5.69$: $p=.001$) which appears to be primarily due to a comparatively greater increase in performance time for the command line interface as the number of parameters increased. However the performance difference between the two interfaces is greater for the one-parameter condition than for the two-parameter condition.

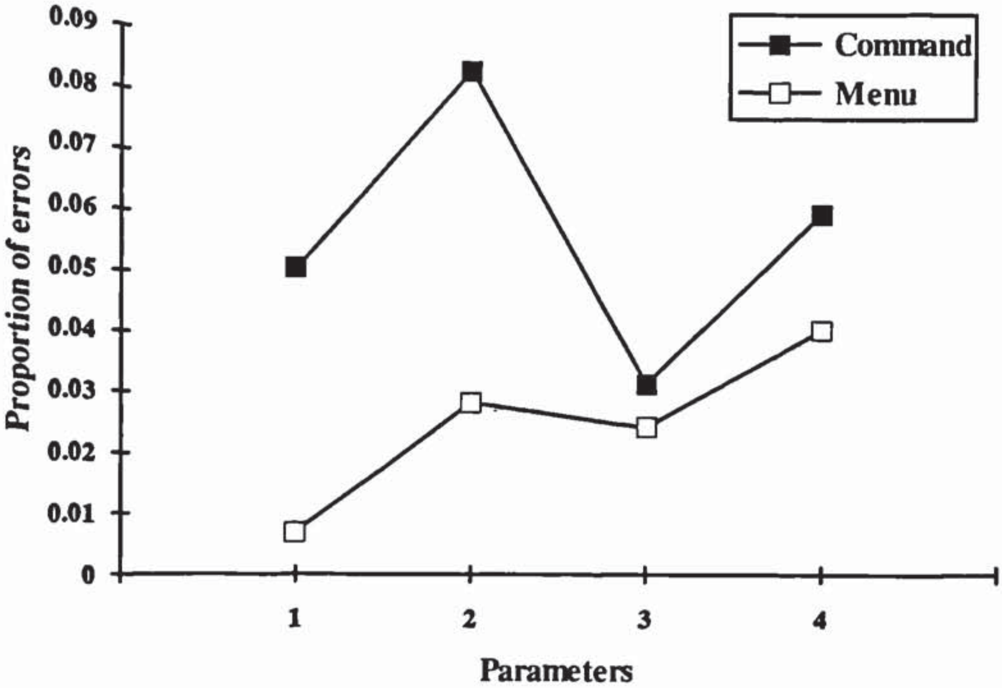
Fig: 5.02 : Response times for the command and menu interfaces



There were very few errors generated during task performance (see Figure 5.03) and, consequently the error distribution was strongly positively skewed. In addition, it was not possible to adjust error performance for individual differences in command recognition as was done for response times. For these reasons individual differences in error performance have not been examined, and analyses presented below relating to command and parameter conditions must be treated with a degree of caution.

There was a significant main effect of command interface ($F(1,63)=20.32 : p<.001$) with performance accuracy being greater when using the menu interface. There was also a significant effect of the number of parameters required ($F(3, 189)=4.79 : p<.01$), in which the greatest number of errors was generated in the two-parameter condition ('SORT'), and the fewest errors were generated in the one- and three-parameter conditions, for which error performance was very similar. The interaction between command interface and the number of parameters required was significant ($F(3, 189)=3.02 : p<.05$) with error performance increasing in line with the number of required parameters for the menu interface, but the greatest number of errors being generated in the two-parameter condition whilst using the command interface.

Fig. 5.03 : Proportion of errors for command and menu interfaces



5.3.4 Individual differences in command generation: Statistical approach

The effects of individual differences upon the performance of the command generation tasks were examined using three sets of analyses. Firstly, separate analyses were conducted for the menu and command interfaces to examine the main and interactive effects of cognitive ability and expertise. 2 x 4 (expertise x parameters required) and 2 x 2 x 4 (cognitive ability x expertise x parameters required) ANOVAs were used for each command condition. The reason for analysing command generation methods separately at this point was to avoid the possibility that a lack of association between cognitive ability or expertise and one command method could mask interesting and significant differences in performance upon the other method. Given that these command methods may be used independently within an interface, a relationship between individual differences in ability and performance upon either one of these methods is of interest. Secondly, a further set of similar ANOVAs were performed for the command line interface condition in which an adjustment was made for the effects of typing speed. As noted above, there were significant differences between the novice and expert groups with respect to typing speed, and also a significant correlation with spatial visualisation. Whilst overall performance in the command line condition is of interest, the cognitive demands of command generation are the primary focus of concern. Consequently for one set of analyses, task performance response times were adjusted for individual differences in typing speed. Average time per keystroke was calculated from the initial typing test, and this figure was then used to adjust response times for each command interface parameter condition proportionately. This method was used in preference to the inclusion of typing speed as a covariate in ANOVA because it takes account of the interactive effects of typing speed and parameters required. Finally, the potential interactive effects of expertise, cognitive ability and command methods were examined using a 2 x 2 x 4 (expertise x command method x parameters required) or a 2 x 2 x 2 x 4 ANOVA (cognitive ability x expertise x command method x parameters required). In these analyses the only results of interest were those involving the command method factor. Even where statistically significant, results which do not include this factor are not reported, as this would largely replicate the other analyses described above.

5.3.5 The effects of expertise

There was a main effect of expertise ($F(1,62)=12.20 : p<.001$) for the menu interface with novices performing more slowly than experts (see Figure 5.04). There was also a significant interaction between expertise and the number of required parameters ($F(3, 186)=6.99 : p<.001$) with the performance disadvantage for novices increasing as the number of required parameters increased. There was a similar main effect of expertise for the command line interface ($F(1,62)=11.70 : p<.001$) and an interaction with the number of required parameters ($F(3,186=6.81 : p<.001$). The nature of these effects was as for the menu interface (see Figure 5.04).

When command line performance was adjusted for the effects of typing speed the effects of expertise ($F(1,62)=3.90 : p=.053$) and the interaction between expertise and the number of parameters ($F(3,186)=2.64 : p=.051$) just failed to reach significance (see Figure 5.05). There was no appreciable difference in performance between the novice and expert groups for the one-parameter condition, and a consistent performance advantage for the expert group for the two-, three-, and four-parameter conditions.

Fig. 5.04 : Response times for novice and expert subjects using command and menu interfaces

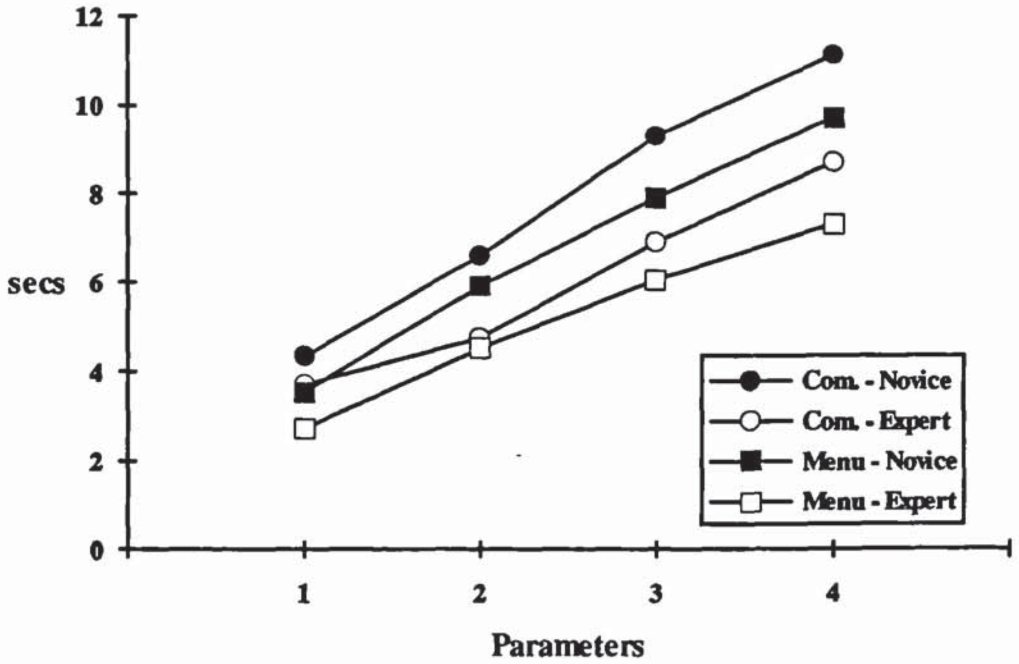
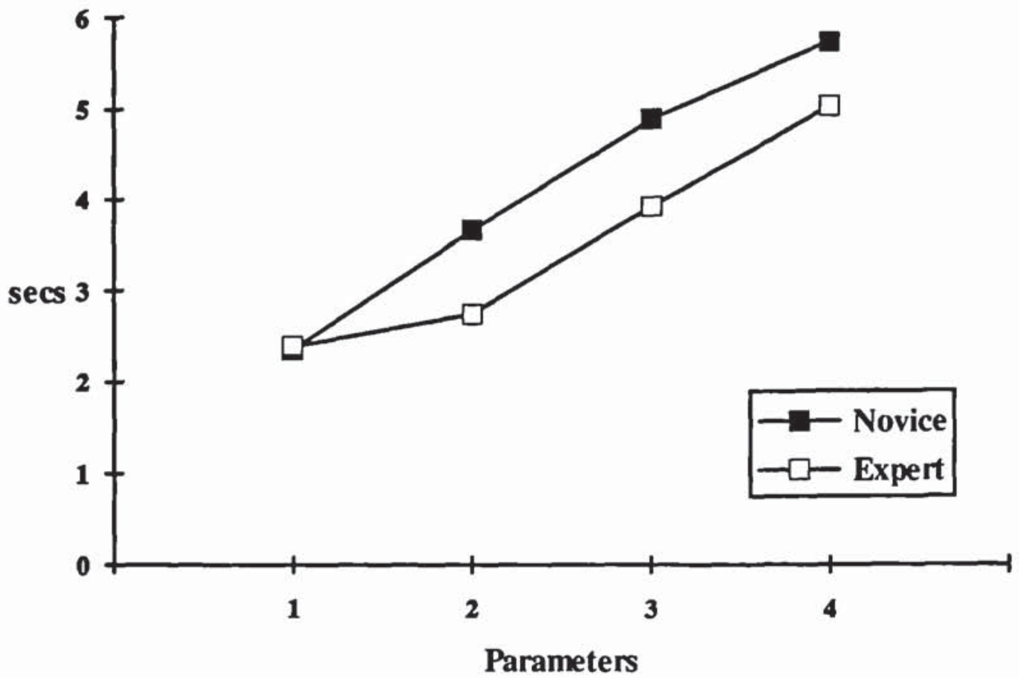


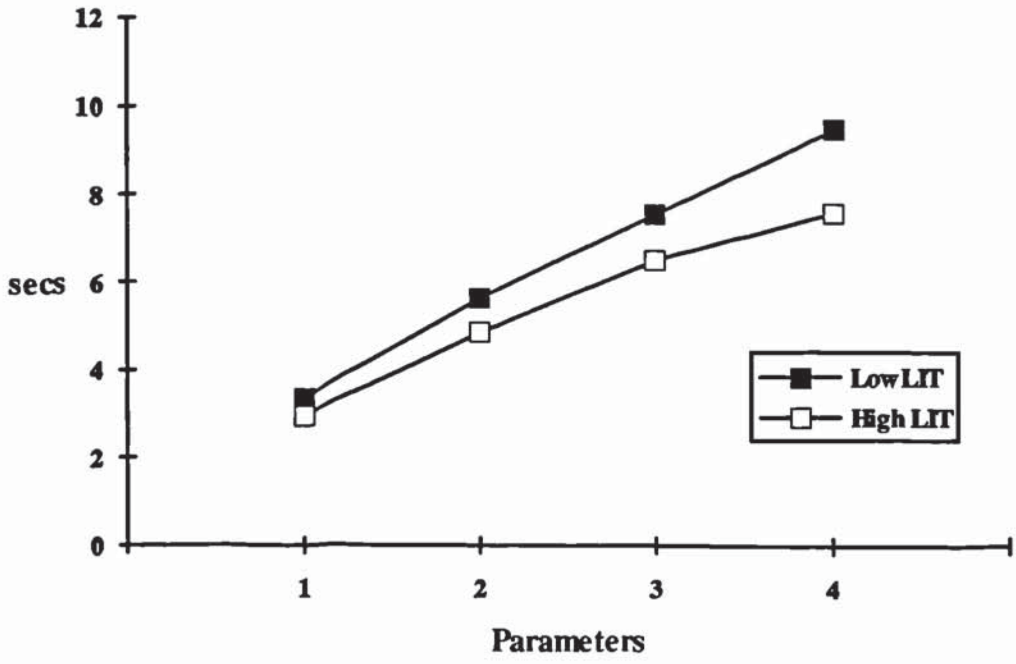
Fig. 5.05 : Adjusted response times for novices and experts using the command interface



When the effects of expertise were examined in relation to command interface, using a full factorial design, there was no significant interaction between these factors.

Similar analyses were used to examine the effects of computer literacy. Expertise was not included as an additional factor because the comparatively strong group differences in computer literacy (see Table 5.01) would result in highly unequal cell sizes. There was a main effect of computer literacy for the menu interface ($F(1,62)=4.31 : p<.05$) with the high computer literacy group performing more quickly. There was also a significant interaction with the number of required parameters ($F(3,186)=6.20 : p<.001$) with the performance advantage of the high computer literacy group increasing in line with the number of command parameters (see Figure 5.06). Neither the main or the interactive effects of computer literacy were significant for the command line interface before or after adjusting for the effects of typing speed.

Fig. 5.06 : Response times for subjects high and low in computer literacy using the menu interface



When the effects of computer literacy were examined in relation to command interface, using a full factorial design, there was no significant interaction between these factors.

5.3.6 Cognitive ability and command generation

The relationship between cognitive ability and command generation performance were examined in conjunction with the effects of expertise.

5.3.6.1 Spatial visualisation

There was a main effect of spatial visualisation for the menu interface ($F(1,60)=10.57$: $p<.01$) with the high spatial group performing more quickly (see Figure 4.07). There was also a significant interaction between spatial visualisation and parameters required ($F(3,180)=3.03$: $p<.05$) with the performance advantage for the high spatial group increasing in line with the number of required parameters. The main effect of spatial visualisation was also significant for the command line interface ($F(1,60)=15.02$: $p<.001$) as was the interaction with parameters required ($F(3,180)=7.41$: $p<.001$). When command interface performance was adjusted for typing speed there was still a main effect of spatial visualisation ($F(1,60)=6.36$: $p<.05$), however, the interaction with parameters was no longer significant (see Figure 5.08).

Fig. 5.07 : Response times for subjects high and low in spatial ability using the command and menu interfaces

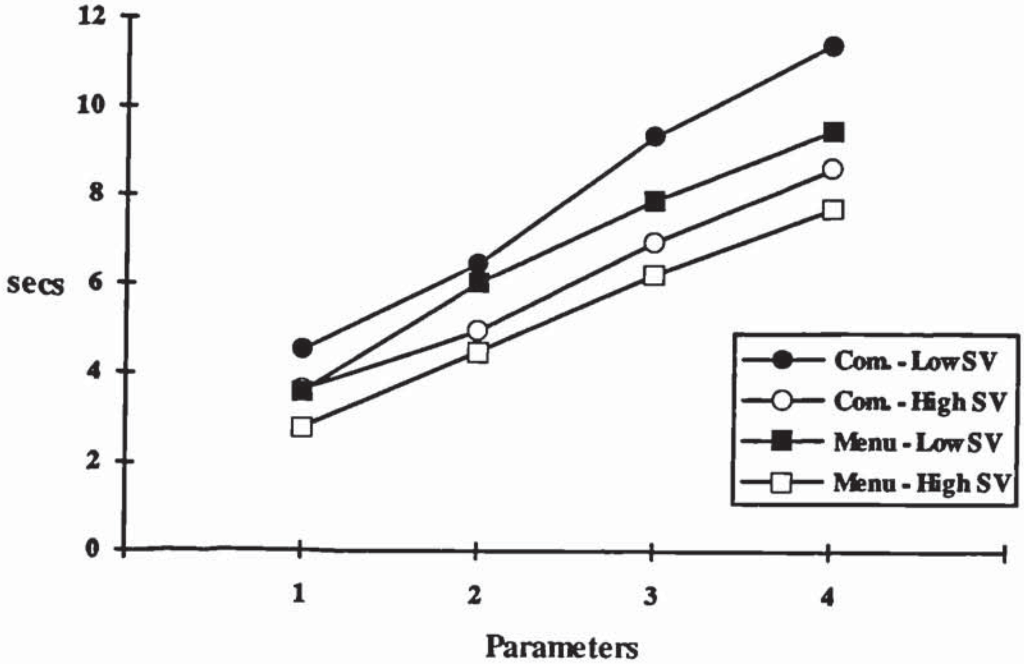
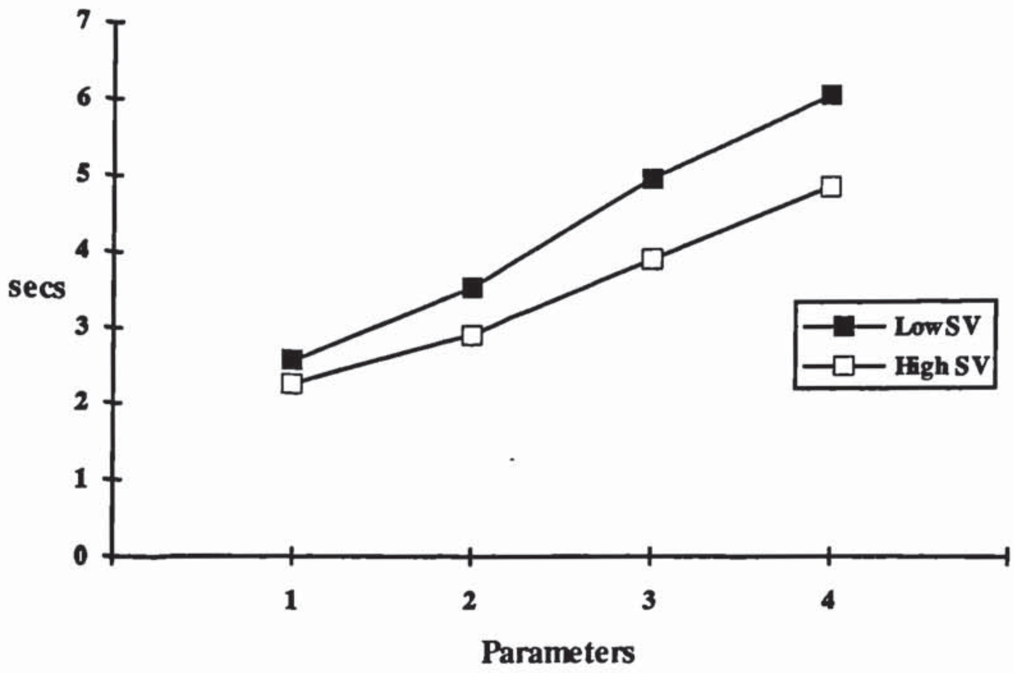


Fig. 5.08 : Adjusted response times for subjects with high and low spatial visualisation scores using the command interface



When the effects of spatial visualisation were examined in relation to command interface, using a full factorial design, there was no significant interaction between these factors.

5.3.6.2 Spatial memory

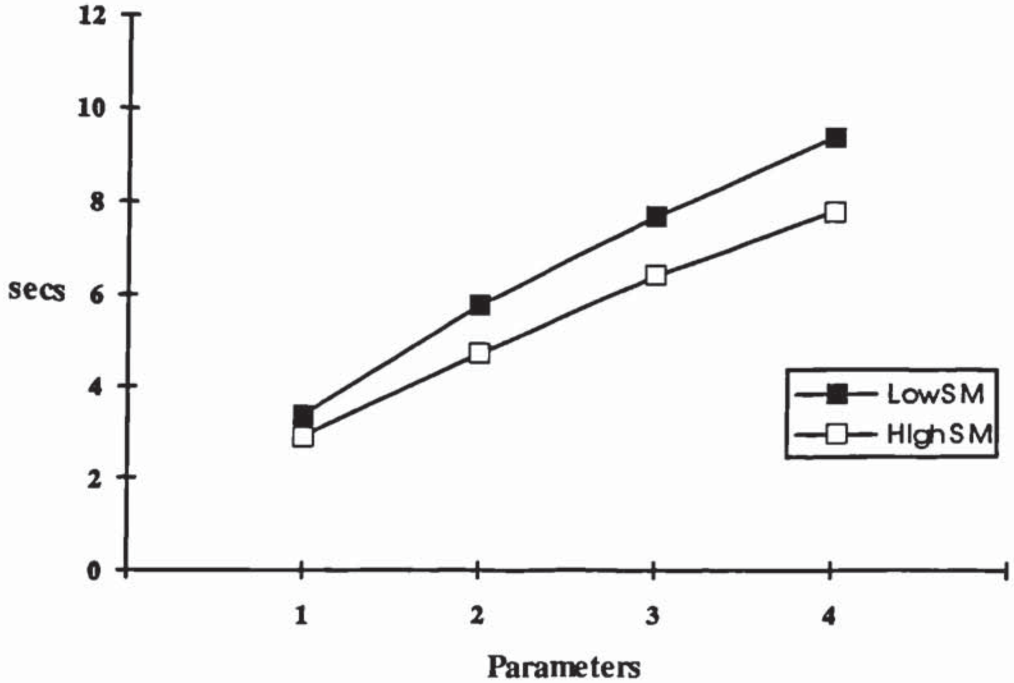
There was a main effect of spatial memory for the menu interface ($F(1,60)=6.37$: $p<.05$) with the high spatial memory group performing more quickly. There was also a significant interaction with expertise ($F(1,60)=4.23$: $p<.05$) such that the performance of high spatial memory novices was similar to that of expert subjects, whilst low spatial memory novices performed considerably more slowly (see Table 5.06).

Table 5.06 : Interaction between spatial memory and expertise for response time (secs) in the menu interface condition.

	Novice	Expert
Low SM	7.850	5.263
High SM	5.852	5.059

The interaction between spatial memory and parameters required was significant for the menu interface ($F(3,180)=3.84 : p<.05$) with the performance advantage of the high spatial memory group increasing in line with the number of required command parameters (see Figure 5.09).

Fig. 5.09 : Response times for subjects high and low in spatial memory using the menu interface

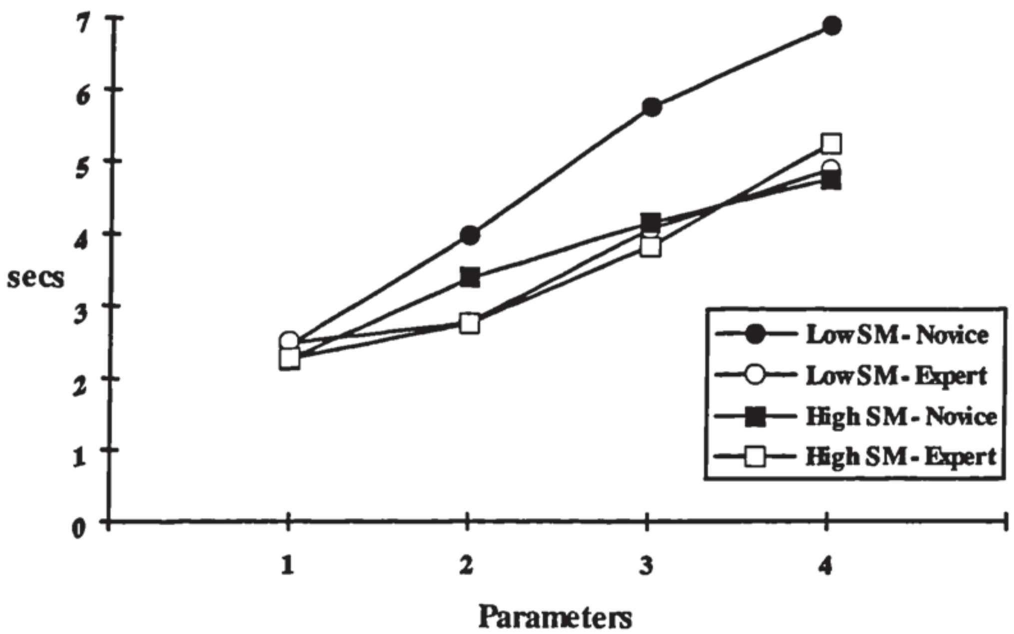


There was also a main effect of spatial memory for the command line interface ($F(1,60)=6.61 : p<.05$) and a significant interaction between spatial memory and parameters required ($F(3,180)=4.61 : p<.01$). The nature of these effects was the same as those for the menu interface. The three-way interaction between spatial memory, expertise, and parameters required was also significant ($F(3,180)=6.09 : p<.001$) with the performance of high spatial memory novices, and high and low spatial memory experts being very similar. However, the performance of low spatial memory novices showed a greater comparative deterioration as the number of required command parameters was increased (see Table 5.07). A similar interactive effect approached significance for the menu condition ($F(3,180)=2.61 : p=.053$).

Table 5.07 : Breakdown of response times (secs) for spatial memory, expertise and required parameters for the command line condition.				
Parameters	1	2	3	4
Low SM - Novice	4.776	7.446	10.955	13.240
Low SM - Expert	3.886	4.832	7.153	8.659
Mean	4.321	6.139	9.054	10.950
High SM - Novice	3.928	5.881	7.885	9.308
High SM - Expert	3.562	4.677	6.703	8.762
Mean	3.745	5.279	7.294	9.035

When command interface performance was adjusted for typing speed the main effect of spatial memory and the interaction with required parameters were no longer significant. However there was still a significant three-way interaction between spatial memory, expertise and the number of parameters required ($F(3,180)=3.98$: $p<.01$). As can be seen from Figure 5.10 performance in the single parameter condition is similar for all groups. However, as the number of parameters increased, there was a greater deterioration in performance for the low spatial memory novice group. When the effects of spatial memory were examined in relation to command interface, using a full factorial design, there was no significant interaction between these factors.

Fig. 5.10 : Adjusted response times for novices and experts with high and low spatial memory scores using the command interface



5.3.6.3 Verbal ability and logical reasoning

There were no main or interactive effects of verbal ability or logical reasoning with respect to either the menu or command line interfaces. This was also true of analyses in which command line performance was adjusted for typing speed, and in full factorial analyses investigating the interactive effects of command method.

5.3.6.4 Associative memory

The main effect of associative memory was not significant for the menu interface condition. However, there was a significant interaction between associative memory and expertise ($F(1,60)=4.27$: $p<.05$). This was due to low associative memory experts performing slightly faster than high associative memory experts, whilst low associative memory novices performed much more slowly than high associative memory novices (see Table 5.08).

Table 5.08 : Interaction between associative memory and expertise for response time (secs) in the menu interface condition.

	Novice	Expert
Low associative memory	7.532	4.997
High associative memory	5.946	5.284

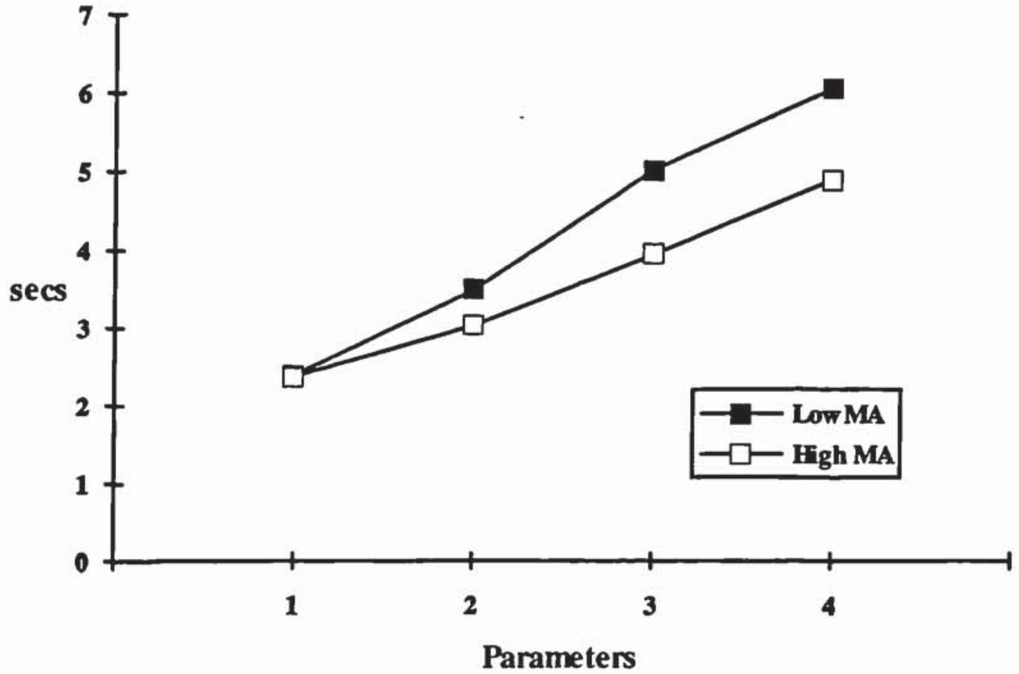
There was a main effect of associative memory for the command line condition ($F(1,60)=4.64$: $p<.05$) with the low associative memory group performing more slowly (see Table 5.08). There was also a significant interaction between associative memory and the number of required parameters ($F(3,180)=5.03$: $p<.01$) such that the performance advantage for high associative memory subjects increased in line with the number of required parameters (see Table 5.09).

Table 5.09 : Breakdown of response times (secs) for high and low associative memory groups and number of required parameters in the command line condition.

Parameters	1	2	3	4	Mean
Low associative mem.	4.160	6.130	8.986	10.913	7.547
High associative mem	3.895	5.308	7.373	9.080	6.414

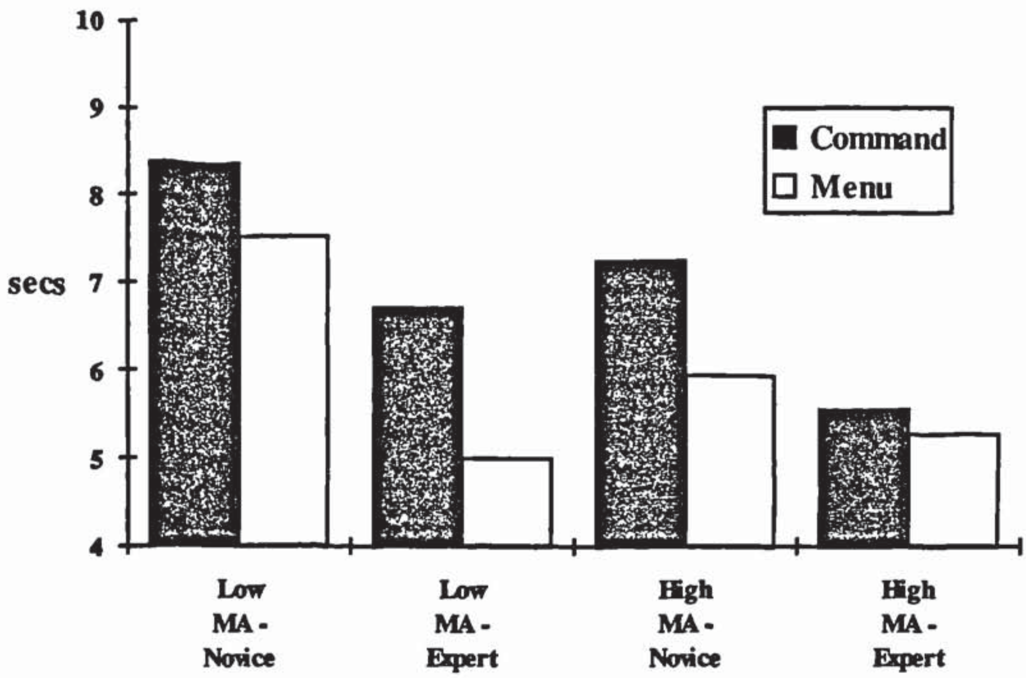
When performance in the command line condition was adjusted for typing speed there was still a similar main effect of associative memory ($F(1,60)=4.74$: $p<.05$) and a significant interaction with the number of parameters required ($F(3,180)=3.75$: $p<.05$) (see Figure 5.11). Whilst performance was very similar for all groups in the single parameter condition, there was a greater deterioration for the low associative memory group as the number of required parameters increased.

Fig. 5.11 : Adjusted response times for subjects with high and low associative memory scores using the command interface



When the effects of associative memory were examined in relation to command interface using a full factorial design there was a significant interaction between associative memory, expertise and command condition ($F(1,60) = 6.75$: $p<.05$). This interaction is shown in Figure 5.12. As can be seen, the performance advantage for the menu interface is much greater for low associative memory experts than for high associative memory experts. Whilst this pattern is reversed for the novice groups, the performance of low associative memory novices is considerably slower for both interfaces.

Fig. 5.12 : Response times for novices and experts with high and low associative memory scores



5.3.7 Workload

Differences in self-report workload were examined for each cognitive ability measure separately, using a series of 2 x 2 x 2 (cognitive ability x expertise x command interface condition) ANOVAs. The only significant effect was a main effect of logical reasoning ($F(1,60)=9.26$: $p<.01$) such that subjects in the high logical reasoning group reported lower overall workload than subjects in the low logical reasoning group (see Table 5.10).

Table 5.10 : Mean self-report workload for high and low logical reasoning groups

	Command	Menu	Mean
Low Logical reasoning	53.44	51.33	52.39
High Logical reasoning	42.00	41.34	41.68

5.3.8 Strategy differences

The main and interactive effects of interface and parameters were examined using a 2 x 4 ANOVA. The effects of expertise and computer literacy were examined separately using 2 x 4 (expertise x parameter) ANOVAs. In these instances the interface factor was not required because the total number of (error free) commands generated is constant for all subjects and knowledge of frequency of use for either command enables a comparison to be performed. Similarly, the effects of cognitive ability were added to the latter design such that a 2 x 2 x 4 (ability x expertise x parameter) ANOVA was used.

There was a main effect of interface condition ($F(1,63)=11.94$: $p=.001$) with subjects using the menu interface more frequently than the command line (see Table 5.11). The main and interactive effects of parameter were not significant.

Table 5.11 : Command selection strategy differences for each parameter condition

Parameters	Command	Menu
1	2.969	6.031
2	3.031	5.969
3	2.844	6.156
4	2.484	6.516

The main effect of expertise was not significant either. However, there was a significant interaction between expertise and parameters required ($F(3, 186)=2.67$: $p<.05$) such that the frequency with which novices used the command line was relatively invariant across the four levels of the parameter factor (see Table 5.12). However, the frequency with which experts used the command line varied in direct relation to the number of parameters, with frequency reducing as command complexity increased.

Table 5.12 : Number of commands issued using the command line by novice and expert subject groups.

Parameters	Novice	Expert
1	2.031	3.906
2	2.187	3.875
3	2.125	3.563
4	2.125	2.844

The were no significant main or interactive effects of computer literacy, spatial visualisation, spatial memory, verbal ability or logical reasoning.

There was a significant interaction between associative memory and expertise ($F(1, 60)=5.78 : p<.05$) such that novices with low associative memory tended to use the command line more frequently than experts with low associative memory, whereas experts with high associative memory used the command line much more frequently than any other group (see Table 5.13).

Table 5.13 : Command selection strategy breakdown by expertise and associative memory group.

	Low MA	High MA
Novices	2.588	1.584
Experts	1.500	4.948

5.4 Discussion

The response time advantage which was apparent for the menu system is contrary to the results of Whiteside et al. (1985) and Antin (1988). This may be attributable to the small command set used in the present experiment, which resulted in a comparatively small range of menu options. However, this must be offset against the number of menu panels to be traversed when generating the more complex commands. As the interactive effects of 'interface type' and 'parameters required' show, increases in complexity proved comparatively less costly for the menu interface. An alternative explanation, which was addressed earlier, relates to design flaws in the implementation of the menu systems used in some of these earlier studies. It may be that the mechanics of menu command generation in these cases unduly biased the results. An interesting area for future research would be to investigate the performance time trade-off between the number of menu options on a panel and the number of command line arguments required. The increased performance accuracy when using the menu is in keeping with the syntactic guidance which is provided by such interface designs (Norman, 1983). This performance advantage also increased in line with the complexity of the command. It would appear that subjects were aware of the performance advantages associated with the menu interface, as when they were given the choice of command generation method in the strategy condition, significantly more commands were generated using the menu. Given these performance differences it is interesting to note that there was no significant effect of interface condition for self-report workload.

5.4.1 Expertise

There was a main effect of expertise for both interface conditions which interacted with the number of required command parameters, supporting the hypothesis that novices would be subject to a comparatively greater performance disadvantage in conditions of high cognitive load. However, there was no interaction involving expertise and interface type, indicating that whilst novices perform more slowly this is a relatively constant effect across both command and menu interfaces. Whilst this is contrary to the experimental prediction based upon the cognitive demands associated with each interface, it is consistent with the results of Antin (1988). The fact that typing speed appears to account for much of the variance associated with expertise in the command line condition further accentuates this position. It is interesting to note, however, that when the effects of typing speed were removed there was a relatively constant performance difference between novices and experts in all but the most simple command parameter condition.

Whilst there was no main effect of expertise with respect to performance strategy, the trend indicated that the expert group preferred to use the command interface. This is contrary to the results of Antin (1988). There was a significant interaction between expertise and 'parameters required'. Novice command selection strategy remained relatively invariant regardless of the number of parameters required. However, experts appear to have adjusted command selection upon the basis of command complexity. When complexity was low they selected the command line method of generation more frequently than when complexity was high. Whilst this pattern of results is consistent with the experimental hypotheses, there is no ability related basis for this strategy. Experts consistently performed more quickly when using the menu interface. It is possible, however, that this is a transfer effect and the result of a strategy learned from other systems. Instead of applying a 'default method' (Card, Moran, and Newell, 1983) experts may be applying a default strategy.

5.4.2 Cognitive ability

There was a main effect of spatial visualisation for both menu and command line interface conditions, and also a significant interaction with the number of parameters required in both cases such that the performance advantage of high spatial ability subjects increased in line with command complexity. This latter finding is consistent with the results of Davis and Bostrom (1992). However, contrary to experimental prediction there was no significant interactive effect which included interface type. This does not support the hypothesis (Egan and Gomez, 1985) that the ordering of command symbols is an important factor in the association between spatial ability and command generation performance. If this were the case, then spatial ability should be more strongly associated with performance in the command line interface condition, where the sequence of command symbols is subject to greater flexibility. When the effects of typing speed were controlled, the main effect of spatial visualisation was still significant but the interactive effect of parameter complexity was no longer significant. It would appear that the main effect of spatial ability for this interface condition cannot be attributed to differences in typing speed.

In keeping with the results of Egan and Gomez (1985), there was a main effect of spatial memory. The interactive effects of parameters required supports the experimental prediction that this association would increase in strength in line with command complexity. However, as was the case for spatial visualisation, there was no significant interaction involving interface type. The consistency of these results would appear to indicate that the command complexity variable is distinct from the hypothesised differences in cognitive demand associated with each interface.

There was a significant interaction between spatial memory and expertise such that the performance of high spatial memory novice subjects was similar to that of the expert subjects, whilst the performance of low spatial memory novices was comparatively worse. This only became apparent in the command line interface for the conditions of greater command complexity, and could not be attributed to differences in typing speed. There was little difference between the scores of experts and novices upon the test of spatial memory. Consequently, these results cannot be attributed to a curvilinear relationship between spatial memory and performance. These results are also consistent with the theory of individual differences in skill acquisition put forward by Ackerman (1988) such that the effects of cognitive ability upon task performance are greater for novices than for experts.

The only significant effect of logical reasoning was a main effect for self-report workload such that subjects in the low logical reasoning group reported greater workload than subjects in the high logical reasoning group. This lack of significant performance difference is surprising given the association between logical reasoning and command generation found in the experiment reported in Chapter 2. This apparent contradiction may be explained if, instead of allowing performance to deteriorate, the low ability group allocated additional attentional resources to the task.

The main effect of associative memory was not significant for the menu interface condition. However, there was a significant interaction with expertise, with low associative memory novices performing much more slowly than high associative memory novices, whose performance was slightly slower than the expert groups. Once again this supports Ackerman's (1988) theory, with the strongest association between cognitive ability and performance being found for the novice group. There was a main effect of associative memory for the command line interface condition, along with a significant interaction with parameters required. This was not an artefact of typing speed. It would appear that the predicted effects of expertise, i.e. comparable menu performance for novices but superior performance of experts in the command line condition, may be heavily dependent upon the level of associative memory skills in the novice group. However, this was only partially supported in the significant interaction between associative memory, expertise, and command interface. Low associative memory experts were found to be the fastest performers in the menu condition, but were disproportionately slower for the command line interface. There was a strong performance advantage for the menu interface in the case of high associative memory novice subjects, whilst the low associative memory novice group performed more slowly with both interfaces. These difference in performance speed are very much in keeping with differences in cognitive ability for the expert group, and this may reflect differences in typical command usage. If low associative memory experts match their computer interactions to their cognitive capabilities, then they may become comparatively more skilled in the use of menu interfaces. This pattern would not be expected to be apparent for novice users, and as a result high associative memory novices may have yet to develop suitable 'micro' strategies to allow their performance in the command line interface condition to reflect their cognitive ability. Further to this, expertise related differences in typing speed will also be reflected in this interaction.

This practice hypothesis is supported by the significant interaction between associative memory and expertise with respect to the type of commands used in the strategy condition. The high associative memory expert group used the command line interface much more frequently than the low associative memory expert group or either novice group. Also, the low associative memory novice group used the command line more frequently than the low associative memory expert group. Possibly this latter relationship reflects differences in the efficiency with which strategy is modelled upon ability, such that the expert group have a clearer understanding of their most effective strategy options.

5.4.3 Conclusions

In summary, contrary to several previous studies, there was a performance advantage for the menu interface which increased in line with command complexity. The predicted interaction between expertise and interface type was not supported, with a similar performance advantage for experts in both interface conditions. However, it would appear that the effects of spatial memory, and in particular associative memory, are central factors in this relationship with novice subjects obtaining low scores on these cognitive ability tests performing particularly poorly. These findings are consistent with Ackerman's (1988) theory of individual differences in skill acquisition. Spatial visualisation and spatial memory were both predictive of performance in both interface conditions, and also interacted with command complexity in the predicted direction. Given the weak correlations between these factors and associative memory, it would appear that both spatial and general memory second-order ability factors may be associated with the process of command generation. However, the fact that there were no significant effects of verbal ability suggests that a more general intelligence factor cannot explain these differences. An interesting area for future research would be to examine the potential interactive effects of anxiety and memory during the process of command generation (cf. Eysenck, 1976).

Chapter 6

The 'Generating' Component: Icons vs Text

6.1 Introduction

This chapter presents an experimental investigation of individual differences in the use of icons and text labels. A search and select paradigm was used to examine performance in relation to spatial and verbal ability.

Over recent years the use of icons within the human-computer interface has become increasingly prevalent (Maguire, 1990). Essential to this development has been a growth in computer processing and display capabilities (Smith, Irby, Kimball, Verplank, and Harslem, 1982; Gittings, 1986). These changes have been closely associated with the advent of direct manipulation interfaces (DMI : cf. Hutchins, Hollan, and Norman, 1986) and the use of 'real world' comparisons for computer task environments such as the 'office' or 'desktop' metaphors (Smith et al. 1982). Computer programming languages have made similar progressions with moves toward object oriented programming (Morrill-Tazelar, 1990), encouraging the reification of previously abstract concepts, in what is seen as a means to increase the ease with which computer systems are learned and used (Rogers, 1989). Icons are central to many such processes, providing an important means of access to program objects.

There are a number of advantages thought to be associated with the use of icons, many of which relate to hypothesised differences in the fundamental processing mechanisms associated with verbal and non-verbal material. Support for such a distinction comes from several sources, and relates to the ongoing controversy over the nature of mental representation, as discussed in Chapter 4. Paivio (1971) proposed that there are distinct verbal and non-verbal processing codes which relate to separate but inter-connected systems of long term retention. An experimental method which has been used to support this hypothesis involves the comparison of specific attributes, e.g. size or weight, associated with objects presented as pictures and text. Making relative assessments of 'real life' size for two objects presented as pictures (e.g. a mouse and an elephant) is usually found to be quicker than when the same objects are presented as text (Paivio, 1991). Further to this, greater processing disruption occurs when the stimuli are presented as pictures with sizes that contradict 'real life' than when the same is done with text. This led Paivio to suggest that the transfer of information from the non-verbal system to the verbal system is more efficient than the reverse. A similar distinction is employed by Baddeley (1986) in his model of working memory which incorporates separate phonetic and visuo-spatial components, and by Wickens (1984b) in his model of attentional resources.

Icons are held to provide a means of conveying large amounts of information in a concise manner (Barker, Najah, and Manji, 1987; Gittings, 1986; Rogers, 1989; Thimbleby, 1990). They enable a universality of communication which may cross language barriers and application packages (Rogers, 1989), they may be used to represent classes of objects (Ziegler and Fahnrich, 1988), and many of the methods associated with their use (e.g. DMIs) reduce syntax errors (Gittings, 1986). Recognition memory for pictures has been found to be superior to that for words (Sheppard, 1967) and this is also thought to generalise to the use of icons in the human-computer interface (Blankenberger and Hahn, 1991). However, this memorial advantage is not apparent when recall is required. Lansdale (1988), conducted two experiments in which icons were used to facilitate database access from which he concluded that whilst total recall of icons was not subject to great performance advantage, the shape, location, and colour of icons appeared to be separately encoded, thereby facilitating partial recall. Blankenberger and Hahn (1991) also identified 'articulatory distance' (cf. Hutchins, Hollan, and Norman, 1986) as being an important factor in successful recall. However, for icons, comprehensibility is a double edged sword. Icons can lack the precision of language (Rogers, 1989), and acceptable representation of abstract concepts may prove difficult (Rogers and Osborne, 1987). One method of overcoming this problem is to combine icons and text labels. Support for this type of design is provided by Guastello, Traut, and Korienek (1989) who found that subjects rated combined icons and text labels to be more meaningful. Similarly, Kacmar and Carey (1991) found that menu selection performance was less error prone, although no faster, when items were composed of both text and icons. Merwin, Dyre, Humphrey, Grimes, and Larish (1990) investigated the effect of adding icons to text labels in an information retrieval task. There was no significant improvement in performance times, however, consistent with the findings of Guastello et al. (1989), subsequent recall of the database was significantly greater for subjects who had performed in the text with icons condition.

Despite the number of apparent advantages associated with the use of icons, empirical support for their use within the computer interface is somewhat mixed. Muter and Mayson (1986) examined the effects of adding icons to videotex menus. No improvement was found for information retrieval response time when compared to a text only condition, but error rates were significantly reduced. Confirmation of these results is provided by a similar experiment conducted by MacGregor (1992) which examined the effect of adding either icons or additional text descriptors to a videotex system. No response time differences were found but both icons and descriptors produced a significant reduction in errors when compared to the original system. Kerr (1990) investigated the use of text, icons, and colour as navigation aids to a college catalogue database. No significant performance differences were found. Questionnaire measures of awareness of cues, ease of use, and satisfaction with the interface indicated that subjects were initially more aware of the text and icon cues; but in a follow up test, six weeks later, subjects from the colour cues condition were found to have better recollection of database attributes, and recalled greater satisfaction and ease of use. In a similar study, Lansdale, Simpson and Stroud (1990) compared text and icons as 'cue enrichers' in an information retrieval task. In addition the effects of 'choice' and 'no choice' in allocating text and icons to advertisements was also examined. Consistent with Lansdale's (1988) earlier findings, subject choice in the allocation of 'enricher' was an important factor in the retrieval of the item. Lansdale et al. were unable to draw any firm conclusions with respect to advantages of either text or icons when used for this purpose. However, they suggest that the evidence points to equivalent levels of performance associated with each form of 'cue enricher'.

In an experiment mentioned in Chapter 5, Whiteside, Jones, Levy, and Wixon (1985) compared novices, experts and transfer users whilst performing a file manipulation task when using command line, menu and iconic applications. The performance of the novices and transfer users was worst for the iconic applications. These results are contradicted by those of Davis and Bostrom (1992) who compared novice performance whilst using command line and DMIs. A performance advantage was found when using the DMI. Similar results were obtained by Rauterberg (1992), who compared the performance of novice and expert users upon an information retrieval task using either a menu interface or a DMI. A performance superiority for both groups was found for the DMI. In a comparison of performance using word processors with and without iconic interface components, Rohr (1986) reports lower levels of help requests for the iconic condition in the first of two experimental trials. In a second experiment time spent in error was found to be less for the iconic interface when subjects were performing complex tasks, but the reverse applied for simple tasks. Rohr attributes this to the importance of relational (spatial) coding when task complexity is high (see Chapter 4). However, as Benbasat and Todd (1993) point out, these studies confound the use of icons with the process of command generation. In two experiments Benbasat and Todd used an electronic mail project management task to consider these interface elements independently. Whilst there was a performance superiority for DMI conditions over menu interface conditions during early trials, the effect of using icons or text was not significant.

6.1.1 Search and select paradigms

A small number of studies have used a search and select paradigm to directly compare responses when using icons to those when using text. Arend, Muthig, and Wandmacher (1987) compared performance whilst matching either simple abstract icons which differed only in global features, standard set of 'desktop' icons, or a set of text labels. Simple icons were found to be selected more quickly than either text or the more complex 'desktop' icons, and this effect increased with the size of the menu to be searched. There was no evidence of a speed-accuracy trade-off which would account for this. Contrary to this, in an experiment mentioned above, Kacmar and Carey (1991) found no difference in response times when selecting icons, text, or a combination of both, although less errors were made in the combined condition. This might be due to the fact that menu selection was made upon the basis of a text description of a required computer function, as opposed to a matching procedure. It may be that the variance in reading and comprehension times outweighed between condition differences.

Lansdale, Jones, and Jones (1989) conducted two experiments in which the speed of search was compared for icons and text. Subjects were required to search a matrix of icons or text, each cell of which represented two attributes of a computer document. The target consisted of one or two attributes displayed prior to the onset of the field, and corresponded to the mode of field presentation (icon or text). Using a touch screen, subjects were required to select each document which matched the target attributes. Results showed significantly faster reaction times in the icon condition with no cost to the number of false positives generated, but a non-significant trend towards greater missed targets in the icon condition. Lansdale et al. (1989) suggest that this decrease in reaction time may be attributable to the use of peripheral vision to identify icons, enabling a degree of parallel processing, whilst serial search was required in the verbal condition. In order to test this hypothesis Lansdale et al. (1989) conducted a second experiment in which subjects were instructed to use a serial search strategy for both icons and text conditions. They found no significant difference in the speed with which icons were scanned between the first and second experiment, whereas there was a significant increase in the speed of text search. Icons were still searched faster than text. They interpret these results as supporting the importance of serial search for text, based upon the assumption that their instructions reduced the number of unstructured searches in the text condition.

Blankenberger and Hahn (1991) examined the effects of iconic articulatory distance (Hutchins, Hollan, and Norman, 1986) upon a search and select task. An initial subject sample generated three different sets of icons for a series of text editing commands. A second subject sample was then required to perform a sorting task in which icons and commands were matched. The degree of conformity with which each icon was allocated to a particular command across the subject sample was then used to form three groups of icons relating to articulatory distance (near, middle, and far). In other words, if an icon was frequently held to be representative of a particular command, the articulatory distance associated with this icon could be said to be 'near'. Conversely, if an icon was rarely associated with a particular command the articulatory distance could be said to be 'far'. A final subject sample was then used to compare search and select performance whilst using these three new icon sets and one set of word descriptors. Subjects were presented with one of the text editing commands and required to search and select the appropriate icon or text descriptor. Icons were selected faster than text, with speed differences indicating an advantage when the articulatory distance was small. However, performance accuracy was greatest when using the text descriptors. In a further experiment, constant screen positioning was used for icons and text. No differences were found between any of the sets in this latter experiment.

Scott and Findlay (1991) used a matching task in order to examine the performance of subjects whilst searching text and icon fields. Subjects were presented with a central stimulus surrounded by a field of 20 items. They were required to press a button as soon as they had visually located the target, and then had to verbally identify the position of the target within the field. Icons were all selected from those "in computer usage" (p. 247). Four conditions were employed: simple words (four characters), complex words (six characters), simple icons ('relatively simple visual structure'), and complex icons ('icons of a relatively complex pattern'). Scott and Findlay found that icons were matched more quickly than words, supporting the findings of Lansdale et al. (1989), and that simple icons/text were searched more quickly than complex icons/text. There was no significant difference in performance accuracy between icons and text, but more errors were made when locating complex targets than simple ones. In order to examine the hypothesis (Lansdale et al., 1989) that icon search times are reduced by the use of peripheral vision a follow up study was conducted. A small sample performed the same experimental task and a scleral coil was used to monitor eye movement. Results supported the use of peripheral vision enabling a more efficient search pattern, involving fewer saccades, to be used for icons and for simple targets. These search and select findings are also supported by experiments using the attribute comparison method described above in relation to Paivio's dual code hypothesis. Pictures have generally been found to be processed more quickly than text (Pellegrino, Rosinski, Chiesi, and Siegel, 1977; Hogaboam and Pellegrino, 1978).

In summary, whilst there are a number of performance advantages thought to be associated with the use of icons, empirical studies present a less than emphatic picture. The use of icons as information retrieval aids has not proved consistently beneficial to performance, although it would appear that when added to text labels some reduction in error may occur. It may be that the number of attributes associated with retrieval cues is an important factor in facilitating recall. Comparison of iconic systems with other forms of interface has provided mixed although generally favourable results, however recent empirical evidence suggests that the important variable in these comparisons may be the use of direct manipulation techniques and not the use of icons. One area in which superior performance has consistently been apparent is in studies which have used the search and select paradigm to compare icons and text. In these studies response times are faster for icons, although this may be at the cost of some reduction in performance accuracy.

6.1.2 Individual differences relating to the processing of icons and text

Very little attention has been given to exploring individual differences in the use of icons and text. Experiments conducted by Jennings, Benyon, and Murray (1991) and Davis and Bostrom (1992) indicate that spatial ability is not significantly related to performance when using DMIs. However, as mentioned above, such findings cannot necessarily be attributed to the use of icons in the interface and may be concerned with other elements associated with the DMI. If substantial individual differences do exist with respect to the relative processing of icons and text, this might help to explain why, when so many advantages are thought to be associated with the use of icons, comparatively little in the way of performance benefits are apparent. The available research evidence is less than conclusive, but there is some evidence to support the hypothesis that spatial ability is more strongly related to the processing of icons, whilst verbal ability is more strongly associated with the processing of text labels.

Loo (1978) investigated the effects of field-independence / field-dependence (FI-FD) upon the speed of recognition of road traffic signs. As mentioned in the introductory chapter, the FI-FD construct has been criticised (cf. McKenna, 1984) and may in fact relate to individual differences in spatial ability. Subjects were presented with slides of verbal or symbolic traffic signs which were either embedded (presented with a street scene background) or non-embedded (presented with a plain background). FI-

FD was found to be predictive of overall performance, with field independence being associated with faster responses. In addition there was an interaction with the slide background condition such that the performance advantage for field independents was greater when the signs were embedded. However, there was no interaction with verbal vs symbolic sign type. The fact that the semantic content of verbal and symbolic signs was different makes this last result difficult to interpret. In addition, responses were made vocally and this may be more compatible with the processing requirements of verbal signs (Wickens, Sandry, and Vidulich, 1983), although whether such an effect would interact with FI-FD is not clear. Consistent with this position, Mumaw and Pellegrino (1984) report an experiment in which spatial ability was found to be significantly related to the speed of search in a task which required shape comparison. Dupree and Wickens (1982) investigated the relationship between spatial ability and performance of a physical identity letter matching task. Three target conditions were employed in which either the target alone was presented, the target was presented with compatible noise (three instances of the same letter presented on either side), or the target was presented with incompatible noise (three instances of a dissimilarly shaped letter presented on either side). High spatial subjects performed better than low spatial subjects, and were not affected by the target presentation conditions, however performance for low spatial subjects decreased with the level of stimulus/noise compatibility. Dupree and Wickens (1982) also report an experiment in which performance upon a visual shape comparison task was analysed. Spatial ability was found to be correlated with the slope relating response time to stimulus similarity, with high spatial ability being associated with quicker judgements when shape similarity was high. Further evidence suggesting an association between spatial ability and the processing of iconic information is provided by Egan (1979) and Mumaw, Pellegrino, and Glaser (1980 : cited in Cooper and Mumaw, 1985, p. 75) who examined the relationship between spatial ability and the component elements associated with spatial transformation tasks. They found that spatial visualisation was primarily predictive of the process of encoding, and of the efficiency of the mental representation of spatial information. Similarly, Paivio (1991) reports experiments in which individual differences in spatial ability were found to be predictive of the difference in time to make a symbolic comparison, as previously described, for pictures and words. With respect to the hypothesised association between verbal ability and the processing of text labels, Hunt, Lunneborg, and Lewis (1975; Hunt, 1978) report a number of studies in which verbal ability was shown to be associated with the speed of lexical symbol access and manipulation. Tasks such as the Posner, Boies, Eichelman, and Taylor (1969) letter matching paradigm indicated a performance superiority for subjects of high verbal ability, which was strongest in conditions which required name identity matching of letters. However, contrary to these findings of increased speed of lexical access for high verbals, Hogaboam and Pellegrino (1978) investigated individual differences in semantic processing of pictures and words and found no relationship between verbal ability and performance. As mentioned in Chapter 4, Hunt (1978) also presents evidence which suggests that individual differences in verbal and spatial ability differentiate the method of encoding verbally presented spatial relational information. High verbal subjects were thought to encode this information propositionally, whereas high spatial subjects were thought to encode the same information spatially.

These experiments do not directly address individual differences in the relative processing of icons and text. However, they do generally support an association between cognitive ability and the processing of verbal and non-verbal information, which is not inconsistent with a differential processing hypothesis. This experiment was designed to investigate whether spatial ability is more strongly associated with the processing of icons than text, whilst the reverse is true with respect to verbal ability. It may be that the use of icons is beneficial for certain individuals whereas the use of text is beneficial for others. A search and select matching task was used, incorporating four main experimental conditions in which subjects either searched for an icon target in an icon field, for an icon target in a text field, for a text target in an icon field, or for a text target in a text field. In addition to measures of performance speed and accuracy, self-report workload was also assessed in order to further examine individual differences in the attentional resource requirements of each experimental condition. It was predicted that subjects with high spatial ability would perform more efficiently when processing spatial information, whilst subjects with high verbal ability would perform more efficiently when processing verbal information.

6.2 Method

6.2.1 Subjects

48 subjects aged between 18 and 30 (mean 20.71 years) were recruited from the student population of Aston University. Equal numbers of male and female subjects were randomly allocated to each experimental sequence. Subjects had not taken part in any of the previous experiments, were right handed, reported vision which was normal or corrected to normal, and spoke English as their first language.

6.2.2 Measures of individual differences

Subjects initially completed tests of spatial ability (VZ2 - Ekstrom, French, and Harman, 1976), verbal ability (Nelson-Denny vocabulary test), and computer literacy (CALIP; Poplin, Drew, and Gable, 1984).

6.2.3 The experimental task

The experimental task was presented on an Acorn Archimedes computer. Subjects were seated at approx. 80 cm from the CRT, with a mouse and mouse pad situated directly in front of them.

Fig. 6.01 : Text display

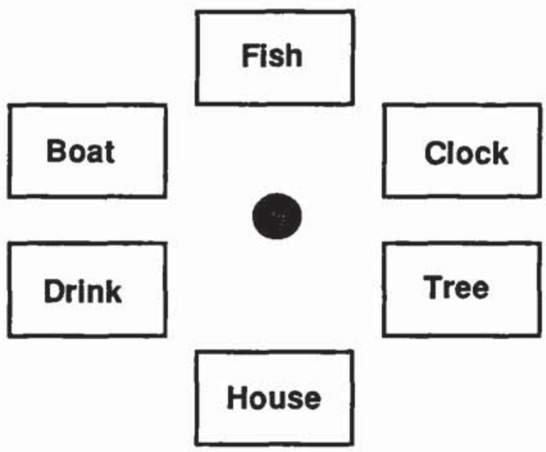
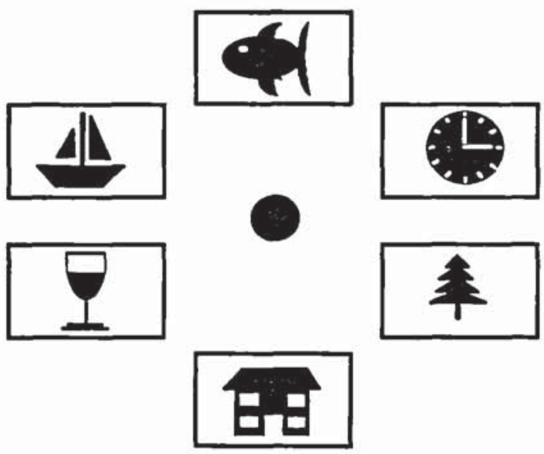


Fig. 6.02 : Icon display



Four within subjects search and select conditions were employed in which the initial target stimulus could be either an icon or text label, and the field to be searched could also be composed of either icons or text labels. These performance blocks were presented in a counterbalanced sequence, with subjects completing the NASA TLX measure of self-report workload following each condition. For the sake of brevity, when referring to these conditions, the target type will be given first followed by the field type. Therefore 'icon-text' would refer to the condition in which an icon target was presented and subjects were required to search a field composed of text labels. The screen layout for text and icon search conditions is shown in Figures 6.01 and 6.02, and is described below. Twelve practice and 36 performance trials were used in each condition, with each target occurring at each position within the circle, for each stimulus during the performance section. Given this constraint, the sequence with which targets appeared at each point was random.

Each trial began with a target stimulus (selected in a random sequence from the six alternatives) being presented in the centre of the screen for a duration of one second. This was followed by an 'offset' (a clear screen) for a period of one second in order to ensure that target images were not being retained in iconic memory (Eysenck and Keane, 1990). The field of six targets was then presented, arranged in a circle subtending an angle of approximately four degrees, surrounding a central positioning dot, as shown in Figures 6.01 and 6.02. Subjects were required to first select the central positioning dot with the mouse in order to ensure a constant starting distance. They were then required to select from the six field items the one which matched the target.

The stimuli used in this experiment were loosely derived from the Battig and Montague (1969) Category Norms. Six items were selected (boat, fish, house, tree, clock, drink) based upon the criteria that each word was four or five letters long, had one syllable, and was capable of being depicted by a simple icon. As far as possible icons were matched for complexity, with each consisting of a relatively simple 'global' shape (Arend, Muthig, and Wandmacher, 1987). An attempt was also made to ensure dissimilarity of global features between icons.

A mouse was chosen as the method of target selection because of its widespread association with iconic interfaces. However, previous experiments have indicated that mouse performance may be related to spatial ability (Taylor and Hinson, 1988; Barker, Carey, and Taylor, 1990). Such differences were not the primary focus of this experiment. Consequently, prior to the search and select conditions, as described above, subjects performed a mouse familiarisation task which was designed to reduce the performance variance associated with this device, and to allow an independent assessment of the association between mouse performance and cognitive ability. A similar technique of familiarisation was used by Blankenberger and Hahn (1991) and in the command generation experiment reported in the previous chapter. The screen layout and trial procedure was essentially the same as used in the search and select conditions. Each trial began with the presentation of a cross, centrally positioned on the VDU. After an offset of one second, six boxes were displayed in a circle surrounding a central positioning dot. One of these boxes contained a cross which subjects were required to select. Performance of this task was also followed by completion of the NASA TLX.

6.3 Results

Related t-tests were used to compare mouse performance to the mean search and select condition performance. There was a significant difference for both response times ($t = -11.68$: $df = 47$: $p<.001$) and errors ($t = -8.27$: $df = 47$: $p<.001$). Mean scores are presented in Tables 6.01 and 6.02.

Table 6.01 : Mean response times (centisecs.) for mouse only and mean search and select performance

	Mean	SD
Mouse	84.14	19.88
Search and select	111.60	16.01

Table 6.02 : Mean proportion of errors for mouse only and mean search and select performance

	Mean	SD
Mouse	.001	.004
Search and select	.019	.016

6.3.1 Target and field conditions

The effects of target and field types were analysed using 2 x 2 (target type x field type) ANOVAs for response times, proportion of errors and self-report workload. Error rates were consistently extremely low, as can be seen from Table 6.04, and no significant main or interactive effects were found. Consequently, this dependent measure was not included in any analyses relating to individual differences.

The effect of target condition was not significant for either response time or for self-report workload (see Tables 6.03 and 6.05). However, there was a significant effect of field type for both response time ($F(1,47)=91.58 : p<.001$) and workload ($F(1,47)=21.67 : p<.001$) with subjects performing more quickly and reporting lower levels of workload when searching an icon field. There was also a significant 'target type' x 'field type' interaction for both response times ($F(1,47)=8.75 : p<.01$) and workload ($F(1,47)=10.04 : p<.01$) with mean scores for each measure following the same pattern. Subjects responded most quickly and reported least workload in the icon-icon condition. However, when an icon target was paired with a text field then speed of performance was at its slowest and self-report workload was at its highest. Pairing a text target with an icon field, however, resulted in performance speed and workload which was only slightly greater than the icon-icon pairing. It would appear that the field type is the more critical variable, and that the process of translation involved when an icon stimulus is presented followed by a text field increases the load beyond that which occurs for the text-text pairing.

Table 6.03 : Means and standard deviations for response times (millisecs) in each experimental condition

	Mean	SD
Mouse	841	199
Text-Text	1201	191
Icon-Icon	989	174
Text-Icon	1040	160
Icon-Text	1234	222

Table 6.04 : Means and standard deviations for the proportion of errors in each experimental condition

	Mean	SD
Mouse	0.00	0.00
Text-Text	0.02	0.03
Icon-Icon	0.02	0.02
Text-Icon	0.01	0.02
Icon-Text	0.02	0.02

Table 6.05 : Means and standard deviations for self-report workload in each experimental condition		
	Mean	SD
Mouse	28.49	14.18
Text-Text	37.67	15.11
Icon-Icon	33.10	17.15
Text-Icon	35.72	13.98
Icon-Text	41.49	16.80

6.3.2 Cognitive ability and search performance

Table 6.06 presents the mean scores for the measures of spatial and verbal ability. The correlation between these measures was not significant ($r = .10$).

Table 6.06 : Means and standard deviations for cognitive ability measures.		
	Mean	SD
Spatial visualisation	11.35	4.18
Nelson-Denny vocab.	64.44	16.84

In order to examine the relationship between cognitive ability and search performance, regression analyses were calculated in which vectors were coded to represent the 'target type' and 'field type' factors. These were entered into the regression equation with each cognitive ability measure, and additional vectors to account for interactive effects (Pedhazur, 1982). Separate regression analyses were performed for both spatial ability and verbal ability. Whilst neither spatial nor verbal ability was significantly related to performance in the 'mouse' condition (see Table 6.07), this variance was controlled, and therefore the main effects of cognitive ability relate only to the cognitive process of search and not to the physical selection components of the task.

Table 6.07 : Correlations between spatial and verbal ability and performance in the 'mouse only' condition		
	Response time	Workload
Spatial visualisation	-.23	-.03
Nelson-Denny vocab.	-.01	.12

Fig. 6.03 : Regression of response times on spatial visualisation for each experimental condition

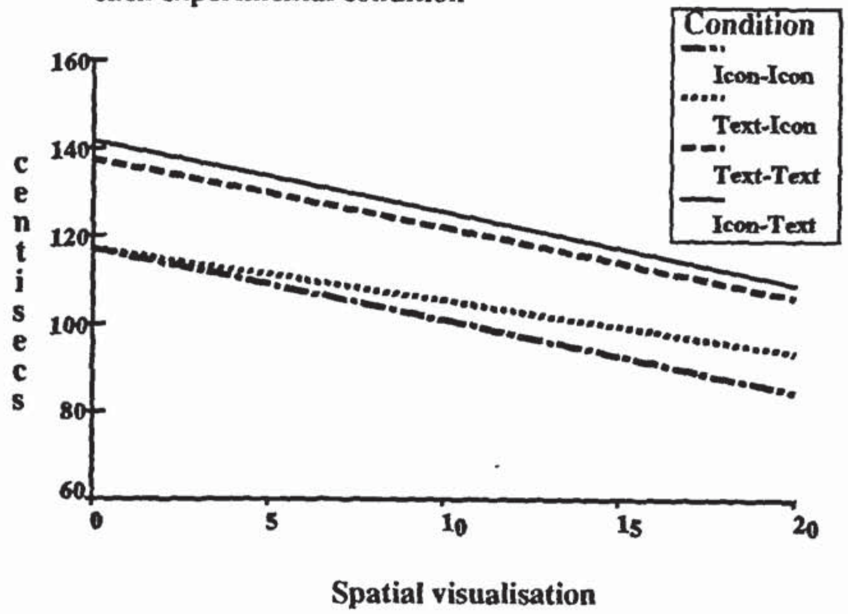


Fig. 6.04 : Regression of response times on vocabulary scores for each experimental condition

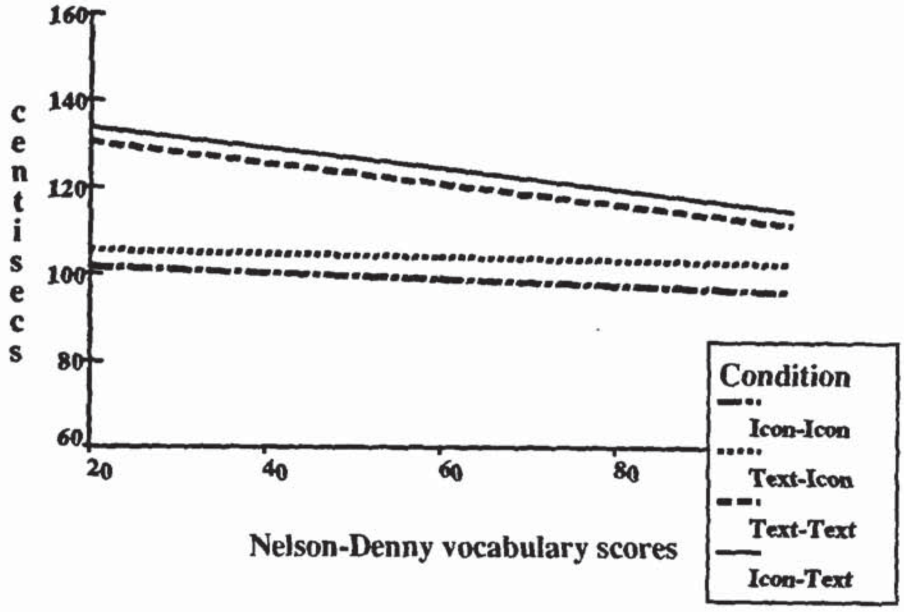


Fig. 6.05 : Regression of workload upon spatial visualisation for each experimental condition

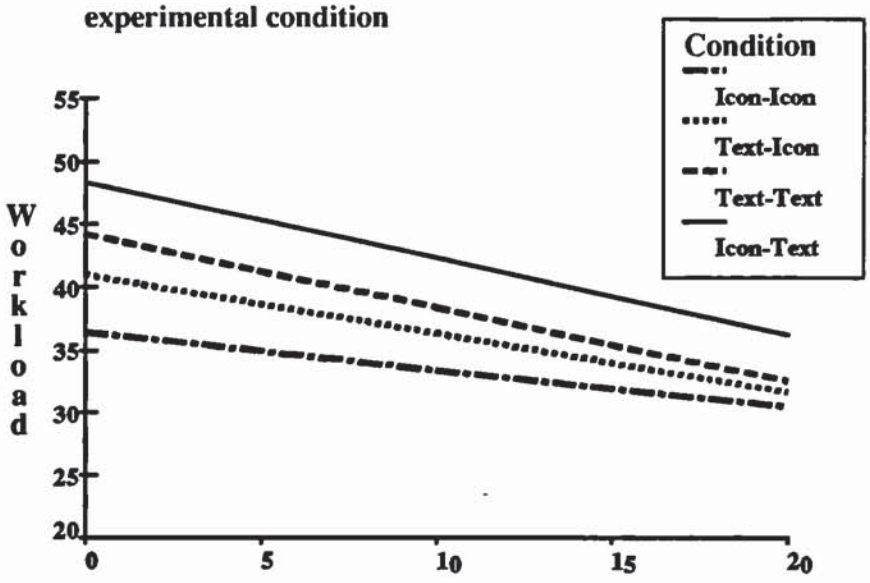
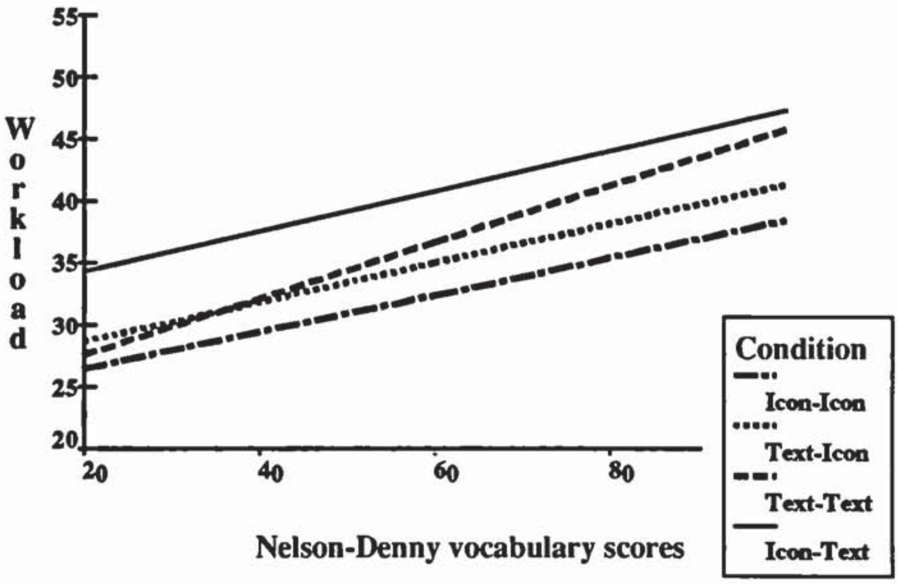


Fig. 6.06 : Regression of workload upon vocabulary scores for each experimental condition



The results of these analyses are shown in Figures 6.03 to 6.06. With respect to response times, there was a significant main effect of spatial visualisation

$(F(1,47)=5.74 : p<.05)$, however, the interactions with target type and field type were not significant. There were no main or interactive effects of verbal ability upon response time performance. There were also no main or interactive effects of spatial or verbal ability upon self-report workload. This lack of interactive effect in all conditions is illustrated by the similarity of the regression slopes for each cognitive ability and each experimental condition. It is worth noting, however, that whilst there is no significant relationship between spatial or verbal ability and workload, there are opposing trends for each ability measure, with workload decreasing with increasing spatial ability, but increasing for verbal ability.

6.4 Discussion

As anticipated, mouse performance was substantially easier than performance in any of the search and select conditions. This is consistent with the increased cognitive demands associated with the matching process. The correlation between spatial ability and mouse performance was not significant but was in the predicted direction and broadly supports previous findings (Taylor and Hinson, 1988; Barker et al., 1990).

Consistent with previous research, which has compared search and select performance for icons and text (Lansdale et al., 1989; Blankenberger and Hahn, 1991; Scott and Findlay, 1991), there was an overall response time advantage for conditions which required the search of an icon field. This was also supported by self-report workload scores which indicated that subjects were experiencing less attentional resource demand in these conditions. This did not appear to be the result of a trade off against performance accuracy, as the number of errors was very low and fairly uniform across all conditions. It would appear, therefore, that in situations of high icon dissimilarity there is less cognitive demand associated with searching an icon field than a text field. There was also a significant target type x field type interaction which was evident for both response times and self-report workload. This interaction indicates an increased performance cost incurred when the target and field types were not matched. This pattern supports the existence of dual processing codes associated with working memory (Baddeley, 1986; Wickens, 1984b), and may also indicate that subjects tend mentally to represent information in the presented format until required to make a comparison.

The hypothesised interaction between individual differences in performance and mode of presentation were not apparent. Whilst there was a main effect of spatial ability which could not be attributed to mouse performance, there were no significant interactive effects between cognitive ability and the search and select conditions. As evidenced by the similarity of the regression slopes for each cognitive ability with respect to each dependent measure neither spatial nor verbal ability appeared to be sensitive to manipulations of target type or field type. As text stimuli were displayed in the same case as the text field it is possible that subjects were engaging in a process of physical identity matching. This may have resulted in high spatial ability subjects using a spatial strategy in order to process text targets. However, this explanation could not account for the fact that a similar pattern of results was found when subjects were matching an icon target to a text field. In this condition physical identity matching was not possible and yet no differential association with spatial ability was apparent.

Whilst there was no significant relationship between spatial or verbal ability and the level of self-report workload, it is interesting to note that there were opposing trends in the nature of these correlations. High workload was associated with low spatial ability but high verbal ability. Generally it would appear that the NASA TLX was sensitive to experimental manipulations of search and select condition. However, there was no evidence to indicate that the lack of predicted interactive effects between cognitive ability and presentation mode could be attributed to systematic variance in workload.

On the basis of this evidence it would appear that individual differences in spatial and verbal ability cannot account for the lack of a consistent pattern with respect to the advantages of using icons within the interface, as described above. However, there are a number of related areas of investigation which might produce different results. In the present experiment very simple icons were used. It may be that an important dimension of individual differences relates to the complexity of icons. As mentioned earlier, spatial ability has been shown to covary with stimulus similarity in a shape comparison task (Dupree and Wickens, 1982). Possibly if more complex, or more similar icons had been used the effects of spatial ability would have been more pronounced. This would be consistent with the findings of Loo (1978), as previously described, and the interactive effects of FI-FD and embedded vs non-embedded street traffic signs. As mentioned earlier, the use of icons is thought to provide certain benefits with respect to the recall of information (Lansdale, 1988). This may also be a source of individual differences in performance. Mumaw and Pellegrino (1984), in an experiment mentioned above, found that individual differences may relate to the efficiency with which subjects are able to mentally represent information. It may be that, whilst the use of a search and select paradigm produces very little in the way of an interaction between cognitive ability and presentation mode, tests of recall would produce very different results. A further possibility would be to examine individual differences in relation to manipulations of the articulatory distance of icons (Blankenberger and Hahn, 1991). This might be achieved via a category matching task from which retrieval of complex verbal and non-verbal information from long-term memory could be assessed (Paivio, 1991).

Chapter 7

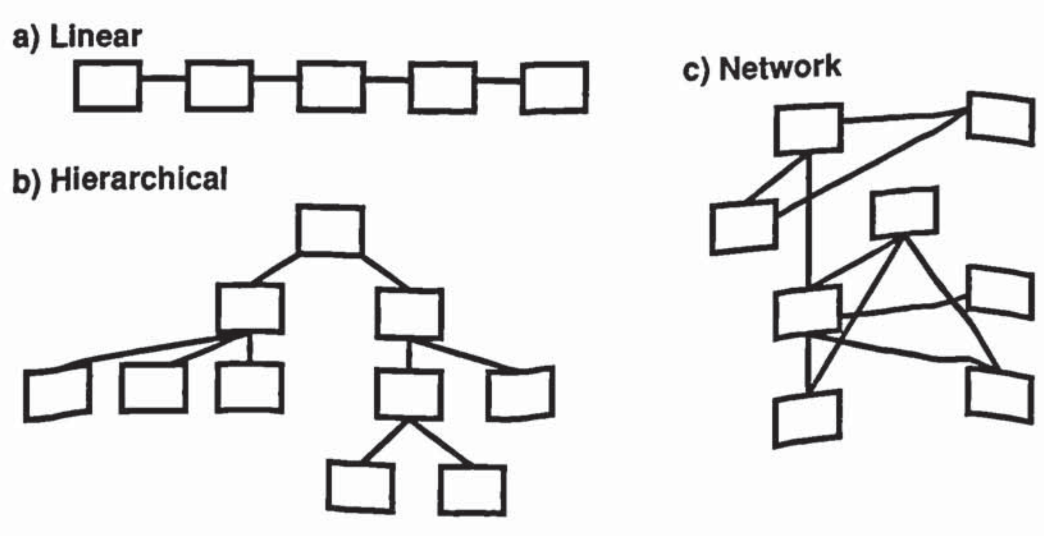
Information Retrieval

7.1 Introduction

This chapter presents an empirical investigation of individual differences in the process of information retrieval from a textual database. In particular, the interactive effects of database structure and node selection methods are examined in relation to age and cognitive ability. This introductory section discusses the existing evidence relating to each of these factors.

As the range of database applications increases (Kerr, 1990; Nielsen, 1992) and powerful hardware becomes more readily available, so the information retrieval user-base expands. Mediated searching (the use of information retrieval experts) is no longer the norm, and library and information point users are frequently left to their own devices. Consequently systems must be able to cater for a wider range of user abilities and requirements. New methods of linking and displaying information (Conklin, 1987; Halasz, 1988; Smith and Weiss, 1988) are associated with these developments. To date comparatively little experimental attention has been given to the cognitive demands which arise from the various linking structures and interface styles which are being used. Still less attention has been given to the potential interaction between these factors and individual differences in cognitive ability. This experiment is designed to explore some of these issues in relation to both data structure and interface design.

Fig. 7.01: Linear, hierarchical, and network data structures



7.1.1 Database structure

Computer technology makes possible a number of linking structures for the presentation and storage of textual material (Horton, 1990), the most predominant of which are the linear, hierarchical, and network structures (see Figure 7.01).

The linear structure is familiar to most people as the method of display generally used for paper-based information. It allows sequential access of information in a manner which is navigationally undemanding. Linear index systems mean that all options are immediately available to the user, allowing freedom of selection. However, there are also potential disadvantages associated with linear structures. The user may be presented with large quantities of information which are irrelevant to their current area of interest. Allied to this, it may be necessary to process a larger proportion of the contents of the database in order to locate a given target. This may not be a problem if the contents of the search are well defined, the database is small, or the cognitive processing costs per item are low. However, if there is a degree of uncertainty as to the target information, perhaps leading to an exhaustive search if the database is large, or if the cognitive or computing processing costs per item are great, then this system of information presentation may become inefficient. Many of these difficulties can be overcome by a process of classification (Engel, Andriessen, and Schmitz, 1983) and the attendant benefit of 'insulation' of information (Paap and Roske-Hofstrand, 1988) which allows the user to be guided away from material irrelevant to their current requirements. A hierarchical database structure enables such a pattern of interaction. Based upon the user's selection from a small number of broad categories, a limited section of the database is presented for examination and further selection. The number of alternatives presented in each menu panel and the number of selection processes which must be completed in order to arrive at the required information can be viewed as a depth vs. breadth trade off. There are several studies which have examined menu organisation in this respect. Increased depth can lead to a process of 'funnelling' in which the database user does not locate all the information which is relevant to their current area of inquiry because of the difficult in moving 'horizontally' within the structure. Lee and McGreggor (1985) present a model of menu search in which between four and eight alternatives per menu panel is optimal. However, the nature of the search (exhaustive vs self-terminating: MacGreggor, Lee, and Lam, 1986), the type of target matching (McDonald, Stone, and Liebelt, 1983; Norman and Chin, 1988), the menu content (Liebelt, McDonald, Stone, and Karat, 1982; Giroux and Belleau, 1986; Schwartz and Norman, 1986) and the domain knowledge of the user (Somberg and Picardi, 1983) are all important factors in determining the most effective structure. Various schemes have been proposed to determine the optimal structure (Shurtleff, Jenkins, and Sams, 1988; Fisher, Yungkurth, and Moss, 1990; Ukelson and Makowsky, 1992). However, there is convincing evidence to suggest that, if the database is not too large, a categorical organisation of menu items in a single index (i.e. maximum breadth) is preferable (Snowberry, Parkinson, and Sisson, 1983).

A network structure of linking allows for a greater flexibility in the design of a database (Smith and Weiss, 1988). Relational links can be established which may facilitate 'browsing', allowing the user to pursue 'content-oriented information-seeking strategies' (Marchionini and Shneiderman, 1988, p. 71). Greater efficiency of movement within the database is arguably possible as the user is not required to move up and down hierarchical chains in order to locate related nodes (Conklin, 1987). However, it has been suggested that network systems may be less efficient for fact retrieval (Marchionini and Shneiderman, 1988) and that the user may be more prone to becoming lost or disoriented within the database (Nielsen, 1990). This may be attributed to the greater structural complexity associated with an increased number of links, a reduction in the spatial regularity of the database structure, or both. Such an effect can be seen from the importance attached to the use of 'landmarks' or maps as a means of navigational reference within network structures (Valdez, Chigwell, and Glenn, 1988; Frisse and Cousins, 1992).

An important consideration is the cognitive structure which will most readily be adopted by the user when processing the contents of a database. Various theories have been proposed to describe the structure imposed by human memory upon stored information (e.g. Collins and Quillian, 1969, 1970). Soderston (1986) suggested that experimental evidence supporting a performance superiority when using a hierarchical as opposed to a linear structure was due to a greater compatibility with the structure of human memory. An experiment conducted by Durding, Becker, and Gould (1977) indicates that optimal database structure may vary depending upon the nature of the information. This experiment investigated the basis upon which individuals select and utilise data structures. Four data structures were considered: (i) lists (index), (ii) hierarchical, (iii) network, and (iv) table. Durding et al. (1977) found that the semantic content of the presented information was central to the efficiency with which subjects used data structures to represent that information. Hollands and Merikle (1989) have further extended these findings by demonstrating the interactive effects of task type and domain expertise. In this experiment domain experts were initially used to create a hierarchical data structure. Additional hierarchies were also created using alphabetic and random item allocations. The performance of users of varying domain expertise was then compared upon term matching (subjects required to locate a match for a target term) and definition matching (subjects required to locate a term to match a target definition) tasks. Results indicated an interaction between menu organisation, expertise and task type, with domain novices performing better with alphabetical organisation when term matching was required, but categorical organisation proving to be superior regardless of expertise when definition matching was required. Some of the issues surrounding the cognitive compatibility which exists between an individual's model of a database and the information content are discussed by Allen (1990) who also examines the possibility of using individually tailored advance knowledge organisers which would assist the individual to develop an appropriate cognitive structure of the information to be searched.

An empirical comparison of linear and hierarchical information structures was conducted by Gordon, Gustavel, Moore, and Hankey (1988) who assessed linear and hypertext (a small hierarchy two / three levels deep with embedded menus selected using cursor keys) presentations of documents in two conditions of reading comprehension. In one condition subjects were presented with general interest articles and instructed to read 'as if for interest', and in the other condition they were presented with technical material and instructed to study for a short test. Linear presentation of the material was found to be superior in the general interest condition, in which subjects reported greater cognitive effort associated with the use of the hypertext format. However, there was little difference between the two interface conditions when subjects were studying technical material. Consistent with this latter result, Marchionini and Shneiderman (1988) briefly mention an unreferenced experiment conducted by Koved which examined the use of hierarchical or linear database structures for fact retrieval. No speed or error differences were found between the two systems.

An examination of the benefits of network structures is provided by Gray and Shasha (1989) who compared subject performance whilst using two versions of a textual database (n=50). In one version the contents could only be accessed by reference to specific properties of an individual file, whilst in the other version semantic links were created between files, forming a network which could also be used for navigation. Subjects were required to use a query language to locate the answers to five questions. Gray and Shasha (1989) found that subjects in the condition without additional links performed better than those in the network condition. They report that some subjects in the network condition found that the use of the additional links resulted in navigational problems. However, it should be noted that the interface to this database is far more complex than those typically found in hypertext applications.

McKnight, Dillon, and Richardson (1989) conducted an experiment which was more representative in this respect. Fact retrieval performance was compared for two network structures and two linear structures (one was paper). Subjects were more accurate and spent less time using the index in the linear conditions. An advantage for linear over network structures is also reported by Monk, Walsh, and Dix (1988). Subjects were presented with computer program code using linear or network structures and required to locate the answers to program comprehension questions. Performance was quicker when using the linear structure. However, the provision of a map in a second experiment resulted in similar performance speed for both linear and network conditions. Mohageg (1992) examined linear, hierarchical, network, and combined network and hierarchical structures in relation to the access of a geographical database. The linear structure did not take the form of an index and required subjects to page from one topic to the next in sequence. Subjects located the answers to questions which had been categorised as 'hierarchical' or 'relational', on the basis of database programmers' ratings as to the suitability of these structures for the location of each answer. In the first experiment the navigational distance to target information was the same for the hierarchical, network and combined structures. Mohageg found task completion times were significantly shorter for the hierarchical and combination structures than for the network or linear structures. Whilst the network task produced greater path uncertainty than the hierarchical or combined structures (Frick and Miller, 1951; Miller and Frick, 1949), the greatest deviation from the optimal path was found for the hierarchical structure. This difference primarily resulted from questions which had been designated 'relational'. In the second experiment only 'relational' tasks were used and additional network links were established such that the number of links optimally required to locate target information was half that required in the hierarchical condition. This was designed to represent a greater facility for information access 'with no regard for levels' which Mohageg (1992, p. 357) suggests is one of the prime advantages of relational structures. Task completion times were found to be significantly shorter for the network structure than for the hierarchical or linear structures. The only significant difference in navigational efficiency was a greater deviation in the combination condition when compared to the network condition.

7.1.1.1 Summary

The experimental evidence comparing linear, hierarchical, and network linking structures is complex. It would appear that the nature of the task (Hollands and Merikle, 1987; Gordon et al., 1988), the organisation of information (Snowberry et al., 1983), the information content, and the cognitive structure applied by the user

(Durding et al. 1977; Soderston, 1986) are all factors which influence the optimal database structure. On balance, it would appear that, when navigational distance is equated, the weight of empirical evidence favours hierarchical or linear structures as opposed to network structures. However, as Mohageg (1992) points out, this does not take account of the potential reductions in navigational distance which are possible when using network structures. The advantage found by Mohageg (1992) for hierarchical over linear structures may be due to the absence of a linear index system in this case.

Of particular interest to the present discussion is the spatial complexity and semantic content associated with each structure. Within this context the linear structure can be seen to be the least spatially complex, whilst the network structure has the greatest spatial complexity. As indicated by the experiment of Durding et al. (1983) the semantic content associated with each of these structures will be dependent upon the information content of the database. However, as discussed above, the inclusion of relational links in the network structure can be predicted to increase semantic content, albeit at the expense of spatial complexity (Nielsen, 1990).

7.1.2 Menu design

The use of explicit or embedded menus is an interface design issue which is closely related to many of the structural considerations presented above (Koved and Shneiderman, 1986). An explicit menu provides the user with a series of contextually divorced alternatives from which a selection can be made. An embedded menu uses words and objects within the database as a means of selection. Embedded menus are thought to result in the availability of increased contextual information, highlighting semantic rather than spatial relationships within the database (Marchionini and Shneiderman, 1988). This approach to menu selection is in keeping with the increased interest in direct manipulation interfaces (DMIs: Peverett, 1992) and object oriented approaches to program design (Morrill-Tazelar, 1990). Users are able to directly manipulate the object of concern, in this case the menu item. Marchionini and Shneiderman (1988, p. 78) suggest that this may improve navigational performance because information is provided in "...meaningful task domain (as opposed to computer domain) terms and concepts, thereby reducing disorientation".

This prediction is supported by an experiment reported by Shneiderman (1987c) in which explicit and embedded menus were used in a repeated measures examination of database access. Subjects answered significantly more questions, were navigationally more efficient, and preferred using embedded menus. Marchionini and Shneiderman

(1988) briefly discuss a number of experiments relating to the use of embedded menus in 'Hyperties', a hypertext database. In one such experiment 14 out of 16 subjects performing a fact retrieval task were found to use an index as opposed to embedded menus. However, Marchionini and Shneiderman pointed out that subjects were initially more familiar with this method of retrieval. In a further similar study, subjects were found to perform more quickly when using the index, however this performance advantage diminished over time. An examination of a log of use for another Hyperties database, which was installed in a museum, revealed that embedded menus were the preferred method of use. The importance of contextual information is consistent with results reported by Wright and Lickorish (1989) who found that under conditions of increased cognitive load, a performance disadvantage was associated with the use of an index navigation system presented separately from the database text contents, as opposed to a 'with text' navigation system.

7.1.2.1 Summary

The evidence with respect to the use of explicit or embedded menus suggests that there may be navigational benefits associated with embedded menus due to increased semantic information content, but that they are also subject to response time costs, with subjects taking longer to locate and select menu items. However, it would appear that task demands may be an important factor in determining user preference and performance with either of these methods.

The comparative advantage of hierarchical over network structures which was apparent in the first of Mohageg's (1992) experiments, discussed above, may have been due to differences in the types of menu used. Whilst it is not completely clear from the descriptions given, it would appear that an explicit menu was used for the hierarchical condition and an embedded menu was used for the network condition. It may be that the speed advantage for the hierarchical condition was not due to a reduction in the cognitive demands associated with navigating this structure, but to differences in the speed with which subjects were using explicit and embedded menus.

7.1.3 Individual differences

The increased cognitive overhead incurred by non-linear systems (Conklin, 1987), and the increasing use of such systems by infrequent and casual users who are in the early stages of skill acquisition, suggests that the variance in performance accounted for by individual differences in cognitive ability will also increase (Ackerman, 1988).

Certainly there are grounds to believe that substantial individual differences in information retrieval exist (Egan, 1988; Nielsen, 1989). However, to date, very little empirical attention has been paid to the potential interactive effects between individual differences in cognitive ability or age and database structure or menu style (explicit vs embedded). In Chapter 4 the evidence relating spatial ability to the process of navigation was considered. Some of the literature most pertinent to the current discussion will now be reviewed in greater detail.

Spatial ability has consistently been found to be the strongest predictor of information retrieval performance. Campagnoni and Erlich (1989) examined the relationship between spatial visualisation (VZ2; Ekstrom, French, and Harman, 1976) and information retrieval performance whilst using a 'hypertext-based help system'. Subjects (n=12) located the answers to six queries within a hierarchical database. Half of the questions were thought to be most suited to a browsing information retrieval strategy, whilst the other half were thought to be suited to what was termed an analytical (use of index) strategy. All subjects located all answers. Results indicated that, regardless of question type, most users preferred the browsing strategy and would often only use the index if this strategy failed. It should be noted, however, that there was very little difference between the two strategies in the number of steps required. There was a significant (r = -.75) correlation between spatial visualisation and solution time, and between spatial visualisation and the number of times subjects returned to the top level of the database (r = -.65). Whilst all subjects were unfamiliar with the system used, they varied widely in their previous computer experience. Expertise, as measured on a five point Likert scale (beginner to expert) was not significantly associated with performance. However, whilst non-significant, there was a positive correlation between expertise and spatial visualisation (r = .47) with high spatial ability being associated with greater computer experience.

Spatial ability was also found to be a strong predictor of performance in a study conducted by Vicente, Hayes and Williges (1987). A variety of predictors were 'assayed' in relation to the search of information relating to the armed forces, which was organised in a three level hierarchy consisting of 15 files. The predictors variables included flexibility of closure , perceptual speed, spatial orientation, spatial scanning, spatial visualisation, spatial memory, vocabulary, reading rate, comprehension, sex, computer experience, computer courses taken, abstractness, field dependency, anxiety and information processing rate (derived from a choice reaction time task). The dependent measures used were 'time taken', 'total number of commands' and 'number of different commands used'. Six variables were significantly correlated with at least one dependent measure and these all related to either spatial ability or verbal ability. A stepwise multiple regression revealed two variables which accounted for 45% of the variance in time taken, these being the vocabulary section of the Nelson-Denny Reading Test and a test of spatial visualisation (VZ2 - Ekstrom et al., 1976), with the latter being the stronger predictor. In a further analysis, subjects were divided about the median to form low and high spatial ability groups. There was a significant difference in performance, with low spatial ability subjects taking, on average, twice as long to locate targets. Using a command polling procedure devised by Elkerton and Williges (1985), Vicente et al. (1987) found that low spatial subjects showed a different pattern of command usage. Low spatial ability subjects used the SCROLL commands more frequently, as well as the ZOOM OUT command which allowed subjects to re-orient themselves within the file structure. In a subsequent experimental report Vicente and Williges (1988) hypothesised that this pattern of command use was due to a greater need for visual momentum (Woods, 1984: cf. Chapter 3) on the part of low spatial ability subjects, and that the provision of additional navigational information would improve the performance of this group. In order to test this hypothesis a 'graphical' interface was developed in which a map of the database and information as to current file position were provided. Subjects were selected to two groups on the basis of extreme spatial visualisation scores. It was hypothesised that, due to a ceiling effect upon the performance of high spatial subjects, low spatials would show the greatest benefit from the adapted interface. There was a main effect of spatial visualisation in the predicted direction, and performance was better in the 'graphical' interface condition than the control condition. However, the predicted interaction with spatial ability did not materialise, with both groups benefiting equally. The variance in scores for the 'graphical' interface was found to be reduced. This effect was significant for the high spatial group, and approached significance for the low spatial group. With respect to command usage, the SCROLL DOWN command was used significantly less frequently in the graphical interface condition, and there was a non significant trend in the predicted direction with respect to the ZOOM command. There was a corresponding increase in the use of the SEARCH-AND command.

The association between spatial ability and the provision of navigational information was also examined by Billingsley (1982). Subjects were required to search a menu hierarchy in one of three conditions. In the first of these subjects were provided with no navigational assistance. In the second, subjects were provided with a map of the hierarchy, whilst in the third condition subjects were provided with a list of links which could be followed to targets. There was a tendency for performance to be best in the map condition. However, spatial memory (MV2: Ekstrom et al, 1976) was reported to be consistently predictive across all conditions, for both response times and navigational efficiency. Associative memory (MA3: Ekstrom et al., 1976) was also significantly related to response times, although the correlation was weaker than that for spatial memory. This pattern of results can be seen to be consistent with that of Vicente and Williges (1988) in that the provision of spatial information was of equal assistance to high and low spatial ability subjects.

Whilst the experiments outlined above have examined the association between spatial ability and the provision of additional navigational information, Seagull and Walker (1992) considered the impact of variations in the structure of a database hierarchy. Four conditions were used, each of which contained 64 nodes. One condition used a linear structure with all 64 nodes accessed by means of a single menu (this had to be spread over two screens, and therefore required paging). Another condition used a two level hierarchy with an 8 x 8 format. A further condition used a three-level hierarchy with a 4 x 4 x 4 format. The final condition used a mixed two and three level hierarchy with six alternatives presented on the highest menu. Subjects were experienced in the use of computers (more than 1 year). Spatial visualisation (VZ2; Ekstrom et al., 1976), flexibility of closure (CF2; Ekstrom, et al., 1976), perceptual speed (P2; Ekstrom et al., 1976), spatial scanning (SS2; Ekstrom et al., 1976) and verbal comprehension (V4; Ekstrom et al., 1976) were examined as predictors of performance. Spatial visualisation was the strongest and most consistent predictor of response time, although correlations were modest ($r < .4$). However, there was no interactive effect of spatial ability and structure. Verbal ability was not significantly correlated with performance times. Seagull and Walker also examined the strength of correlation for the first and the second halves of the performance trials separately. Correlations with spatial visualisation were slightly stronger for all interface conditions in the second block of trials. On the basis of Ackerman's (1988) theory of individual differences in skill acquisition, Seagull and Walker suggest that this reflects the influence of individual differences in processing speed as opposed to individual differences in "the ability to remember or learn the organisation structure of the information" (p. 378). They draw support for this hypothesis from the lack of a significant correlation between spatial ability and navigational efficiency. However, this conclusion does not seem warranted as it is based upon the assumption that subjects are able to immediately achieve an efficient mental representation of the hierarchy, whereas it seems more plausible that this will be developed over time. Seagull and Walker's experimental design would not facilitate such development as subjects performed only four trials in sequence with each hierarchical structure. In addition, the lack of significant correlation between spatial ability and navigational efficiency may be attributable to relatively error free performance. It may be that a more complex task would produce different results.

These experiments suggest that there is no strong association between spatial ability and the provision of navigational information or changes to the underlying spatial structure of the database. Further experimental evidence to this effect was also presented in Chapter 4. However, an interaction between spatial ability and experimental manipulations of interface variables was apparent in results reported by Jennings, Benyon, and Murray (1991), who required subjects to locate product information in a shopping catalogue. Correct retrieval of information required subjects to specify a product and three attributes. Five different interfaces were used: (i) 'button', in which users "select an item type by clicking the mouse button on named boxes representing narrowing choices of categories of item types, and clicking on their required values for the attributes of the item type" (p. 246), (ii) 'command', in which users must type the product name and attributes using the correct syntax, (iii) 'iconic', which is similar to the button interface except that icons are used to represent the products, (iv) 'menu', in which subjects were presented with a menu of category choices, and pull-down menus for selecting product attributes, and (v) 'question', in which subjects type responses to a series of prompts. Spatial ability, verbal ability, field independence, short-term memory, and the Myers-Briggs Type Inventory were used as predictors. The only significant correlations with performance were obtained between spatial ability and performance using the 'command' and 'question' interfaces, and between verbal ability and performance using the 'question' interface. Jennings et al. (1991) suggest that this pattern of results may be attributable to additional navigational demands imposed by these conditions. However, the reasoning behind this hypothesis is less than clear. Whilst little detail is provided, Jenkins et al. (1991) indicate that additional mode changes are required when using the 'command' interface. However, the 'question' interface is apparently similar to the 'button' and 'iconic' interfaces in this respect. As an alternative hypothesis Jenkins et al. (1991) propose that spatial ability may be related to "a user's ability to cope with interfaces allowing a very open and flexible dialogue". Whilst it could be argued that the 'question' interface is highly structured (cf. Macaulay and Norman, 1984), this position is more consistent with results presented in Chapter 4, which suggest that the semantic content of the interface is the important factor in determining the strength of the association between spatial ability and performance. It may be that the 'command' and 'iconic' interfaces present a more abstract model to the user.

7.1.3.1 Summary

Spatial ability has consistently been found to be the strongest predictor of information retrieval performance. The results of Vicente et al. (1987), Vicente and Williges

(1988) and Campagnoni and Erlich (1989) suggest that the association between spatial ability and performance may be due to a navigational advantage, with high spatial subjects navigating the database more efficiently than low spatial subjects. However, this was not supported by the findings of Seagull and Walker (1992) who propose that this association may be due to differences in processing speed. An interaction between spatial ability and interface type was apparent in the study of Jenkins et al. (1991). Whilst the relative navigational demands of each interface were not entirely clear, it would seem that variation in navigational demand cannot explain this interactive effect. In this respect this study is consistent with those previously mentioned. An alternative explanation was proposed which concerns the semantic distance (cf. Hutchins, Hollan, and Norman, 1986) associated with the interface. As was demonstrated in Chapter 4, the association between spatial ability and performance appears to strongest when the semantic content of the interface is low.

The evidence relating verbal ability to the process of information retrieval is mixed. Vicente, Hayes and Williges (1987) found a significant association with performance, Seagull and Walker (1992) did not, and Jennings et al. (1991) found a significant correlation only when the 'question' interface was used.

7.1.4 Experimental aims

One of the aims of this experiment was to examine the association between spatial ability and the semantic and spatial information content of the interface within the context of a complete information retrieval setting. As mentioned above, manipulations of the spatial complexity and semantic content of the interface can be achieved through variations in the linking structure applied to data, and the use of explicit or embedded menus. Upon the basis of the evidence presented above, it was predicted that the effects of spatial ability would be inversely related to the semantic content of the interface, whilst spatial complexity would have relatively little influence. Previous studies of individual differences have been limited to hierarchical or linear (index) structures. An additional aim of this experiment was therefore to establish the nature of the association between information retrieval performance and spatial ability when using a network database structure.

In order to examine these issues subjects were required to locate the answers to 18 questions about tropical fish within a database. This topic was chosen in order to avoid the problems of domain experience and age or sex bias. All subjects were naive to this subject area. Four interface conditions were used: (i) linear structure (explicit menu), (ii) hierarchical structure (explicit menu), (iii) hierarchical structure (embedded menu), and (iv) network structure (embedded menu). This selection of conditions allowed a comparison of linear and hierarchical structure with explicit menus, a comparison of hierarchical structures with differing menu systems, and a comparison of hierarchical and network menu structures with the same menu system. Prior to the performance of this task subjects also completed two menu selection component tasks which enabled an assessment of the relationship between cognitive ability and the process of explicit and embedded menu selection which was independent of the more complex processing associated with information retrieval.

Upon the basis of the evidence presented above, the importance of verbal ability as a predictor of information retrieval is unclear. An further aim of this experiment was to clarify this position and to examine possible interactive factors, such as linking structure and menu type, which might account for these conflicting previous results. In addition to tests of spatial and verbal ability, tests of logical reasoning and associative memory were also included. A number of studies have found tests of logical reasoning to be predictive of computer-based performance (Egan and Gomez, 1985; Green, Gomez, and Devlin, 1986; Czaja, Joyce, and Hammond, 1989), and the inclusion of this test provided an indication of the importance of spatial ability in relation to the more general contribution of fluid intelligence. Similarly, a test of associative memory was included in order that the effects of spatial memory could be interpreted in terms of its association with second-order spatial and general memory factors.

A deficiency within current research knowledge in this area concerns the effects of age upon information retrieval performance. Whilst age has been shown to be strongly predictive of word processing (Egan and Gomez, 1985; Czaja, Joyce, and Hammond, 1989) and the use spreadsheets (Gist, Rosen, and Schwoerer, 1988) this apparently important variable has not been investigated with respect to the performance of textual information retrieval. A related study by Greene, Gomez and Devlin (1986; Greene, Devlin, Cannata, and Gomez, 1990) found age to be predictive of the use of a database Structured Query Language (SQL). However, this study was not computer-based. A thorough review of the relevant literature has revealed no information relating to age differences in the navigation of data structures. There are a number of age-related changes in cognitive function which can be predicted to interact with the spatial and semantic content of the interface during the process of information retrieval (as discussed in Chapter 4). Spatial ability has consistently been found to decline with age independently of practice or experience (Salthouse and Mitchell, 1990). Conversely many aspects of verbal ability and processing related to semantic content have been found to remain relatively unimpaired, or even improve, well into old age (Mitchell and Perlmutter, 1986; Bolla, Lindgren, Bonnacorsy, and Bleeker, 1990; Davies, Taylor, and Dorn, 1992). This pattern of cognitive ageing suggests that the inclusion of additional semantic information within the interface may be comparatively more advantageous to older users, whilst an emphasis upon the spatial content of the interface will have the reverse effect. However, consideration must also be given to the results of studies which have demonstrated that increases in task complexity, and the requirement for the deployment of additional processing resources may disproportionately disadvantage the older individual (Cerella, Poon, and Williams, 1980; Horn and Hofer, 1992; Salthouse, 1992). It may be that any performance advantage for older subjects associated with the inclusion of additional semantic links will be offset by the additional complexity of the network structure.

7.2 Method

7.2.1 Subjects

48 subjects were recruited from local job clubs and training agencies. Of these 24 were aged between 18 and 25 (mean = 21.79 years : SD = 2.75) and 24 were aged 45 and over (mean = 50.79 years : SD = 4.30). In each age group equal numbers of males and females were tested. Subjects were computer novices with less than 20 hours experience of interactive computing.

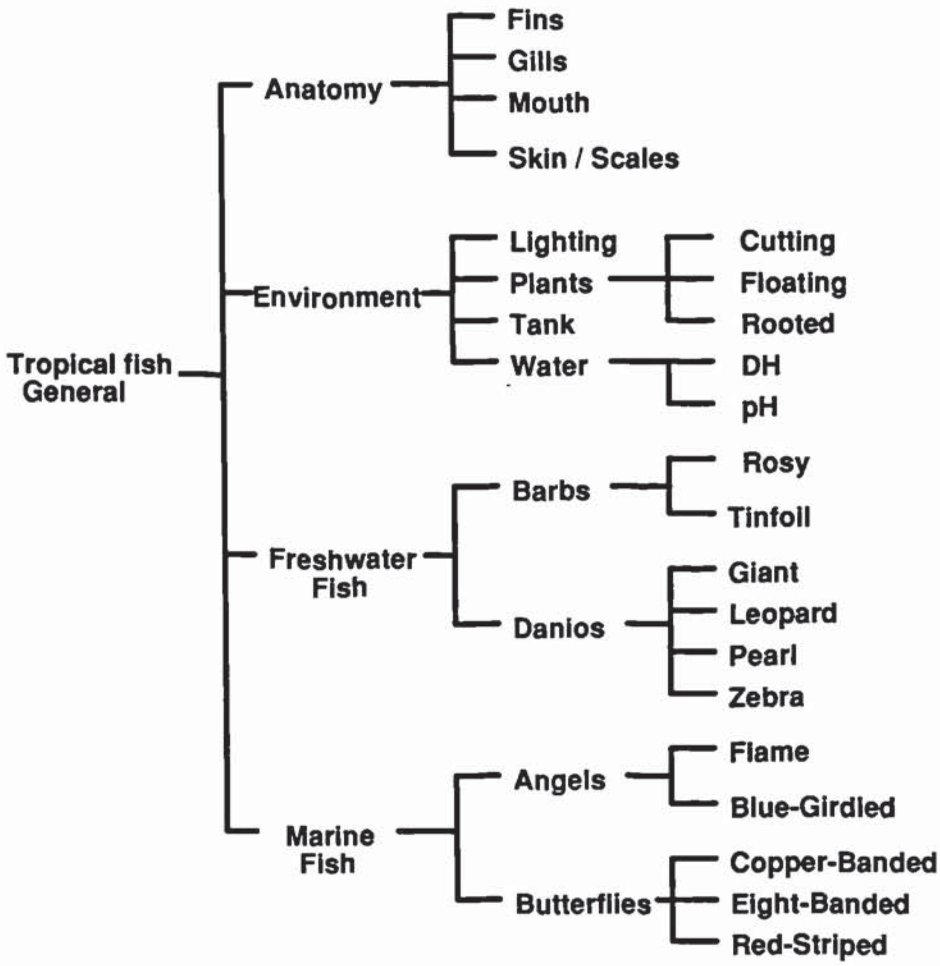
7.2.2 Measures of cognitive ability

Subjects initially completed tests of spatial visualisation (VZ2: Ekstrom et al., 1976), spatial memory (MV2: Ekstrom, et al., 1976), and verbal ability (Nelson-Denny Vocabulary Test, 1973), logical reasoning (LR2: Ekstrom, et al., 1976) and associative memory (MA2: Ekstrom, et al., 1976).

7.2.3 Experimental task overview

The information retrieval tasks were conducted on PC compatible computers fitted with 80286 processors and Hercules graphics cards. The database and component tasks were specially programmed for the experiment in 'C' and Assembly language. These programs used text mode (80 x 25 characters) and logged subject performance data with a timing accuracy of approx. ± 5 msecs. The database contained 800 lines of textual information which were divided into 33 separate files (nodes), the longest of which was 44 lines and the shortest of which was six lines. This information was included for compilation within the program code in order to avoid timing delays associated with disk access.

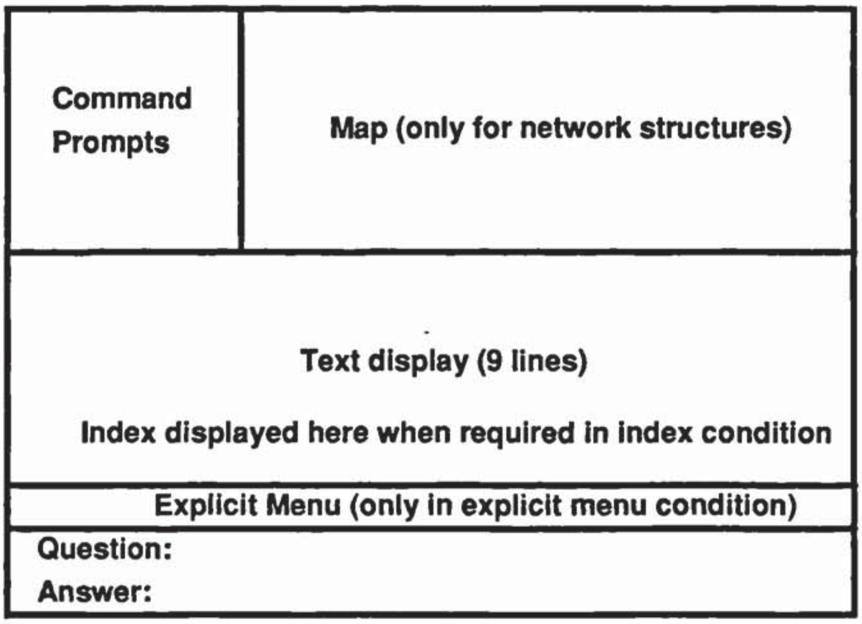
Fig. 7.02 : Tropical Fish Database: Hierarchical File Structure



Four experimental interface conditions were used in which menu characteristics and underlying file structure were manipulated. Experimental conditions included linear structure, hierarchical structure with explicit menu, hierarchical structure with embedded menu, and network structure with embedded menu. Each of these conditions is described in more detail below. Figure 7.02 shows the files contained within the database linked in the hierarchical structure. A between subjects design was used, therefore six subjects in the younger age group and six in the older group were tested in each of these interface conditions with an equal number of males and females allocated to each.

The screen layout for the database is illustrated in Figure 7.03. Whilst this basic layout was the same for all conditions, certain elements (e.g. the map or the explicit menu) were only displayed in the appropriate experimental conditions.

Fig. 7.03 : Screen Layout



7.2.4 Menu selection component tasks

Subjects were initially required to perform two component tasks which related to the processes of explicit and embedded menu selection. These tasks were presented in counterbalanced sequence.

The explicit menu component task presented subjects with a four item horizontal menu. Three of these items were "OOOO" and were used as distractors, and the fourth was "XXXX" and was the target. Subjects were required to use the left and right cursor keys to position a highlighted bar over the target and were then required to press the ENTER key. At this point the screen cleared for a 1 second interval prior to the next menu presentation. Subjects completed four practice trials followed by sixteen trials for which speed and accuracy of performance were recorded. Targets were randomly positioned within blocks of four trials such that every position was used.

The embedded menu component task presented subjects with a passage of text upon the computer screen which contained four highlighted 'keywords', these being represented by three sets of "OOOO"s which were distractors and one set of "XXXX"s which was the target. Subjects were required to used the left, right, up, and down cursor keys to position the cursor anywhere on the target and then press the ENTER key. Subjects completed four practice trials followed by sixteen trials for which the speed and accuracy of performance was recorded. 20 different passages of text were used. The position of the target in relation to the distractors was randomly allocated over blocks of four trials such that each position was used. Texts were allocated to either practice or performance sections and then presented in random sequence. Each piece of text was 25 lines long, displayed in a window of 9 lines. If the target was not immediately displayed upon the screen subjects were required to scroll down until it came into view.

NASA self-report workload ratings were collected at the end of each component task.

5.2.5 Information retrieval task conditions

In all interface conditions six blocks of three questions were presented to subjects, the first of which was used as a demonstration / practice block and was excluded from all analyses. In order to answer a question, having located the required information, subjects pressed the 'A' key. A prompt to this effect was placed in the 'command prompts' section of the screen (see Figure 7.3). This resulted in the cursor moving from the currently selected text to the bottom line of the screen next to the 'Answer' prompt, where the subjects could then type in their response. Typing time was not included in response times. If the question was the first or second in the block of three, then the current node position was maintained, and once the subject pressed the ENTER key the cursor was returned to the currently selected text. Following the third response for each block the current position was returned to the original starting position ('Tropical fish - General'). This sequence of questions was designed to allow performance to be examined in a situation which approximated general information retrieval performance over a prolonged query without imposing artificial landmarks. The five performance blocks were presented in random sequence. The number of links required in each condition for each block was controlled so that for the hierarchical structure conditions, 12 links was the shortest route which could be taken in order to correctly answer each of the three questions. In the index condition only three links were required per block. In the complex hypertext condition additional links were provided (as described below) and blocks required between 6 and 10 links to be navigated in order to locate the necessary information.

7.2.5.1 Linear structure

In this condition subjects were presented with an index of all files within the database. Given the importance of categorisation within indexes (Snowberry, Parkinson and Sisson, 1983) these files were hierarchically grouped such that files of increasing specialisation were indented to the right and grouped under files covering more general information about the same topic. Subjects used the UP and DOWN cursor keys to navigate the index, and the ENTER key to select a file. Having selected a file the first 9 lines of text were displayed, with the UP and DOWN cursor keys now being available to scroll the contents of the file upon the screen. Subjects could return to the index at any time by pressing the "I" key. A prompt to this effect was placed in the 'command prompts' section of the screen (see Figure 7.3).

7.2.5.2 Hierarchical structure with explicit menus

In this condition subjects were presented with a hierarchical file structure as shown in Figure 7.2. A menu of available nodes was displayed in the area shown in Figure 7.3. The LEFT and RIGHT cursor keys moved a highlighted bar along the menu, and the ENTER key selected a file for examination. The first 9 lines of text of the appropriate file were displayed on the screen. The UP and DOWN cursor keys were used to scroll the contents of the file on the screen. Subjects could use the 'Backup' command by pressing the 'B' key in order to move back up the menu hierarchy. A prompt to this effect was placed in the 'command prompts' section of the screen (see Figure 7.3).

7.2.5.3 Hierarchical structure with embedded menus

In this condition the same hierarchical file structure was used as in the previous condition, however, the explicit menu was not available. Subjects were required to navigate between files by placing the cursor upon highlighted key words within the text files and pressing the ENTER key. This resulted in the first 9 lines of the relevant piece of text being displayed on the screen. Each keyword occurred only once in each file. The 'Backup' command was again available to subjects and allowed them to retrace their steps, one step at a time. A prompt to this effect was placed in the 'command prompts' section of the screen. A map showing the current node and all nodes immediately connected was displayed in the section of the screen shown in fig 7.3. This ensured that the prompts available to subjects with regard to node availability were the same for explicit and embedded menus. The provision of a map has also been shown to be an important element in hypertext interface usability (Monk, Walsh, and Dix, 1988; Simpson, 1990).

7.2.5.4 Network structure with embedded menus

In this condition embedded menus were used, as in the previous condition, however, the pattern of links was more complex. All of the hierarchical links were maintained, and additional links were created on the basis of semantic content.

7.2.6 Dependent measures

Dependent measures were response time, excluding typing time (milliseconds), accuracy (number of questions correct), and a measure of navigational efficiency which was derived from the ratio of the number of links used and the optimal number of links required. Response time and navigational efficiency data were only used for those performance blocks in which all three questions had been correctly answered. This ensured that the optimal navigational distance was the same for all subjects.

7.3 Results

7.3.1 Outlying or missing data

Three subjects were excluded from all data analyses because of outlying or missing data. Subject 31 had made a response error in each block of trials and consequently no meaningful response time data could be obtained. Similarly, subject 29 had made a response error in all but one block of trials, and reaction time data was extraordinarily slow throughout all trials (well in excess of three standard deviations from the mean). A further subject (09) was found to be an extreme outlier with respect to navigational performance. Subjects 09 and 29 had both performed the retrieval task using the explicit menu interface, and during the experimental debriefing reported some confusion as to whether the currently highlighted item on the menu referred to the current text or to the next available option. This confusion may have given rise to these extreme outlying scores. Full details of each of these subjects are given in Table 7.1.

Table 7.01 : Details of excluded subjects								
Subject No.	Interface condition	Age	Sex	SV	SM	ND	LR	MA
09	Explicit menu	21	M	-.75	5.25	23	5.50	6.00
29	Explicit menu	49	M	3.00	4.25	39	2.75	8.00
31	Network	59	M	-2.75	6.00	67	5.50	7.00

If the scores relating to the cognitive predictors presented in Table 7.1 are compared to the mean scores for the rest of the sample as presented in Table 7.2 it is interesting to note that all of these subjects had extremely low spatial visualisation scores. The exclusion of these subjects did not dramatically alter the results of any statistical test with respect to interface condition. However, where individual differences were being considered the magnitude of effects was generally reduced by their exclusion, i.e. results are more conservative than they might otherwise have been.

Table 7.02 : Sample mean, standard deviation and range for cognitive predictors.				
	MEAN	SD	MIN	MAX
SV	7.96	4.16	-.75	16.25
SM	10.96	5.62	3.00	22.75
ND	55.98	21.71	15	98
MA	9.58	6.23	1.00	28
LR	8.80	5.93	-1.00	25

Two further subjects (one from each age group) performed particularly inaccurately in the explicit menu task component condition. Whilst the average level of accuracy was 95%, one subject missed all targets and the other hit only 19% of targets. As this was an apparently simple task it can only be assumed that these subjects misunderstood what was required. Their scores were excluded from analyses involving this component task only.

7.3.2 Menu component tasks

Performance accuracy was not examined for these tasks because of the extremely small number of errors. Performance accuracy was 100% for the embedded menu component task.

Unrelated t-tests were used to examine age differences in response time and self-report workload. There were no significant differences for either task with respect to self-report workload. However, there were significant age differences for response time in both the explicit ($t=-5.31$; $df = 41$; $p<.001$) and embedded ($t=5.09$; $df = 43$; $p<.001$) menu component tasks with the older age group performing more slowly than the younger age group. Means and standard deviations are shown in Table 7.3.

Table 7.03 : Means and standard deviations for response times (secs.) for the explicit and embedded menu component tasks.

	Explicit		Embedded	
	Mean	SD	Mean	SD
YOUNG	1.377	0.378	9.730	1.609
OLD	2.328	0.732	14.815	4.411

Correlations between cognitive predictors and performance upon the menu task were calculated, and are shown in Table 7.4. There were significant, although modest correlations between spatial visualisation and response times for each of the component tasks. Logical reasoning was significantly correlated with performance on the explicit menu task but not the embedded menu task. However, spatial memory was the strongest predictor for both tasks.

Table 7.04 : Correlations between cognitive predictors and response times for the explicit and embedded menu component tasks.

	Explicit menu (n=43)	Embedded menu (n=45)
SV	-.27 *	-.26 *
SM	-.42 **	-.42 **
ND	.15	-.18
MA	-.20	-.24
LR	-.32 *	-.18

ns = non-significant : * = $p<.05$: ** = $p<.01$: *** = $p<.001$

In order to examine age differences in the relationship between cognitive ability and performance upon the component tasks multiple regression analyses were used in which age was entered as a dummy variable and cognitive ability was entered as a continuous variable (Pedhazur, 1982). A third vector representing the interactive term (cognitive ability x age) was the only variable of interest.

Fig. 7.04 : Embedded menu component task : Regression of response times on spatial visualisation for younger and older age groups

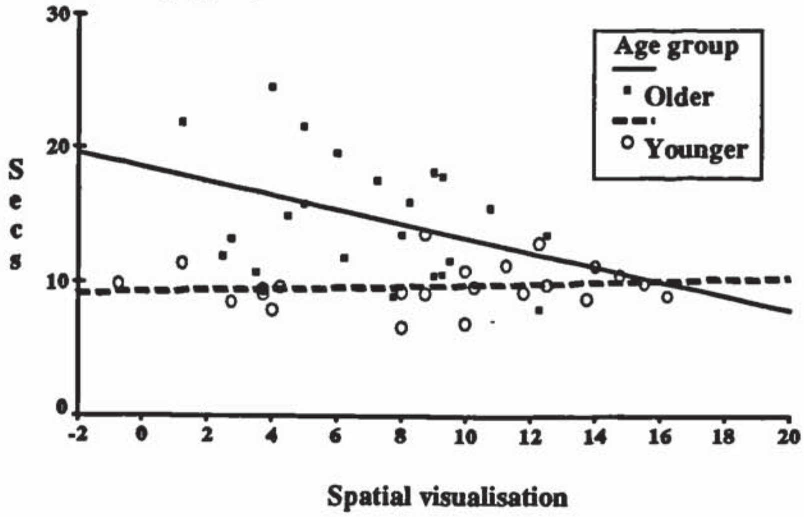


Fig. 7.05 : Embedded menu component task : Regression of response times on spatial memory for younger and older age groups

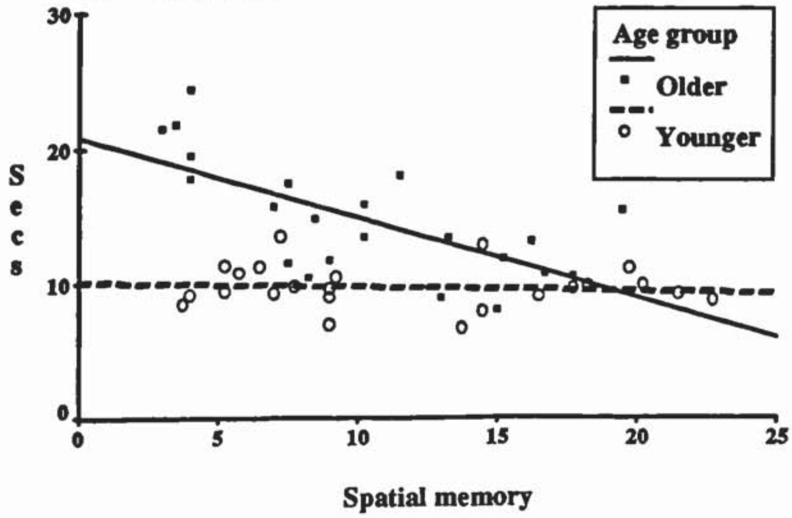
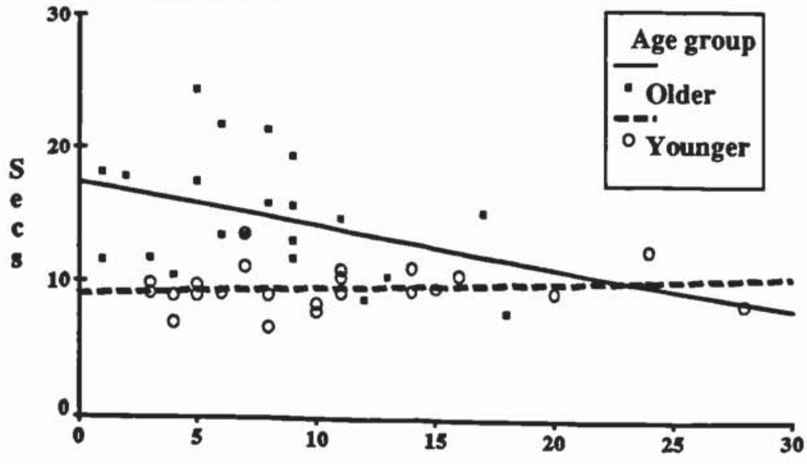


Fig. 7.06 : Embedded menu component task : Regression of response times on associative memory for younger and older age groups



There were no significant differences in the regression slopes for younger and older subject groups for the explicit menu component task, although the interaction with spatial memory approached significance ($F(1,39)=3.99$). With respect to the embedded menu component task there was a significant interaction between age group and spatial visualisation ($F(1,41)=4.98 : p<.05$), spatial memory ($F(1,41)=15.43 : p<.001$), and associative memory ($F(1,41)=5.34 : p<.05$). These interactions can be seen in Figures 7.04 to 7.06. As can be seen from the regression lines, there is a consistent pattern across each of these effects such that cognitive ability is more strongly predictive of performance for the older subject group.

7.3.3 Information retrieval task: Statistical approach

The distribution of accuracy scores was restricted in range and negatively skewed. A Kolmogorov-Smirnov test indicated that the distribution was significantly different from a normal distribution ($p<.05$). None of the usual methods of data transformation improved this situation. Unfortunately, the data analysis required parametric tests, and whilst the particular tests used are relatively robust with respect to non-normality of sample distribution, this must be noted and results pertaining to accuracy of performance treated with a degree of caution.

Data for each block were only included for analysis if all three questions were answered correctly. Consequently, because 'block' was a within subjects factor, the data for any subject who had incorrectly answered one or more questions would be completely excluded from the analysis, thereby reducing the cell numbers unacceptably. In order to overcome this problem, the interactive effects of individual differences and performance block were analysed separately from those involving interface effects, and cell means were substituted for missing data.

A 2 x 4 ANOVA was used to examine the effects of age (2 levels) and interface condition (4 levels). Newman Keuls tests were used in order to assess the difference between interface condition means. Whilst this test is less conservative than a Sheffé test it was selected for reasons of comparability of results with those of Mohageg. The effects of age and performance block (5 blocks) were examined using a 2 x 5 ANOVA.

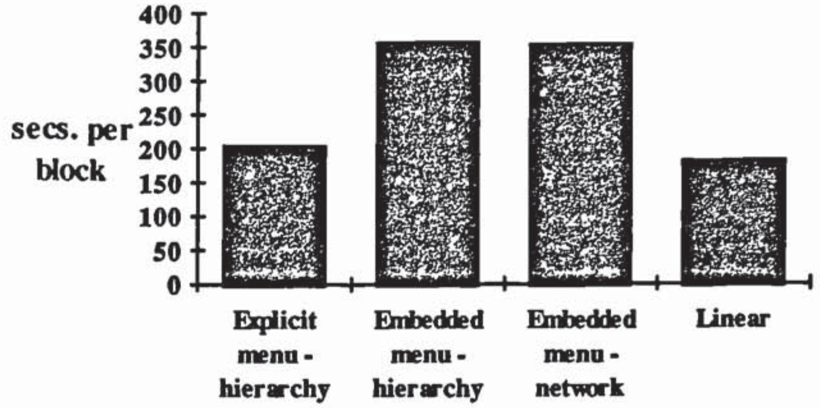
Separate analyses of the relationship between cognitive ability and performance were conducted using a series of multiple regression equations for each dependent measure with each cognitive predictor. Interface condition (4 levels) was coded as a factorial variable whilst each of the measures of cognitive ability were treated as continuous variables. This method of analysis avoids the loss of information which occurs when continuous variables are categorised (Pedhazur, 1982). A further series of multiple regression analyses was performed to examine possible age differences in the relationship between cognitive ability and performance. Age was entered as a dummy variable in the regression equation, whilst cognitive ability measures were again entered as continuous variables. The latter analysis examined overall performance (mean for all interface conditions) as a full factorial design would involve regression slopes being plotted on the basis of too few observations.

7.3.4 The effects of interface condition and age

7.3.4.1 Response Times

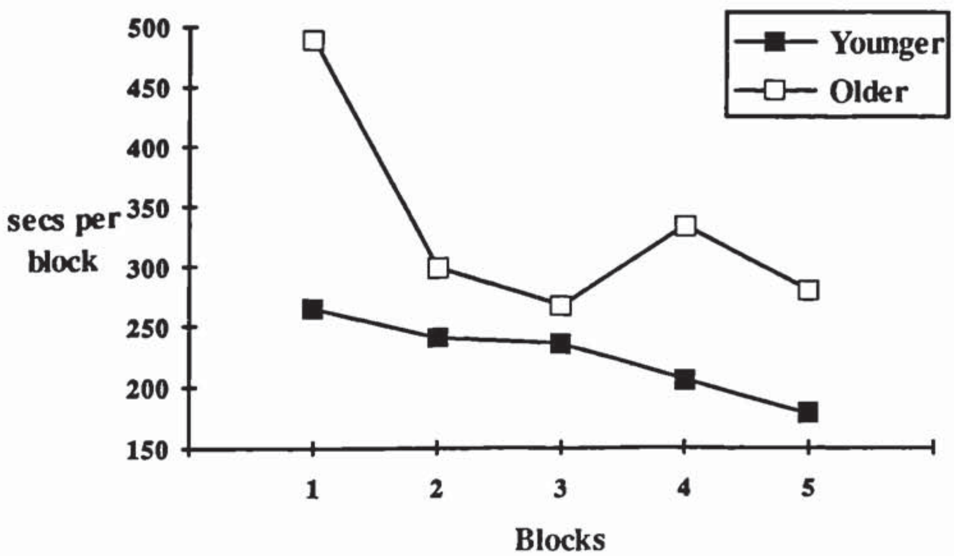
There was a significant main effect of interface condition ($F(3,37)=11.16 : p<.001$) with subjects performing most quickly in the linear condition, and most slowly in the 'embedded menu - hierarchical structure' condition (see Figure 7.07). A Newman-Keuls test (alpha = .05) revealed that the mean response times for the embedded menu conditions were significantly different from those for the explicit menu conditions. There was also a significant main effect of age ($F(1,37)=14.76 : p<.001$) with the younger age group performing more quickly than the older age group. However, there was no significant interaction between age and interface condition.

Fig. 7.07 : Response times for each interface condition



There was a significant main effect of performance block such that performance improved over time ($F(4,172)=10.83 : p<.001$). There was also a significant interaction between age group and performance block ($F(4,172)=4.62 : p=.001$) such that the older subject group performed much more slowly in the first block, and despite an initially steeper acquisition slope, there was a persistent performance advantage for the younger age group (see Figure 7.08).

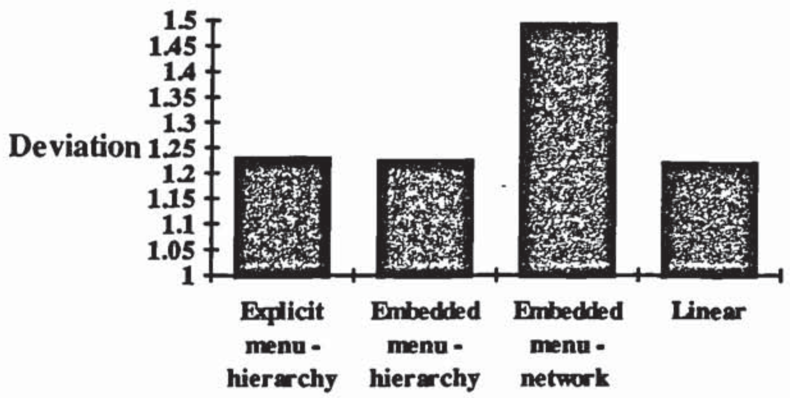
Fig. 7.08 : Response times for younger and older subjects for each block of questions



7.3.4.2 Navigational efficiency

There was a main effect of interface condition ($F(3,37)=3.19 : p<.05$) with performance in the linear, explicit menu - hierarchy, and embedded menu - hierarchy conditions being very similar. However navigational efficiency in the embedded menu - network condition was poorer (see Figure 7.09). A Newman-Keuls post hoc test revealed that the embedded menu - network condition was significantly different from the other three interface conditions. The main effect of age and the interaction between age and interface condition were non significant.

Fig. 7.09 : Deviation from optimal path for each interface condition

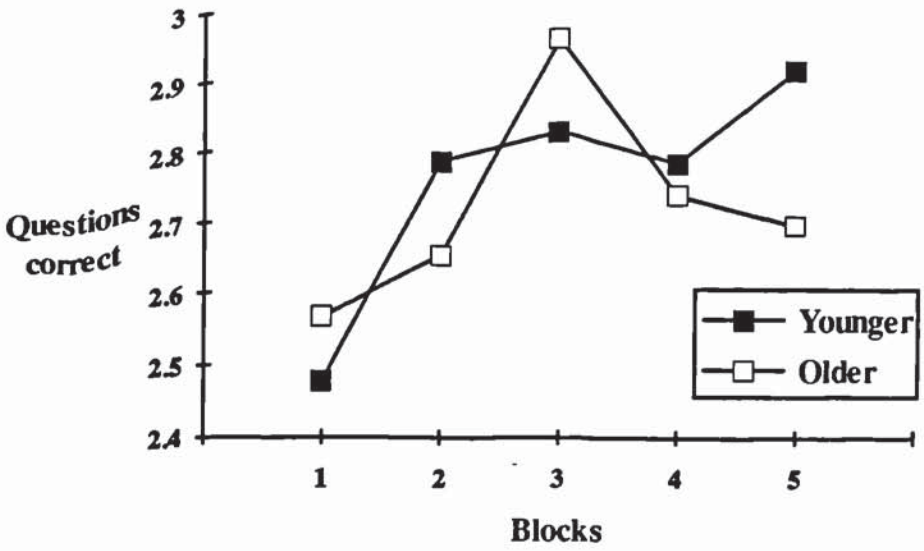


There was a significant main effect of performance block (F(4,172)=4.28 : p<.005) with navigational efficiency generally improving over time. However there was no significant interactive effect of age group with block.

7.3.4.3 Accuracy

Accuracy was analysed using a 2 x 4 x 5 analysis of variance (age x interface condition x block). There were no significant main or interactive effects of age or interface condition with respect to accuracy. A Newman-Keuls test indicated that no two interface conditions were significantly different. There was a significant main effect of performance block (F(4,148)=3.13 : p<.05) which was broadly consistent with performance accuracy increasing over time (see Figure 7.10). Examination of Figures. 7.08 and 7.10 reveal no obvious speed accuracy trade-off over time for either age group, although performance accuracy does peak in the third performance block for the older age group whilst performance speed is relatively stable.

Fig. 7.10 : Performance accuracy for younger and older subjects for each block of questions



7.3.5 The effects of interface upon acquisition

The effects of interface upon acquisition were examined using a 4 x 5 (interface x blocks) ANOVA for each dependent measure. The interactive effects of interface and block were non-significant.

7.3.6 Age differences related to network complexity

There was a non-significant trend such that the older subject group experienced greater comparative difficulty in the 'embedded menu - network structure' condition than in the 'embedded menu - hierarchical structure' condition. This pattern was consistent for response time, navigational efficiency and performance accuracy (see Figures 7.11 to 7.13). With respect to response times, whilst the performance of the older group deteriorated in the network condition the opposite was true for the younger group. Both age groups navigated comparatively more efficiently in the 'embedded menu - hierarchical structure' condition, although the older group were the more efficient of the two. In the network condition, however, this pattern was reversed. With respect to accuracy, both age groups were more accurate in the 'embedded menu - hierarchical structure' condition, however, the performance advantage for the older subject group in this condition was noticeably reduced in the network condition. Combined measures of performance efficiency were examined (combinations of speed and accuracy of performance, e.g. accuracy / square root rt), however this interaction still failed to reach significance. This may be due to the small cell numbers involved.

Fig. 7.11 : Response times for younger and older subjects in each interface condition

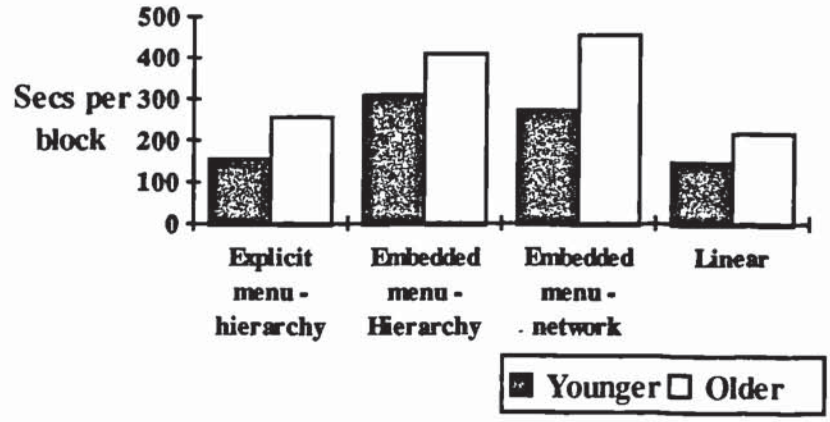


Fig. 7.12 : Deviation from optimal path for younger and older subjects for each interface condition

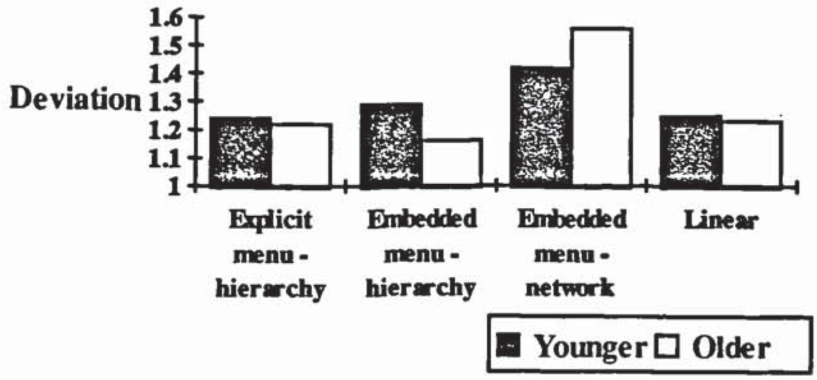
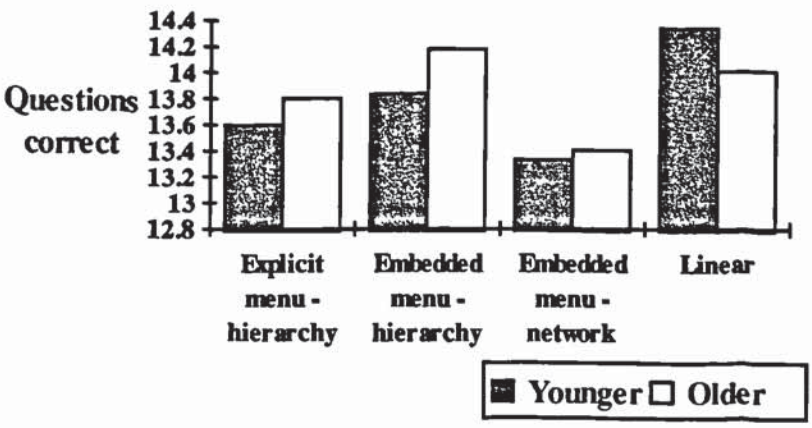


Fig. 7.13 : Performance accuracy for younger and older subjects in each interface condition



In order to identify more precisely the locus of age differences in response time, the importance of the variance associated with the process of node selection was considered in relation to the variance associated with interface condition. Multiple regression analyses were used in which information retrieval task response times was the dependent variable and either explicit or embedded menu component task response times were entered as the first predictor variable. Age was then entered into the equation to ascertain the variance associated with age which was not related to the basic node selection process. As can be seen from Tables 7.05 and 7.06 the variance in response times associated with age which cannot be predicted from performance upon the component tasks is extremely small.

Table 7.05 : Regression of retrieval task response time upon explicit menu component task response time, followed by age group

	r2 change	df	SS change	MS	F	p
Explicit RT	.32659	1	250945.48	250945.48	19.82	<.001
Age	.01416	1	10878.92	10878.92	<1	ns
Residual		40	506557.35	12663.93		

Table 7.06 : Regression of retrieval task response time upon embedded menu component task response time, followed by age group

	r2 change	df	SS change	MS	F	p
Embedded RT	.3998	1	319169.67	319169.67	27.98	<.001
Age	.0001	1	78.67	78.67	<1	ns
Residual		42	479033.47	11405.56		

7.3.7 Cognitive predictors of information retrieval

Table 7.07 shows the correlation matrix for these ability tests.

Table 7.07 : Correlation matrix for cognitive ability tests

	SM	ND	MA	LR
SV	.50 ***	.12	.17	.41 **
SM		.23	.41 **	.34 **
ND			.28 *	.27 *
MA				.18

ns = non-significant : * = p<.05 : ** = p<.01 : *** = p<.001

7.3.8 The relationship between cognitive ability and interface condition

As a precursor, separate regression equations were generated for each cognitive predictor which included the linear, quadratic, and cubic terms. No significant improvements in prediction were found when the quadratic or cubic terms were included for any of the cognitive predictors. Consequently, a linear model was used throughout.

Fig. 7.14 : Regression of performance accuracy on spatial visualization for each interface condition

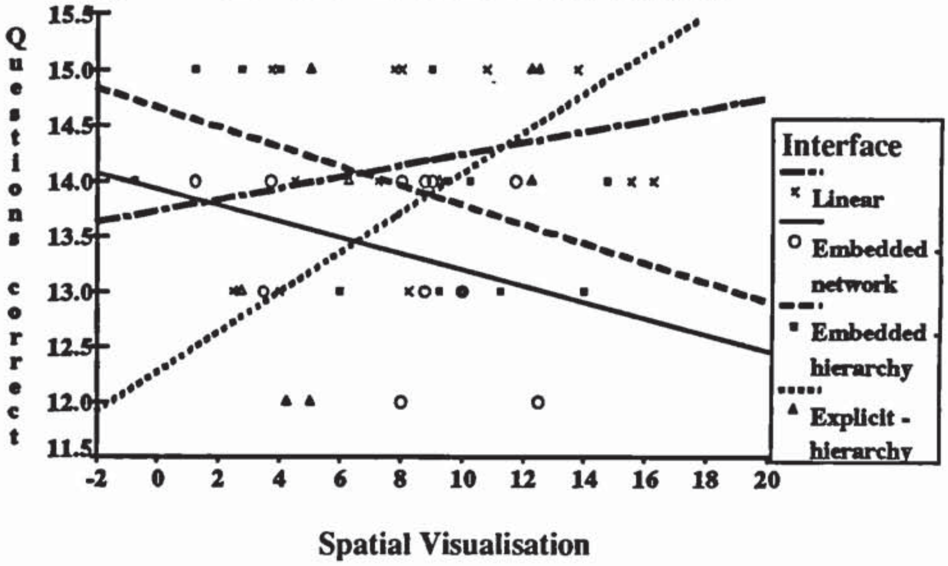


Fig. 7.15 : Regression of response times on spatial visualisation for each interface condition

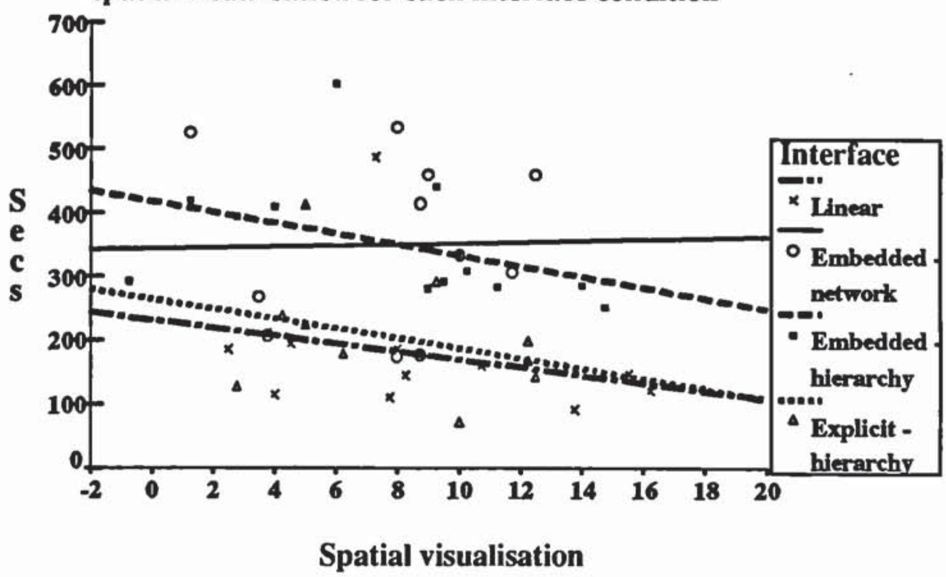
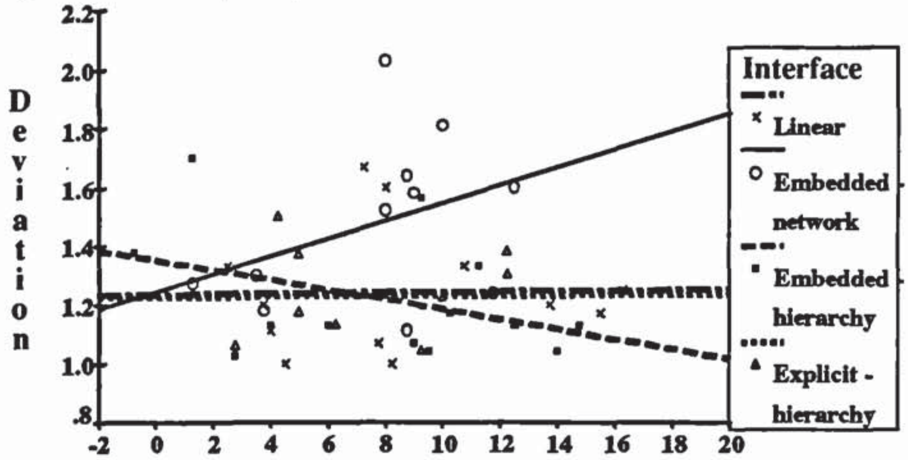


Fig. 7.16 : Regression of navigational efficiency (no. links taken / optimal no. links) on spatial visualisation for each interface condition



The main effect of spatial visualisation was not significant for any of the dependent measures. However, there was a significant interaction between spatial visualisation and interface condition with respect to accuracy of performance ($F(3,37)=3.35$: $p<.05$). The slopes for the regression of accuracy upon spatial visualisation for each of the interface conditions are shown in Figure 7.14. As can be seen, there is a positive relationship between spatial visualisation and accuracy for the 'explicit menu - hierarchical structure' and linear conditions, with high spatial subjects performing more accurately. However, for both of the embedded menu conditions this position is reversed, and high spatials perform less accurately. It should be remembered, however, that the distribution of accuracy scores is restricted and negatively skewed. These findings must consequently be treated with some caution. However, further support is given by the non-significant trend in this direction with respect to response times for the network structure (see Figure 7.15) in which there is a positive association between spatial visualisation and response times, i.e. high spatials perform slower in this experimental condition, and a similar trend for navigational efficiency (see Figure 7.16).

There was a main effect of spatial memory ($F(1,37)=6.74$: $p<.05$) with respect to response times in the predicted direction. Subjects with high spatial memory scores performed more quickly than those with low spatial memory scores. Whilst the main effect for logical reasoning and response time approached significance ($F(1,37)=4.02$), there were no other significant main or interactive effects for any of the other cognitive predictors.

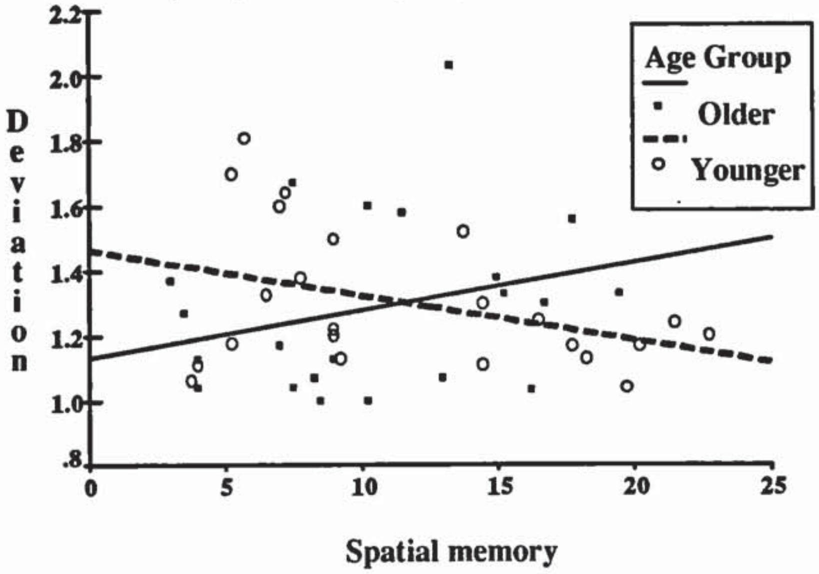
7.3.9 The relationship between cognitive ability, age, and information retrieval

Age differences in cognitive ability were examined using a series of unrelated t-tests. The only significant difference in scores was for the Nelson-Denny vocabulary test ($t=-3.66$; $df = 43$; $p=.001$) with older subjects performing better. For all other tests there was a non significant trend for the younger subjects to achieve higher scores (see Table 7.08).

Table 7.08 : Means and standard deviations for cognitive ability test scores for younger and older subject groups.

		Mean	SD
SV	Younger	8.90	4.83
	Older	6.98	3.15
SM	Younger	11.66	6.15
	Older	10.23	5.05
ND	Younger	45.74	16.38
	Older	66.68	21.71
LR	Younger	9.39	6.37
	Older	8.18	5.51
MA	Younger	10.61	6.64
	Older	8.50	5.72

Fig. 7.17 : Regression of navigational efficiency (no. of links taken / optimal no. links) on spatial memory for younger and older age groups



The only significant interaction between any cognitive predictor and age group was with spatial memory for navigational efficiency ($F(1,41)=4.74$: $p<.05$). The scatterplot of scores and group regression lines are shown in Fig. 7.17. As can be seen, there is positive association between spatial memory and navigational efficiency for the older age group, whilst the opposite is true for the younger age group.

7.4 Discussion

7.4.1 Database structure and interface design issues

Performance in both of the embedded menu conditions was found to be significantly slower than performance in the explicit menu (hierarchy and linear) conditions. This may be largely attributable to the method of node selection, in that embedded menus required greater cursor movement, including the scrolling of text. Whilst this is a feature of many hypertext applications, it places an additional demand upon the process of navigation. Ewing, Mehrabanzad, Sheck, Ostroff, and Shneiderman (1986) found that hypertext navigation using jump arrow keys, which moved the user to the next embedded keyword in the selected direction, was more effective and preferred when compared to the use of a mouse. It may be that such a system would improve the comparative performance of the embedded menu conditions. However, the associated discontinuity of movement which would occur, particularly if the next keyword was not displayed on the screen might disadvantage users of low cognitive ability (Woods, 1984).

The inclusion of additional links in the network condition, providing shorter optimal navigation routes, made very little difference to response times. The small difference between the network condition and the comparable hierarchical structure (embedded menu) was non-significant. This is contrary to the results of Mohageg (1992), who found that when additional network links were provided performance times improved. However, the comparison is not precise. In the present experiment additional links were built upon a hierarchical structure, whereas Mohageg's initial network was non-hierarchical. Performance times in Mohageg's network and hierarchical conditions, when the required navigational distance was the same, indicated a significant advantage for the hierarchical structure. Only when the required navigational distance for the network structure was halved, and the questions used were of a type thought to be particularly suited, did a performance time advantage materialise for this condition. In the current experiment navigational efficiency in the network condition was significantly worse than in any of the other three conditions. It would appear that subjects were not taking advantage of the reduced distance to targets. Whilst this is consistent with the results of Gray and Shasha (1989) it is less compatible with those of Mohageg. In Mohageg's study, when the required navigational distance was the same for network and hierarchical tasks, there was greater path uncertainty when using the network structure but navigational efficiency was higher. When the required navigational distance was reduced for the network structure there was no significant difference in navigational performance between the two conditions. It is possible that, in the current experiment, the decrease in navigational efficiency when using the network structure is attributable to an increased tendency to become disorientated in a more complex structure. This may be allied to an incompatibility with the cognitive structure employed by subjects to mentally represent the contents of the database (Durding, Becker, and Gould, 1977). The hierarchical structure may provide a better match between the system representation and the individual's mental representation of the data (Soderston, 1986). This area of investigation is worth pursuing. It may be that there are disadvantages associated with the provision of additional semantic links in networks structures.

There was very little difference in response time between the two explicit menu conditions, with the linear structure being slightly faster than the hierarchical structure in the predicted direction. This, however, is likely to be influenced by the size of the database. This experiment used a comparatively small database, and it is possible that, given a substantial increase in the number of nodes, this pattern of performance would change. Nevertheless, these findings are broadly consistent with previous results which, on balance, suggest that a categorised index will result in superior performance (Snowberry, Parkinson, and Sisson, 1983).

7.4.2 Cognitive ability and information retrieval

Whilst there was a significant main effect of spatial memory upon response time for the information retrieval task, consistent with the results of Billingsley (1982), the main effect of spatial visualisation failed to reach significance. This is contrary to the results of Vicente, Hayes, and Williges (1987) which indicated that spatial visualisation was the stronger predictor. This may be attributable to the interactive effects of spatial visualisation and interface condition. A weak overall association may be explained by differences between interface conditions, as was also apparent in the study by Jennings et al. (1991). There was a significant interaction between spatial visualisation and interface condition with respect to performance accuracy. This was such that there was a positive association between spatial ability and the number of questions correctly answered in both explicit menu conditions, but a negative association when an embedded menu was used (see Figure 7.14). The similarly sized correlations between spatial visualisation and each of the menu selection component tasks indicates that such differences are not an artefact of the mechanics of the particular menu selection method. These results are consistent with the experimental hypotheses. It would appear that the inclusion of increased semantic content within the interface, in the form of embedded menus, is disproportionately advantageous to low spatial ability subjects. Further support for this position can be found in the interactive trends which were apparent in the response time and navigational efficiency data. With respect to response times, high spatial ability was associated with faster performance in all conditions except the network condition (see Figure 7.15). As discussed above, semantic content was considered to be greatest in the network condition due to relational linking and the use of embedded menus. It may be therefore, that these interface elements serve to negate the association between spatial visualisation and performance. The similarity of the regression slopes for spatial visualisation with respect to the linear and 'explicit menu - hierarchical structure' conditions for both response time and navigational efficiency confirms the findings of Seagull and Walker (1992). It would appear that when explicit menus are used with a linear or hierarchical structure then spatial ability does not interact with the depth vs breadth trade-off.

An alternative explanation for the weak positive correlation between spatial visualisation and response time in the network condition relates to the spatial complexity associated with this structure. Whilst the hierarchical and linear structures have a high degree of spatial regularity, the organisation of node links in the network condition, although more spatially complex, is also less spatially regular. It is possible that such a spatial organisation is less compatible with the mental representation of the database contents used by high spatial subjects (Durding, Becker, and Gould, 1977). The performance advantage associated with high spatial ability may result from a capacity to organise and structure information in a regular manner which is less compatible with the network condition. Consequently, whilst the provision of additional spatial information within conditions of high spatial regularity (e.g. maps; Vicente and Williges, 1988) does little alter the performance balance between high and low spatial individuals, the use of incompatible data structures may radically alter the situation. A process which would help to explain such a relationship is 'rectilinear normalisation' (cf. Wickens, 1992) which refers to a tendency to normalise spatial information into a right-angled grid as part of the process of mental representation. Obviously hierarchical structures are easily represented in this manner, whereas the representation of network structures in this format may be extremely difficult. If such a process was related to spatial ability in this context, then the performance disadvantage of high spatials in the network condition might be explained. This is obviously highly speculative and further research would be required to assess the importance of structural regularity in this context.

A similar pattern of results was also apparent for navigational efficiency, with a tendency for high spatial subjects to be particularly inefficient in their navigation of the network structure (see Figure 7.16). The pattern of regression lines suggest that the overall reduction in navigational efficiency which was found for the network condition is largely due to a the relatively poor performance of high spatial subjects. This trend cannot be explained in terms of a greater determination on the part of high spatial ability subjects to locate the answers to complex questions and therefore appearing less navigationally efficient. The interactive effects of spatial ability and interface condition for response accuracy would suggest exactly the opposite, with low spatial ability subjects locating more correct answers in this condition.

On the basis of these results the effects of spatial visualisation cannot be solely explained by reference to differences in processing speed, as suggested by Seagull and Walker (1992). Whilst the significant correlations between spatial ability and response times for both menu component tasks indicates that processing speed may be an important component, it would appear that higher level cognitive processing is also involved. However, it is possible that the consistent association between spatial memory and response time, and the greater overall correlation when compared to spatial visualisation, was due to spatial memory being more strongly associated with the processing speed demands of this task. This hypothesis is supported by the comparatively large correlations between spatial memory and the menu component tasks, and would explain why spatial memory was less sensitive to interface differences.

Although the main effect of logical reasoning approached significance, none of the other measures of cognitive ability were significantly related to performance. The lack of a significant association between performance and verbal ability as measured by the Nelson-Denny vocabulary test is contrary to the findings of Vicente, Hayes and Williges (1987) but is consistent with those of Seagull and Walker (1992). The lack of significant association between associative memory and performance is contrary to the findings of Billingsley (1982), but does suggest that the effects of spatial memory are attributable to a second-order spatial ability factor rather than a general memory factor (Carroll, 1993). Similarly, it would appear that the higher-order factors of fluid intelligence or general intelligence cannot account for these results.

Table 7.10 : Details of information retrieval studies in which spatial visualisation has been examined as a predictor of response time				
Study	Age	Computer experience	Database structure	Correlation: VZ2 and RT
Vicente, Hayes and Williges (1987)	18-31 yrs.	Mixed	Hierarchical	-.57 (n=30)
Campagnoni and Erlich (1989)	Mixed	Mixed	Hierarchical	-.75 (n=12)
Jennings, Benyon, and Murray (1991)	25-40 years	Regular computer users	Mixed	(n=24) 'Button' = -.37 'Iconic' = -.02 'Menu' = -.07 'Question' = -.48 'Command' = -.58
Seagull and Walker (1992)	Undergrads.	1 year minimum	Linear and hierarchical	-.21 to -.37 (n=44) (linear = -.26)
Present study	Younger = 18-25 years Older = 45+ years	Less then 20 hours	Mixed	(n=45) linear = -.28 ex.- hi = -.31 em.- hi = -.42 net = .02

ex. hi = explicit menu - hierarchical structure : em. hi = embedded menu - hierarchical structure : net = embedded menu - network structure

The strength of the association between spatial visualisation and information retrieval response time was surprising small. This cannot be attributed to an inappropriate application of a linear model as quadratic and cubic alternatives were examined. Similarly, whilst differences across interface conditions serve to depress the overall correlation, those obtained for the linear and 'explicit menu - hierarchical structure' were modest. The magnitude of these correlations is in keeping with those of Seagull and Walker (1992), but substantially smaller than those reported by Campagnoni and Erlich (1989) or Vicente, Hayes and Williges (1987). In an attempt to uncover why this might be so sampling methods and task types were compared across studies. These are shown in Table 7.10

It would appear that restricted range of sample may in part explain the differences in the sizes of the correlations. The study which used the broadest sample in terms of age and experience (Campagnoni and Erlich, 1989) found the largest correlation between spatial visualisation and response time. Vicente, Hayes, and Williges (1987) used a more restricted age range for their sample, but recruited subjects with a wide range of computer expertise, a variable which they found to be significantly related to spatial ability. Seagull and Walker (1992), on the other hand, recruited a restricted sample in terms of both age and expertise. In the present study, the sample was selected from a fairly broad age range, but computer experience was restricted. It would appear, therefore, that sample restriction may play a part in reducing the magnitude of the correlation between spatial visualisation and response time.

7.4.3 The effects of age upon information retrieval performance

There was a significant age difference in the response times for the information retrieval task. This was particularly evident in the first performance block, and whilst the gap was substantially reduced by the second block, a relatively constant response time disadvantage remained. It is suggested that this pattern of results is indicative of an age difference in the initial acquisition of information retrieval tasks, possibly associated with the acquisition of a mental representation of the data structure, overlying a response time difference arising from a fundamental difference in processing speed (Salthouse, 1991b). Evidence supporting a processing speed related performance deficit can be seen in the significant age difference in response times for both of the menu component tasks. These tasks are thought to primarily demand processing speed and psychomotor skill rather than more complex cognitive processing. Further support can be found in the results of the regression analysis in which information retrieval response time was regressed upon response time for the embedded menu component task followed by age. The effects of adding age to the equation were extremely small suggesting that all the variance that would normally be associated with age had been accounted for by the component task. In addition, there is no evidence that the older age group were taking longer to answer questions because they were getting lost more frequently than the younger group. There was also no evidence of an age related speed accuracy trade-off.

The hypothesis that increased semantic content within the interface will lead to a reduction in age differences in information retrieval received little support. Neither of the two interface elements thought to induce greater semantic content obviously improved the performance of the older group. The magnitude of age differences with respect to response times (see Figure 7.11) was relatively consistent across conditions with the exception of the network structure for which the largest difference was evident. Whilst the younger subject group showed an improvement in the network condition when compared to the 'embedded menu - hierarchical condition' the performance of the older group declined. It would seem that either the increased semantic content resulting from the provision of additional relational links did not facilitate the performance of the older group, or that there were other spatial or complexity factors relating to this condition which had more profound age-related effects. The latter possibility will be discussed below. Similarly, a comparison of explicit menu and embedded menu hierarchy conditions for response time indicates no reduction in age related differences. The pattern of results for navigational performance show a small performance advantage for the older age group in all experimental conditions except the network condition (see Figure 7.12). This performance advantage is greatest in the 'embedded menu - hierarchical structure' condition. It is possible that this indicates an advantage for the increased semantic content provided by embedded menus which is not apparent in the network condition because it is overridden by other factors, but this is highly speculative and the effect is small. As regards accuracy (see Figure 7.13), performance is superior for the older group in all conditions except for the linear structure. This does not support the previously stated hypothesis.

There would appear to be two possible explanations for the detrimental effects of the network condition upon the performance of the older group. The first relates to the potential increase in spatial processing which may be a feature of the network structure. The provision of additional links may require the use of additional spatial processing resources in order to maintain a cognitive map of the database structure. Given the documented age related decline in spatial ability (Salthouse, Babcock, Skovronek, Mitchell, and Palmon, 1990; Salthouse and Mitchell, 1990) this may place older subjects at a disadvantage. This possibility, however, is not well supported by the interactive effects of spatial ability relating to the use of the network structure, as described above. The second possible explanation relates to the additional task complexity associated with the network structure. Selection of a node using this interface requires the examination of a greater number of alternatives. At many of the nodes there is an additional processing overhead which arises from the greater number of possible routes from which the user must select. Given the documented effects of certain task complexity manipulations upon age (Salthouse, 1992) this additional processing load may interact with age. These two possible explanations for the age-related performance deficit in the network condition are not mutually exclusive.

To summarise, it would appear that there are two elements which characterise the nature of age differences in information retrieval performance. The first of these relates to differences in processing speed which are fundamental, and which are independent of interface manipulations. The second relates to either the level of task complexity, the level of task spatial demands, or both. This leads to an initial performance disadvantage associated with the development of an adequate mental representation of the task, and also disadvantages the older group in interface conditions which favour high complexity and / or high spatial processing demands.

7.4.4 The relationship between age and cognitive ability

There was a significant difference in the regression slopes for spatial visualisation, spatial memory, and associative memory between the younger and older age groups with respect to the embedded menu component task. Whilst each of these predictors showed very little association with response times for the younger subject group, there were much stronger associations apparent for the older group. It would appear that the effects of cognitive ability upon this task increase in importance with age.

This interaction cannot be explained by an overall curvilinear relationship between cogntive ability and response time such that subjects low in cognitive ability experience a disproportionate performance decline. As mentioned earlier, the best fit for the regression equation is achieved using the linear term. One possible explanation is that these differences reflect age-related cognitive deficits. These are all cognitive abilities which have been found to deteriorate with age (Davies, Taylor, and Dorn, 1992). It may be that this is responsible for the slower response times of the older group. However, in this sample there was no significant age difference in these abilities, although the trend was in the predicted direction. Alternatively, it may be that these differences can be attributed to a shared common variance with a processing speed factor. Given that these component tasks are thought to be demanding of such abilities this is an attractive explanation. There was, however, a surprising interaction between the regression slopes for navigational efficiency and spatial memory for each age group. Contrary to the findings reported earlier with respect to the regression slopes for the component tasks, whilst high spatial memory was associated with improved navigational efficiency for the younger group, the reverse pattern was found for the older group.

7.4.5 Conclusions

The strength of the association between spatial ability and information retrieval performance was smaller than expected. Contrary to previous research spatial visualisation was not significantly correlated with response time. However, there was some evidence to suggest that this may have been partly due to differences between interfaces. Consistent with the experimental hypothesis, the magnitude of the correlation between spatial visualisation and performance was related to the level of semantic content of the interface, as manipulated by database structure and menu design. The weakest associations occurred when semantic content was high. An additional factor which was discussed with respect to the strength of this association was the range of the sample. It was suggested that studies in which weaker correlations have been obtained may have recruited a more restricted sample with respect to age or computer expertise. The significant association between spatial memory and information retrieval response times was in keeping with previous research. Upon the basis of the strength of association between spatial memory and response times for the menu component tasks it was suggested that this may be related to individual differences in processing speed.

There was a significant age difference in response time for the information retrieval task which was particularly evident for the first block of questions. It was suggested that this pattern of results may be indicative of consistent age differences in processing speed coupled with an initial disadvantage associated with the mental representation of the task environment by the older group. No support was gained for the hypothesis that increased semantic content within the interface would reduce the effects of age. However, trends in the data suggested that the complexity associated with the network structure disadvantaged the older group.

Chapter 8

Some Final Considerations

8.1 Introduction

This chapter draws together some of the main findings of this thesis and considers them in the light of earlier research. Where applicable, possible directions for future research and application are indicated. It should be noted, however, that a complete summary of experimental results and conclusions is not attempted. For this information the reader is referred to the closing sections in chapters 2 through 7. Instead, this chapter considers some particular areas of association between the empirical chapters and the initial 'assay' of individual differences in human-computer interaction which was presented in Chapter 1. Consequently, the following discussion is conducted under sub-headings representing the main dimensions of individual difference which were examined.

8.2 Verbal ability

As discussed in Chapter 1, although previous research has generally found verbal ability to be a good predictor of both word processing and, to a lesser extent, information retrieval performance, there are exceptions to this pattern. For example, the first two experiments reported by Egan and Gomez (1985) found that verbal ability was either not reliably associated with performance or was only modestly related. Results obtained in experiments reported in the present thesis throw some light on this situation. In the word processing experiment reported in Chapter 2, verbal ability was found to be associated with the process of 'coarse search' (the use of the page and scroll commands to display target words within computer presented text), but was not found to be generally predictive of the component tasks involved in the process of command generation. This latter position was also confirmed by the results of the command generation experiments reported in Chapters 5 and 6. Given that verbal ability is only significantly associated with the process of 'coarse search', these previously conflicting results can be explained. In the first two experiments conducted by Egan and Gomez (1985), mentioned above, a line editor was used. The interface of this application was such that there were no 'coarse search' demands placed upon the user. In contrast, in the fourth experiment of this series, a display editor was used, and verbal ability was found to be a strong predictor of performance. Previous contrary results have also been obtained within an information retrieval setting. Whilst Vicente, Hayes and Williges (1987) found verbal ability to be a strong predictor of performance, Seagull and Walker (1992) found no significant association. Similarly, the information retrieval experiment reported in Chapter 7 found no significant association between verbal ability and performance. However, a closer examination of the tasks used in these experiments reveals that the explanation for these conflicting results may lie in the comparative 'coarse search' demands. Vicente et al.'s (1987) subjects were required to locate information within 15 text files which varied in length between 55 and 447 lines. In contrast, the experiment by Seagull and Walker (1992), required subjects to locate items within a menu hierarchy with no similar text search demand. The 'coarse search' demands in the information retrieval experiment reported in Chapter 7, in which the maximum file length was 44 lines, were also substantially less than those imposed in the experiment of Vicente et al. (1987). These results are all consistent with the association between verbal ability and computer-based performance being limited to the process of 'coarse search'. However, the mechanics underlying this association are not clear and warrant further research.

Whilst the precise nature of the cognitive processing demands involved have yet to be established, target location appears to be an important factor. The results of the text search experiment presented in Chapter 3 indicated an interaction between verbal ability and target location in a younger sample. High verbal ability subjects were better at locating outer targets and worse at locating inner targets. The results of the experiments reported in Chapters 2 and 3 also indicate that the effects of verbal ability may interact with subject age and computer experience. These factors are discussed further in later sections.

It may be that individuals of low verbal ability would be able to locate target words within text more effectively by utilising performance strategies, such as a string search, which achieve the same end goal as 'coarse search' methods but impose different cognitive demands. This is an area which future research might usefully address.

In summary, verbal ability was found to be significantly associated with word processing performance, but no significant correlation was obtained with respect to the task of information retrieval. The results of the experiment reported in Chapter 2, in which a number of component elements of the word processing task were studied in isolation, indicate that verbal ability is only significantly associated with the process of 'coarse search'. These findings enabled a pattern to be determined with respect to previous experiments, in which the strength of the association between verbal ability and performance was dependent upon the 'coarse search' requirement within the task. Potentially, important factors related to this association include target location, subject age, and subject experience.

8.3 Reasoning ability

Previous studies have found logical reasoning to be predictive of word processing performance (Egan and Gomez, 1985; Czaja, Joyce, and Hammond, 1989). The results of the experiment reported in Chapter 2 are consistent with this position. In addition, logical reasoning was found to be a consistently good predictor of all of the component elements of word processing task performance. Whilst there was no significant main or interactive association between logical reasoning and response times for the command generation experiment reported in Chapter 5, the group of subjects with low logical reasoning ability reported significantly greater workload than the group high in logical reasoning ability. It would appear, therefore, that a process of compensation may have been occurring with additional attentional resources being invested by the low ability group in order to achieve similar levels of performance. On this basis, these results can be seen to be consistent with those of the word processing experiment, in that subjects of high logical reasoning ability can be presumed to be capable of comparatively better performance than subjects of low logical reasoning ability, given an equal allocation of attentional resources.

With respect to information retrieval, there was conflicting previous evidence as to the effects of logical reasoning. The results of on-line studies were mixed, whilst logical reasoning had been found to be significantly related to the use of a Structured Query Language. In the information retrieval experiment reported in Chapter 7 a trend was noted, such that subjects with high logical reasoning scores performed more quickly, however this failed to reach significance. It is difficult to explain why logical reasoning should be predictive of word processing performance but not information retrieval performance. One possible explanation relates not to the type of task but to the level of subject experience. The word processing experiment reported in Chapter 2 used experienced subjects, whilst the information retrieval experiment reported in Chapter 7 used novices. It may be that logical reasoning is more strongly predictive of practised performance. Further research might usefully address this issue.

In summary, on the basis of these results, it would appear that there is consistent evidence to believe that logical reasoning is predictive of word processing performance. However, the position with respect to information retrieval is still less than clear. Whilst these results may indicate that the predictive strength of logical reasoning is task dependent, it may also be related to other factors such as subject experience.

8.4 Spatial ability

Consistent with previous research, spatial ability was generally found to be a good predictor of performance across a wide range of tasks. Both spatial visualisation and spatial memory were predictive of response times for all the component elements of word processing performance in the experiment reported in Chapter 2. Interestingly, there was a tendency for this association to be stronger for the generating components than for the finding components. A correlation with the process of command generation is also supported by results reported in Chapter 5, in which spatial visualisation and spatial memory were found to be significantly associated with performance using both menu and command line as a method of command input, and again in Chapter 6, where spatial visualisation was found to be predictive of the search and selection of icons and text labels. Spatial memory was also predictive of performance accuracy for the text search experiment (Chapter 3), and response times for the information retrieval task reported in Chapter 7. In this latter experiment both spatial visualisation and spatial memory were significantly associated with the performance of the menu component tasks, with spatial memory being the stronger predictor.

Whilst this indicates the importance of spatial ability as a predictor of computer-based performance, little evidence was previously available with respect to the mechanics of the association between spatial ability and the content and demands of the interface. Attempts at reducing the effects of spatial ability by providing additional spatial information, within a database navigation context, had proved unsuccessful (Billingsley, 1982; Vicente and Williges, 1988). Experiments reported in Chapter 4 examined the association between spatial ability and the spatial and semantic content of the interface. Results were obtained which indicated that manipulations of the spatial information content of the interface did little to alter the variance associated with spatial ability. However, increases in the semantic content of the interface improved the performance of 'low spatial / high verbal ability' individuals relative to that of 'high spatial / low verbals'. Whilst this may have been attributable to a ceiling effect, support for this pattern of association was also obtained in the information retrieval experiment reported in Chapter 7. Manipulations of the linking structure and menu type which increased the availability of semantic information improved the relative performance of low spatials. However, manipulations of the spatial complexity of the information structure did not produce a similar effect. A number of potentially important implications follow from these results. For the interface to support the low spatial ability user it would appear to be essential that contextual support is provided. Whilst the provision of spatial information may assist in achieving these ends, the inclusion of spatial information content should not be the primary focus. A spatial representation illustrating that 'A' is linked to 'B' (as in the case of a database map) inevitably conveys a degree of semantic information, however, further contextual information might also include the nature of 'A' and 'B', the reason that 'A' is connected to 'B', and the result of moving from 'A' to 'B'. An example of the provision of such contextual information can be seen in the use of embedded menus, as discussed in Chapter 7. As mentioned in Chapter 4, further research might address the importance of particular types of semantic content in this respect.

A trend in the data reported in Chapter 7 indicated that high spatial ability subjects may be at a comparative performance disadvantage when using network data structures. A process of rectilinear normalisation was suggested as a possible mechanism which would explain such an association. Whilst this is only tentatively advanced, further experimental investigation of this area may prove interesting. It may be that in situations where information retrieval applications will be used predominantly by high spatial ability subjects, as in the case of a help system for a programming language, then optimal performance will be achieved with either linear or hierarchical information structures.

Contrary to prediction, there was no main effect of spatial visualisation upon the performance of the text search task reported in Chapter 3. However, there is some evidence to suggest that this may have been partially due to high spatial visualisation subjects adopting performance strategies which placed them at a disadvantage when the spatial information content of the interface was low. Similar evidence was also apparent with respect to spatial memory. A striking example was revealed when comparing performance using two window sizes. Whilst low and high spatial memory subjects reported similar levels of self-report workload in the large window, there was an interactive effect such that high spatial memory subjects reported comparatively greater workload in the small window condition. It would appear that, during text search, the performance strategy of high spatial ability subjects may rely upon the use of spatial information which places them at an advantage in conditions where there is high spatial information content within the interface, but places them at a disadvantage in conditions of low spatial information. There may be parallels between these results and work reported by Cooper and Mumaw (1985: cf. Chapter 4) relating to individual strategy differences during spatial problem solution. Further research might usefully investigate this possibility. These findings indicate that, when application users will be predominantly of high spatial ability, then the use of windows of less than 12 lines will be particularly detrimental to performance. An example of such a situation can be seen in the case of a display window for the results of a programming language compiler output.

In summary, consistent with previous research, spatial ability was found to be predictive of a wide variety of computer-based tasks. The results of menu and network navigation tasks reported in Chapter 4 indicated that the semantic content of the interface was more important than the spatial content of the interface in determining the performance variance associated with spatial ability. This pattern of results was also supported by the information retrieval experiment reported in Chapter 7. Further to this, an interaction between spatial ability and the spatial information content of the interface was identified in Chapter 3, in which the performance strategy employed by high spatial ability subjects appeared to place them at a comparative performance disadvantage when the spatial information content of the interface was low.

8.5 Short-term memory

As mentioned in the introductory chapter, comparatively little research has been devoted to the association between short-term memory (other than spatial memory) and the process of human-computer interaction. The experiment reported in Chapter 5 focused upon the effects of associative memory and the process of command generation. When using a command line interface there was both a main effect of associative memory and an interaction with command complexity, such that the performance advantage of high associative memory subjects increased in line with the number of required command parameters. When using a menu interface the performance of low associative memory novices was much slower than that of high associative memory novices, which was fairly similar to that of experts. Significant interactions involving the command method (menu vs command line) for both response time and strategy selection was taken to indicate that expert performers used a strategy best suited to their ability, with low associative memory being associated with menu use and high associative memory being associated with command line use. Prolonged use of a particular strategy based upon cognitive ability, in the case of expert subjects, resulted in improved aptitude for that command method. These findings illustrate the importance of a research plan which considers both individual differences in ability and individual differences in performance strategy for both novice and expert groups.

The experiment reported in Chapter 7 examined associative memory as a predictor of information retrieval performance. Contrary to the results of Billingsley (1982), there were no main or interactive effects. More promising results were obtained in an experiment briefly reported in Chapter 4. A measure of working memory efficiency was found to interact with the semantic content and target difficulty in a network navigation task. This measure may be worth pursuing in future research with respect to other information retrieval and wider human-computer interaction (HCI) settings.

In summary, little previous work has been done in this area and therefore experimental results will inevitably be relatively novel. The present experiments have demonstrated a relationship between associative memory and the process of command generation which, depending upon the nature of the interface, interacts with command complexity. Novices with low associative memory ability were found to be particularly disadvantaged, whilst experts were found to match ability and strategy, and become skilled in a preferred command method. Finally, a measure of working memory efficiency (cf. Woltz, 1988) was identified as a promising predictor for future research.

8.6 Expertise

An interaction between expertise and the level of cognitive demand associated with task performance was apparent in a number of the experiments reported in this thesis. In the text search experiment reported in Chapter 3 evidence was obtained which indicated that novices were at a greater performance disadvantage in conditions of high cognitive load. Similarly, in the command generation experiment (Chapter 5) there was a main effect of expertise and a significant interaction with command complexity. The comparative performance advantage of experts increased in line with the number of command parameters required. These results are consistent with Anderson's (1983) model of skill acquisition, as mentioned in Chapter 1. An increased reliance upon declarative knowledge, or controlled processing (Schneider and Shiffrin, 1977), on the part of novice users, results in fewer available attentional resources with which to meet the increased resource demands of more complex task conditions. The results of these experiments stress the importance of 'visual momentum' in this respect (Woods, 1984; cf. Chapter 3). The cognitive costs associated with the process of reorienting within a program task space should be avoided or reduced whenever possible. Similarly, in keeping with the substructured command investigation of Macaulay and Norman (1984), the results of the command generation experiment reported in Chapter 5 indicate that command complexity should be minimised. In keeping with a number of previous experiments (Benbasat, Dexter, and Masulis, 1981; Whiteside, Jones, Levy, and Wixon, 1985; Antin, 1988), there was no interactive effect of expertise and command method. However, as discussed in the previous section, associative memory appears to be an important interactive factor which future studies of command method and expertise must take into account.

The results from a number of experiments supported Ackerman's (1988) theory of individual differences in skill acquisition, as outlined in Chapter 1. The magnitude of the effect of cognitive ability was consistently found to be greater for the novice group than the expert group. In the text search experiment reported in Chapter 3 the interactive effects including verbal ability and expertise supported this position. In the command generation experiment reported in Chapter 5 the interactive effects of both spatial memory and associative memory with expertise were consistent with this position. However, whilst there was evidence to suggest that the effects of cognitive ability were more important with respect to novice performance, the significant correlations reported in Chapter 2, indicates that cognitive ability may remain an important predictor of computer-based performance for experienced computer-users. This is important in terms of the application of such research. As outlined in the introductory chapter, there would be little benefit in making selection decisions based upon criteria which only applied for an introductory period. Within this context, it should also be noted that the tasks used in this experiment were comparatively simple word processing operations, with fairly consistent mapping in many respects (cf. Schneider and Shiffrin, 1977). It is reasonable to expect the impact of cognitive ability to increase in line with cognitive processing load. Nevertheless, future research might usefully examine cognitive ability, processing speed and psychomotor skill as a function of the acquisition of computer-based expertise. Upon the basis of Ackerman's model these predictors would be expected to vary in their comparative importance as skill was developed.

8.7 Age

Consistent with previous research, an age difference in performance speed was found for the word processing experiment reported in Chapter 2. However, whereas the model presented by Egan and Gomez (1985) suggests that age is only associated with the process of command generation, results indicated that it was in fact predictive of all component elements of the task, although this effect was stronger for the command generation elements.

Two possible compensatory mechanisms were identified in the word processing experiment (Chapter 2) which may have enabled some older subjects to maintain performance at or near the level of younger subjects. The first of these was related to verbal ability. As mentioned above, verbal ability was found to be related to the process of 'coarse search'. Regression analyses also revealed an interaction with subject age, such that for older subjects there was a negative association between verbal ability and response time, whilst for younger subjects verbal ability predicted little performance variance. On the assumption that verbal ability is maintained, or improves, into old age (Davies, Taylor, and Dorn, 1992), whilst other abilities such as spatial ability and reasoning ability decline (Horn and Hofer, 1992), it was suggested that older subjects with high verbal ability were employing performance strategies which enabled them to take maximum advantage of available resources. However, it must be remembered that this experiment used an expert sample who may have been able to refine their performance based upon previous experience. A second mechanism of compensation was identified using a process of molar equivalence-molecular decomposition (Charness, 1981a, 1981b, Salthouse, 1984). A tendency was found for older subjects of similar molar task ability to be more skilled at the typing elements of task performance. It was suggested that this was also a compensatory mechanism.

There has been a scarcity of previous research in the area of information retrieval, and whilst the evidence from other areas of human-computer interaction research indicates a probable pattern of results, the findings reported in Chapters 4 and 7 are novel in this respect. With regard to the main information retrieval task (Chapter 7), a performance speed disadvantage was apparent for the older group throughout the task, however this was particularly so in the first performance block. It was suggested that this was indicative of an initial disadvantage for the older group in mentally representing the task, overlying a continuing processing speed disadvantage. Support for a processing speed deficit can be found in the consistently poorer performance of the older group upon all the word processing component tasks, the poorer performance of the older group on the menu component tasks of the information retrieval experiment, and the consistent response time disadvantage for the older group in the menu navigation experiment, briefly reported at the end of Chapter 4. Whilst an interaction was found between age and verbal ability for the 'coarse search' tasks of the word processing experiment (Chapter 2), in the information retrieval experiment (Chapter 7) similar interactive effects were found between age and spatial visualisation, spatial memory, and associative memory with respect to the embedded menu component task. The sample in this latter experiment, however, comprised novice users, unlike that of the word processing experiment. It would appear that the effects of age in relation to this task may be partially explained by a decline in specific age sensitive cognitive abilities.

Little support was apparent for the hypothesis, advanced in Chapters 4 and 7, that the 'classic pattern' of cognitive ageing (cf. Davies, Taylor, and Dorn, 1992) would lead to older subjects disproportionately benefiting from increased semantic content within the interface. There was a non significant trend in the menu navigation experiment (Brief report 1 - Chapter 4) such that spatial information was of greater importance to younger subjects whilst older subjects relied more heavily upon semantic information content. However, in the information retrieval experiment there was a performance trend which suggested that the structural complexity of the database may be a more important factor, with the older group being at a disadvantage when complexity is high. In such situations more selection alternatives would have to be considered throughout the navigation process. This was also supported by the results of the secondary task experiment reported briefly at the end of Chapter 4, the results of which indicated a processing resource disadvantage for older subjects when performing navigation tasks. Taken as a whole, these results suggest that older subjects may differ from younger subjects in the efficiency with which they are able mentally to represent information relating to database navigation, but that the nature of this representation does not vary with age in a novice sample.

In summary, consistent with previous research, a response time disadvantage was continually found for the older group. However, in the word processing experiment (Chapter 2) two potential compensatory mechanisms were identified which may have allowed certain older users to maintain performance levels. The first of these related to verbal ability, whilst the second related to typing speed. A reduction in processing resource availability was identified for older subjects, along with a trend towards poorer information retrieval performance in conditions of high database structural complexity.

8.8 Workload

The importance of assessing workload was demonstrated in this thesis. The NASA TLX measure of self-report workload was found to be sensitive to changes in interface variables. In the text search experiment reported in Chapter 3 there was a significant effect of window size upon self-report workload, indicating the potentially disadvantageous effects of reducing window size below a certain level. There were also instances when the NASA TLX proved to be sensitive to individual differences. In the text search experiment, as mentioned above, there was a significant interaction between spatial memory and window size for self-report workload, and in the command generation experiment reported in Chapter 5 there was a main effect of logical reasoning. This latter effect is a particularly good illustration of the importance of assessing workload. Without this measure there would have been no indication of group differences with respect to logical reasoning. Similarly, whilst there were no significant age differences in self-report workload in either the word processing experiment (Chapter 2) or the information retrieval experiment (Chapter 7), this information provided useful evidence indicating that the effects of age with respect to response times and error rates could not be attributed to individual differences in resource allocation strategies.

8.9 Final comments

An interesting research strategy which may be employed in a future examination of the association between individual differences in cognitive ability and computer-based performance involves the use of molar equivalence-molecular decomposition methodology (Charness, 1981a, 1981b, Salthouse, 1984), as described in Chapter 2. However, instead of examining subject groups which differ in age, as has previously been the case, this experimental method could be applied to groups of extreme cognitive ability. Experimental results presented in Chapters 3 and 5 indicate the importance of the interactive effects of cognitive ability and performance strategy. This research methodology may enable the identification of successful compensatory mechanisms which are employed by subjects of low cognitive ability. Such information would be extremely valuable in the design of training programs, allowing an individually tailored approach based upon cognitive ability. Similarly, this information would have implications for the design of the interface, highlighting important features with respect to the variance associated with individual differences.

Some evidence was presented in Chapter 7 to suggest that sampling methods may account for much of variance between different experiments in the size of the association between cognitive ability and human-computer interaction. In particular the effects of age and computer experience may be of importance in this respect. When much research is conducted within academic settings, samples are often unavoidably restricted. This may lead to an underestimation of the importance of the study of individual differences in human-computer interaction.

Whilst there is still insufficient evidence available in this area to attempt to define a taxonomy of individual differences in computer-based performance, the present thesis has hopefully further clarified this association. Several promising areas for future research have been identified. Certainly these results highlight the importance of considering individual differences with respect to the processes of usability testing, training, interface design, and selection.

References:

Ackerman, P.L. (1987) Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin,* 102, 3-27.

Ackerman, P.L. (1988) Determinants of individual diffences during skill acquisition: Cognitive abiltities and information processing. *Journal of Experimental Psychology: General.* 117, 3, pp. 288-318.

Ackerman, P.L. (1992) Predicting individual differences in complex skill aquisition: Dynamics of ability determinants. *Journal of Applied Psychology*, 77, 5, 598-614.

Ackerman, P.L. & Kanfer, R. (1988) Declarative and procedural knowledge in skill acquisition: An aptitude-treatment interaction framework for training. In *Proceedings of the Human Factors Society - 32nd Annual Meeting*. pp. 1241-1245.

Adamowicz, J.K. (1978) Visual short-term memory, age, and imaging ability. *Perceptual and Motor Skills*, 46, 571-576.

Allen, B.A. (1990) Knowledge organization in an information retrieval task. *Information Processing & Management*, 26, 4, 535-542.

Alvarez, M.G., Murray, J.T., & Hakkinen, M.T. (1984) Search performance as a function of depth of processing and scrolling rate. *Proceedings of the Human Factors Society - 28th Annual Meeting*. p. 709.

Ambardar, A.K. (1984) Human-computer interaction and individual differences. In G. Salvendy (ed.) *Human-Computer Interaction*. North Holland: Elsevier Science. pp. 207-211.

Ambardar, A.K. (1988) Human-computer interaction: Analyses of individual differences and decision making. In *Proceedings of the Human Factors Society: 32nd Annual Meeting*. pp. 826-830.

Anderson, J.R. (1982) Acquisition of cognitive skill. *Psychological Review*, 89, 4, 369-406.

Anderson, J.R. (1983) *The Archictecture of Cognition.* Cambrindge, MA: Harvard University Press.

Anderson, J.R. (1987) Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review*, 94, 2, 192-210.

Anderson, J.R. & Reiser, B.J. (1985) The LISP tutor. *Byte*, 10, 159-175.

Anderson, J.R. & Skwaecki, E. (1986) The automated tutoring of introductory computer programming. *Communications of the ACM*, 29, 9, 842-849.

Antin, J.F. (1988) An empirical comparison of menu selection, command entry and combined modes of computer control. *Behaviour and Information Technology*, 7, 2, 173-182.

Anzai, Y. (1991) Learning and use of representations for physics expertise. In K.A. Ericsson and J. Smith (eds.) *Toward a General Theory of Expertise*. Cambridge: Cambridge University Press.

Arend, U., Muthig, K-P., & Wandmacher, J. (1987) Evidence for global feature superiority in menu selection by icons. *Behaviour and Information Technology*, 6, 4, 411-426.

Babcock, R.L. & Salthouse, T.A. (1990) Effects of increased processing demands on age differences in working memory. *Psychology and Aging*, 5, 3, 421-428.

Baber, C., Usher, D.M., Stammers, R.B. & Taylor, R.G. (1992) Feedback requirements for automatic speech recognition in the process control room. *International Journal of Man-Machine Studies*, 37, 703-719.

Baddeley, A. (1986) *Working Memory*. Oxford: Clarendon Press

Baecker, R.M. & Buxton, W.A.S. (eds.) (1987) *Human-Computer Interaction.* San Mateo, CA.: Morgan Kaufmann.

Barker, D., Carey, M.S., & Taylor, R.G. (1990) Factors underlying mouse pointing performance. In E.D. Megaw (ed.) *Contemporary Ergonomics 1990*:

Proceedings of the Ergonomics Society 1988 Annual Conference. London : Taylor & Francis.

Barker, P.G., Najah, M., & Manji, K.A. (1987) Pictorial communication with computers. *International Journal of Man-Machine Studies*, 27, 315-336.

Barnard, P.J., Hammond, N.V., MacLean, A. & Morton, J. (1982) Learning and remembering interactive commands in a text-editing task. *Behaviour and Information Technology*, 1, 4, 347-358.

Bateson, A., Alexander, R.A., & Murphy, M.D. (1987) Cognitive processing differences between novice and expert computer programmers. *International Journal of Man-Machine Studies*, 26, 649-660.

Battig, W.F. & Montague, W.E. (1969) Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology (Monograph)*, 80, 3, 2, 1-46.

Battmann, W. (1986) Problems in the design of information retrieval systems: User competence and information complexity. In F. Klix & H. Wandke (eds.) *Man-Computer Interaction Research. MACINTER I.* North Holland: Elsevier Science. pp. 131-138.

Bellardo, T. (1985) An investigation of online searcher traits and their relationship to search outcome. *Journal of the American Society for Information Science*, 36, 4, 241-250.

Benbasat, I., Dexter, A.S., & Masulis, P.S. (1981) An experimental study of the human-computer interface. *Communications of the ACM*, 24, 752-762.

Benbasat, I. & Todd, P. (1993) An experimental investigation of interface design alternatives: Icon vs. text and direct manipulation vs. menus. *International Journal of Man-Machine Studies*, 38, 369-402.

Bennett, G.K., Seashore, H.G., & Wesman, A.G. (1972) *Differential Aptitude Tests*. New York: Psychological Corporation.

Benyon, D. & Murray, D. (1988) Experience with adaptive interfaces. *The Computer Journal*, 31, 5, 465-473.

Billingsley, P.A. (1982) Navigation through hierarchical menu structures: Does it help to have a map? In *Proceedings of the Human Factors Society - 26th Annual Meeting.*

Blankenberger, S. & Hahn, K. (1991) Effects of icon design on human-computer interaction. *International Journal of Man-Machine Studies,* 35, 363-377.

Bolla, K., Lindgren, K.N., Bonaccorsy, C., & Bleeker, M.L. (1990) Predictors of verbal fluency (FAS) in the healthy elderly. *Journal of Clinical Psychology,* 46, 5, 623-628.

Borenstien, N.S. (1985) The evaluation of text editors: A critical review of the Roberts and Moran methodology based on new experiments. In L. Borman & B. Curtis (Eds.), *Human Factors in Computing Systems II.* pp. 99-105.

Borgman, C.L. (1984) The users mental model of an information retrieval system: Effects on performance. Unpublished doctoral dissertation: Stanford University.

Borgman, C.L. (1985) Individual differences in the use of technology: work in progress. *ASIS '85 Proceedings of the 48th ASIS Annual Meeting,* Las Vegas, NV, USA, 20-24 Oct, 1985. White Plains, NY, USA: Knowledge Ind. Publications. pp. 243-249.

Borgman, C.L. (1986) The users mental model of an information retrieval system: An experiment on a prototype online catalogue. *International Journal of Man Machine Studies,* 24, 47-64 .

Borgman, C.L. (1989) All users of information retrieval systems are not created equal: An exploration into individual differences. *Information Processing and Management (U.K.),* 25, 3, 237-251.

Breakwell, G.M. & Fife-Schaw, C. (1988) Aging and the impact of new technology. *Social Behaviour,* 3, 119-130.

Brindle, E.A. (1981) The relationship between characteristics of searchers and their behaviors while using an online interactive retrieval system. Unpublished doctoral dissertation: Syracuse University. DAI Order No. 8123887

Bringelson, L.S. & Eberts, R. (1990) The effect of cognitive workload, job enrichment environment, and spatial ability on goal structures and strategies. In *Proceedings of the Human Factors Society - 34th Annual Meeting*. pp. 1514-1518.

Brosey, M.K. & Shneiderman, B. (1978) Two experimental comparisons of relational and hierarchical database models. *International Journal of Man-Machine Studies*, 10, 625-637.

Brown, J.I., Bennett, J.M., & Hanna, G. (1981) *The Nelson-Denny Reading Test. Examiner's Manual, Forms E and F*. Riverside Publishing Co.

Brown, M.L., Newsome, S.L., & Glinert, E.P. (1989) An experiment into the use of auditory cues to reduce visual workload. In K. Bice & C. Lewis (eds.) *Human Factors in Computing Systems. CHI '89 Proceedings*. Austin, Texas. 30th April-4th May 1989. ACM Press. pp. 339-346.

Burnett, S.A. (1986) Sex-related differences in spatial ability: Are they trivial? *American Psychologist*, 41, 9, 1012-1014.

Bush, C.M. & Schkade, L.L. (1985) In search of the perfect programmer. *Datamation*, 31, 6, 128-132.

Campagnoni, F.R. & Ehrlich, K. (1989) Information retrieval using a hypertext-based help system. *ACM Transactions on Information Systems*, 7, 3, 271-291.

Campbell, P.F. & McCabe, G.B. (1984) Predicting the success of freshmen in a computer science major. *Communications of the ACM*, 27(11), 1108-1113.

Campos, A. & Cofan, E. (1986) Rotation of images and primary mental abilities: Influence of information and sex. *Perceptual and Motor Skills*, 63, 2, pt 1, 644-646.

Canter, D, Rivers, R., & Storrs, G. (1985) Characterising user navigation through complex data structures. *Behaviour and Information Technology*, 4, 2, 87-91.

Caplan, P.J., MacPherson, G.M., & Tobin, P. (1985) Do sex-related differences in spatial abilities exist? A multilevel critique with new data. *American Psychologist*, 40, 7, 786-799.

Caplan, P.J., MacPherson, G.M., & Tobin, P. (1986) The magnified molehill and the misplaced focus: Sex-related differences in spatial ability revisited. *American Psychologist*, 41, 9, 1016-1018.

Card, S.K., Moran, T.P., and Newell, A. (1983) *The Psychology of Human-Computer Interaction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Carpenter, P.A., Bishu, R.R., & Riley, M.W. (1990) User characteristics: Are personality types and psychometric factors good predictors? In *Proceedings of the Human Factors Society - 34th Annual Meeting*. pp. 351-355.

Carpenter, P.A. & Just, M.A. (1975) Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, 2, 1, 45-73.

Carroll, J.B. (1986) Lecture 1: Factor-analytic investigations of cognitive abilities. In S.E. Newstead, S.H. Irvine, and P.L. Dann (eds.) *Human Assessment: Cognition and Motivation*. Dordrecht: Martinus Nijhoff. pp. 3-8.

Carroll, J.B. (1993) *Human Cognitive Abilities*. Cambridge: Cambridge University Press.

Carroll, J.M. (1984) Minimalist design for active users. In B. Shackel (Ed.) *Human-Computer Interaction - INTERACT '84*. North Holland: Elsevier Science. pp. 39-44.

Carroll, J.M. & Carrithers, C. (1984) Blocking learner error states in a training-wheels system. *Human Factors*, 26,4, 377-389.

Carroll, J. & Mack, R.L. (1985) Metaphor, computing systems, and active learning. *International Journal of Man-Machine Studies*, 22, 39-57.

Carroll, J.M., Mack, R.L., & Kellog, W.A. (1988) Interface metaphors and user interface design. In M. Hellander (ed.) *Handbook of Human-Computer Interaction*. Amsterdam: North Holland. pp. 67-85.

Carroll, J.M. & Olson, J.R. (1988) Mental models in human-computer interaction. In M. Hellander (ed.) *Handbook of Human-Computer Interaction.* Amsterdam: North Holland. pp. 45-65.

Cattell, R.B. (1957) *Personality and Motivation Structure and Measurement.* Yonkers: New World.

Cattell, R.B. (1971) *Abilities: Their Structure, Growth, and Action.* Boston: Houghton Mifflin.

Cattell, R.B. (1972) The 16PF and basic personality structure: A reply to Eysenck. *Journal of Behavioural Science*, 1, 168-188.

Cattell, R.B. (1973) *Personality and Mood by Questionnaire.* New York: Jossey-Bass.

Cattell, R.B. & Horn, J.L. (1978) A check on the theory of fluid and crystallized intelligence with a description of new subtest designs. *Journal of Educational Measurement*, 15, 139-164.

Cerella, J., Poon, L.W., & Williams, D.M. (1980) Age and the complexity hypothesis. In L.W. Poon (ed.), *Aging in the 1980s: Psychological Issues.* Washington, DC: American Psychological Association. pp. 332-342.

Charness, N. (1981a) Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 2, 467-476.

Charness, N. (1981b) Aging and skilled problem solving. *Journal of Experimental Psychology: General*, 110, 1, 21-38.

Charness, N. (1991) Expertise in chess: The balance between knowledge and search. In K.A. Ericsson and J. Smith (eds.) *Toward a General Theory of Expertise.* Cambridge: Cambridge University Press.

Charness, N. & Bosman, E.A. (1990) Expertise and aging: Life in the lab. In T.M. Hess (ed.) *Aging and Cognition.* Amsterdam: North Holland.

Charness, N., Schumann, C.E., & Boritz, G.M. (1992) Training older adults in word processing: Effects of age, training technique, and computer anxiety. *International Journal of Technology and Aging*, 5, 1, 79-106.

Chen, H-G. & Vecchio, R.P. (1992) Nested IF-THEN-ELSE constructs in end-user computing: Personality and aptitude as predictors of programming ability. *International Journal of Man-Machine Studies*, 36, 843-859.

Cherry, J.M., Fischer, M.J., Fryer, B.M., & Steckham, M.J. (1989) Modes of presentation for on-line help: full screen, split screen and windowed formats. *Behaviour and Information Technology*, 8,6, 405-416.

Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981) Expertise in problem solving. *Cognitive Science*, 5, 121-152.

Clarke, V.A. (1990) Sex differences in computing participation - concerns, extent, reasons and strategies. *Australian Journal of Education*. 34(1), 52-66.

Cohen, G. & Faulkner, D. (1983) Age differences in performance on two information-processing tasks: strategy selection and processing efficiency. *Journal of Gerontology*, 38, 4, 447-454.

Collins, A.M. & Quillian, M.R. (1969) Retrieval time from semantic memory, *Journal of Verbal Learning and Verbal Behaviour*, 8, 240-248.

Collins, A.M. & Quillian, M.R. (1970) Does category size affect categorisation time? *Journal of Verbal Learning and Verbal Behaviour*, 9, 432-438.

Conklin, J. (1987) Hypertext: An introduction and survey. *Computer*, 20, 9, 17-41.

Cooper, L.A. (1976) Individual differences in visual comparison processes. *Perception and Psychophsyics*, 19, 5, 433-444.

Cooper, L.A. & Mumaw, R.J. (1985) Spatial aptitude. In R.F. Dillon (Ed.) *Individual Differences in Cognition Vol. II.*, pp. 67-94.

Cooper, L.A. & Regan, D.T. (1982) Attention, perception, and intelligence. In R.J. Sternberg (ed.) *Handbook of Human Intelligence*. Cambridge: Cambridge University Press. pp. 123-169.

Coventry, L. (1989) Some effects of cognitive style on learning UNIX. *International Journal of Man-Machine Studies*. 31, 349-365.

Craik, F.I.M. & Lockhart, R.S. (1972) Levels of processing: A framewok for memory research. *Journal of Verbal Learning and Verbal Behaviour*, 11, 671-684.

Cronbach, L.J. & Snow, R.E. (1981) *Aptitudes and Instructional Methods*. New York: Irvington.

Czaja, S.J., Joyce, J.B. & Hammond, K. (1989) A comparative study of text editing programs among a sample of older adults. In *Proceedings of the Human Factors Society. 33rd Annual Meeting*. pp. 179-181.

Czaja, S.J., Hammond, K., Blascovich, J.J., & Swede, H. (1986) Learning to use a word-processing system as a function of training strategy. *Behaviour and Information Technology*, 5, 203-216.

Czaja, S.J., Hammond, K., Blascovich, J.J., & Swede, H. (1989) Age related differences in learning to use a text-editing sytem. *Behaviour and Information Technology*, 8, 4, 309-319.

Dambrot, F.H., Watkins-Malek, M.A., Silling, M.S., Marshall, R.S., & Garver, J.A. (1985) Correlates of sex differences in attitudes towards and involvement with computers. *Journal of Vocational Behavior*, 27, 71-86.

Darnell, M.J. & Neal, A.S. (1983) Text editing performance with partial and full page displays. In *Proceedings of the Human Factors Society - 27th Annual Meeting*. pp. 821-825.

Davies, D.R., Taylor, A., & Dorn, L. (1992) Aging and Human Performance. In *Handbook of Human Performance*. Academic Press.

Davies, D.R., Wong, C.S.K., Glendon, A.I., Stammers, R.B., & Matthews, G. (1993) Age, health status, and the acquisition of office technology skills. E.S.R.C. Project Final Report. Aston University, Birmingham.

Davies, S.P., Lambert, A.J., & Findlay, J.M. (1989) The effects of the availability of menu information during command learning in a word processing application. *Behaviour and Information Technology*, 8,2,135-144.

Davis, F.D. (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13, 310-340.

Davis, S. & Bostrom, R. (1992) An experimental investigation of the roles of the computer interface and individual characteristics in the leaning of computer systems. *International Journal of Human-Computer Interaction*, 4, 2, 143-172.

DeGroot, A. (1978) *Thought and Choice in Chess*. The Hague: Mouton.

DeNelsky, G.Y. & McKee, M.G. (1974) Prediction of computer programmer training and job performance using the AABP test. *Personnel Psychology*, 27, 129-137.

De Raad, B, Hendricks, A.A.J., & Hofstee, W.K.B. (1992) Towards a refined structure of personality traits. *European Journal of Personality*, 6, 301-319.

Derrick, W.L., McCloy, T.M., Marshak, W.P., Seiler, G.L., & Reddick, P.A. (1986) The effect of spatial ability on the demand for spatial processing resources. In *Proceedings of the Human Factors Society - 30th Annual Meeting*. pp. 624-627.

Dewier, A. & Kernas, G. (1991) Level of experience in text-editing: The role of mental representations on performance. In M.J. Tauber and D. Ackermann (eds.) *Mental Models in Human-Computer Interaction II*. North Holland: Elsevier Science. pp. 191-201.

Digman, J.M. (1990) Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417-440.

Digman, J.M. & Inouye, J. (1986) Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology*, 50, 116-123.

Dillon, A., McKnight, C., & Richardson, J. (1990) Navigation in hypertext: A critical review of the concept. In D. Diaper, D. Gilmore, G. Cockton, & B. Shackel (eds.) *Human-Computer Interaction INTERACT '90. Proceedings of the IFIP TC Third International Conference on Human-Computer Interaction*, Cambridge, U.K., 27-31 August, 1990. Amsterdam: North Holland. pp. 587-592.

Dillon, A., Richardson, J. and McKnight, C. (1990) The effects of display size and text splitting on reading lengthy text from screen. *Behavior and Information Technology*, 9, 3, 215-227.

Dixon, R.A. (1992) Contextual approaches to adult intellectual development. In R.J. Sternberg & C.A. Berg (eds.) *Intellectual Development*, Cambridge: Cambridge University Press.

Drenth, P.J.D. & Algera, J. (1987) Personnel selection. In P. Warr (ed.) *Psychology at Work*, London: Penguin.

Duchnicky, R.L. & Kolers, P.A. (1983) Readability of text scrolled on visual display terminals as a function of window size. *Human Factors*, 25,6, 683-692.

Dukes, R.L., Discenza, R., and Couger, J.D. (1989) Convergent validity of 4 computer anxiety scales. *Educational and Psychological Measurement*, 49(1), 195-203.

Dupree, D.A. & Wickens, C.D. (1982) Individual differences and stimulus discriminability in visual comparison reaction time. In *Proceedings of the Human Factors Society - 26th Annual Meeting*. PP. 809-811.

Durding, B.M., Becker, C.A., & Gould, J.D. (1977) Data organization. *Human Factors*, 19, 1, 1-14.

Eberleh, E., Korfmacher, W. & Streitz, N.A. (1992) Thinking or acting? Mental workload and subjective preferences for a command code and a direct manipulation interaction style. *International Journal of Human-Computer Interaction*, 4, 2, 105-122.

Egan, D.E. (1979) Testing based on understanding: Implications from studies of spatial ability. *Intelligence*, 3, 1-15.

Egan, D.E. (1988) Individual differences in human-computer interaction. In M. Hellander (Ed.) *Handbook of Human_computer Interaction*. Elsevier Science: North Holland. pp. 543-569.

Egan, D.E., Bowers, C., & Gomez, L.M. (1982) Learner characteristics that predict success in using a text-editor tutorial. In *Proceedings of Human Factors in Computer Systems*. Gaithersburg, MD. March 15-17, 1982.

Egan, D.E. & Gomez, L.M. (1982) Characteristics of people who can learn to use computer text editors: Hints for future text editor design and training. In *Proceedings of the ASIS Annual Meeting*. Volume 19, 75-79.

Egan, D.E. & Gomez, L.M. (1985) Assaying, isolating, and accomodating individual differences in learning a complex skill. In R.F. Dillon (Ed.) *Individual Differences in Cognition*. Accademic Press: London. pp. 173-217.

Egan, D.E., Gomez, L.M., McKeown, K., Soloway, E., Reiser, B., & Marshall, C. (1988) Panel: Dealing with diversity: Approaches to individual differences in human-computer interaction. In *Conference on Human Factors in Computing Systems, CHI '88*, May, 1988, Washington DC. ACM, SIGCHI. pp. 79-81.

Egan, D.E. & Grimes-Farrow, D.D. (1982) Differences in mental representation spontaneously adopted for reasoning. *Memory and Cognition,* 10, 4, 297-307.

Eggemeier, F.T. & Stadler, M.A. (1984) Subjective workload assessment in a spatial memory task. *Proceedings of the Human Factors Society - 28th Annual Meeting*. pp. 680-684.

Ekstrom, R.B., French, J.W., & Harman, H.H. (1976) *Manual for Kit of Factor-Referenced Cognitive Tests*. Priceton, NJ: Educational Testing Service.

Elias, P.K., Elias, M.F., Robbins, M.A., & Gage, P. (1987) Acquistion of word-processing skills by younger, middle-aged, and older adults. *Psychology and Aging,* 2, 340-348.

Elkerton, J., & Williges, R.C. (1984a) Information retrieval strategies in a file-search environment. *Human Factors*, 26, 2, 171-184.

Elkerton, J. & Williges, R.C. (1984b) The effectiveness of a performance-based assistant in an information retrieval environment. In *Proceedings of the Human Factors Society - 28th Annual Meeting.*

Elkerton, J. & Williges, R.C. (1985) A performance profile methodology for implementing assistance and instruction in computer-based tasks. *International Journal of Man-Machine Studies,* 23, 135-151.

Elkerton, J. & Williges, R.C. (1987) A summary of experimental research on command-selection aids. In H.-J. Bullinger (Ed.), *Human-Computer Interaction. INTERACT '87.* North Holland: Elsevier Science. pp. 937-942.

Elkerton, J., Williges, R.C., Pittman, J.A. & Roach, J. (1982) Strategies of interactive file search. In *Proceedings of the Human Factors Society 26th Annual Meeting.* Santa Monica, Ca. : Human Factors Society. pp. 83-86.

Elsom-Cook, M. (1989) Acquisition of computing skills. In A.M. Colley & J.R. Beech (eds.) *Acquisition and Performance of Cognitive Skills.* Chichester: Wiley. pp. 137-165.

Engel, F.L., Andriessen, J.J., & Schmitz, H.J. (1983) What, where, and whence: Means for improving electronic data access. *International Journal of Man-Machine Studies,* 18, 145-160.

Ericsson, K.A. & Smith, J. (1991) Prospects and limits of the emprical study of expertise: An introduction. In K.A. Ericsson & J. Smith (eds.) *Toward a General Theory of Expertise.* Cambridge: Cambridge University Press.

Evans, G.E. & Simkin, M.G. (1989) What best predicts computer proficiency? *Communications of the ACM,* 32,11,1322-1327.

Ewing, J., Mehrabanzad, S., Sheck, S., Ostroff, D., & Shniederman, B. (1986) An experimental comparison of a mouse and arrow-jump keys for an interactive encyclopedia. *International Journal of Man-Machine Studies,* 24, 29-45.

Eysenck, H.J. (1967) *Biological Basis of Personality.* Springfield, IL: Thomas.

Eysenck, H.J. & Eysenck, M.W. (1985) *Personality and Individual Differences.* New York: Plenum Press.

Eysenck, M.W. (1976) Arousal, Learning, and Memory. *Psychological Bulletin*, 83, 3, 389-404.

Eysenck, M.W. & Keane, M.T. (1990) *Cognitive Psychology*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Fetler, M. (1985) Sex differences in the California statewide assessment of computer literacy. *Sex Roles*, 13, 3-4, 181-192.

Fisher, D.L., Yungkurth, E.J., & Moss, S.M. (1990) Optimal menu hierarchy design: Syntax and semantics. *Human Factors*, 32,6, 665-683.

Fisk, A.D. (1987) High performance cognitive skill acquisition: Perceptual/rule learning. In *Proceedings of the Human Factors Society, 31st Annual Meeting*. pp. 652-656.

Fleishman, E.A. (1972) On the relation between abilities, learning, and human performance. *American Psychologist*, 27, 1017-1032.

Flesch, R. (1949) *The Art of Readable Writing*. New York: Harper and Evanston.

Foos, P.W. (1989) Adult age differences in working memory. *Psychology and Aging*, 4, 3, 269-275.

Foreman, K. H-D. (1988) Cogntive style, cognitive ability, and the acquisition of initial computer programming competence. Paper presented at the Annual Meeting of the Association for Educational Communications and Technology, New Orleans, LA, Jan. 14-19, 1988.

Fowler, C.J.H., Macaulay, L.A., & Fowler, J.F. (1985) The relationship between cognitive style and dialogue style: An explorative study. In P.Johnson & S. Cook (Eds.), *People and Computers: Designing the Interface*. Cambridge University Press. pp. 186-198.

Fowler, C.J.H. & Murray, D. (1987) Gender and cognitive style differences at the human-computer interface. In H.J. Bullinger & B. Shackel (Eds.), *Human-Computer Interaction - INTERACT '87*. Elsevier Science: North Holland. pp. 709-714.

Freedman, R.D. & Stumpf, S.A. (1980) Learning Style Inventory: Less than meets the eye. *Academy of Management Review*, 5, 445-447.

Frese, M., Albrecht, K., Altman, A., Lang, J., Papstein, P.V., Peyerl, R., Prumper, J., Schulte-Gocking, H., Wankmuller, I., & Wendel, R. (1988) The effects of an active development of the mental model in the training process: Experimental results in a word processing system. *Behavior and Information Technology*, 7, 3, 295-304.

Frick, F.C. & Miller, G.A. (1951) A statistical description of operant conditioning. *American Journal of Pyschology*, 64, 1, 20-36.

Frisse, M.F. & Cousins, S.B. (1992) Models for hypertext. *Journal of the American Society for Information Science*, 43, 2, 183-191.

Fullerton, A.M. (1983) Age differences in the use of imagery in integrating new and old information in memory. *Journal of Gerontology*, 38, 3, 326-332.

Fulton, M.A. (1985) A research model for studying gender/power aspects of human-computer interaction. *International Journal of Man Machine Studies*, 23, 369-382.

Galotti, K.M., Baron, J., & Sabini, J.P. (1986) Individual differences in syllogistic reasoning: Deduction rules or mental models? *Journal of Experimental Psychology: General*, 115, 1, 16-25.

Garfein, A.J., Schaie, K.W., & Willis, S.L. (1988) Microcomputer proficiency in later-middle-aged and older adults: teaching old dogs new tricks. *Social Behaviour*, 3, 131-148.

Gattiker, U.E. (1992) Computer skills acquisition: A review and future directions for research. *Journal of Management*, 18, 3, 547-574.

Gentner, D. & Stevens, A.L. Eds. (1983) *Mental Models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gilroy, F.D. & Desai, H.B. (1986) Computer anxiety: Sex, race and age. *International Journal of Man-Machine Studies*, 25, 711-719.

Giroux, L. & Belleau, R. (1986) What's on the menu? The influence of menu content on the selection process. *Behaviour and Information Technology*, 5, 2, 169-172.

Gist, M., Rosen, B., & Schwoerer, C. (1988) The influence of training method and trainee age on the acquisition of computer skills. *Personnel Psychology*, 41, 255-265.

Gittins, D. (1986) Icon-based human-computer interaction. *International Journal of Man-Machine Studies*, 24, 519-543.

Gomez, L.M., Egan, D.E., & Bowers, C. (1986) Learning to use a text editor: Some characteristics that predict success. *Human-Computer Interaction*, 2, 1-23.

Gomez, L.M., Egan, D.E., Wheeler, E.A., Sharma, D.K., & Gruchacz, D.K. (1983) How interface design determines who has difficulty in learning to use a text editor. In A. Janda (Ed.) *Human Factors in Computing Systems*. pp. 176-181.

Gopher, D. & Donchin, E. (1986) Workload - An examination of the concept. In K.R. Boff, L. Kaufman, & J.P. Thomas (eds.) *Handbook of Perception and Human Performance, Vol II*. Chichester: Wiley.

Gordon, S., Gustavel, J., Moore, J., & Hankey, J. (1988) The effects of hypertext on reader knowledge representation. In *Proceedings of the Human Factors Society, 32nd Annual Meeting*. pp. 296-300.

Gray, S.H. (1989) The effect of locus of control and sequence control on computerized information retrieval and retention. *Journal of Educational Computing Research*, 5(4), 459-471.

Gray, S. (1990) Using protocol analyses and drawings to study mental model construction during hypertext navigation. *International Journal of Human-Computer Interaction*, 2, 4, 359-377.

Gray, S.H., Barber, C.B., & Shasha, D. (1991) Information search with dynamic text vs paper text: An empirical comparison. *International Journal of Man-Machine Studies*, 35, 575-586.

Gray, S.H. & Shasha, D. (1989) To link or not to link? Empirical guidance to the design of nonlinear text systems. *Behavior Research Methods, Instruments and Computers*, 21, 326-333.

Green, A.J.K. & Gilhooly, K.J. (1990) Individual differences and effective learning procedures: the case of statistical computing. *International Journal of Man-Machine Studies*, 33, 97-119.

Greenberg, S. & Witten, I.H. (1985) Adaptive personilised interfaces - A question of viability, *Behaviour and Information Technology*, 4, 1, 31-45.

Greene, S.L., Devlin, S.J., Cannata, P.E., and Gomez, L.M. 1990 No IFs, ANDs, or ORs: A study of database querying. *International Journal of Man-Machine Studies*, 32, 303-326.

Greene, S.L., Gomez, L.M., & Devlin, S.J. (1986) A cognitive analysis of database query production. In *Proceedings of the Human Factors Society - 30th Annual Meeting*. pp. 9-13.

Grudin, J. & Barnard, P. (1985) When does an abreviation become a word? And related questions. In L. Borman & B. Curtis (Eds.), *Human Factors in Computing Systems II*. North Holland: Elsevier Science. pp. 121-125.

Guastello, S.J., Traut, M., & Korienek, G. (1989) Verbal versus pictorial representations of objects in a human-computer interface. *International Journal of Man-Machine Studies*, 31, 99-120.

Guilford, J.P. (1967) *The Nature of Human Intelligence*. New York: McGraw-Hill

Guilford, J.P. (1985) The structure-of-intellect model. In B.B. Wolman (ed.) *Handbook of Intelligence: Theories, Measurements, and Applications*. Chichester: Wiley. pp. 225-266.

Halasz, F.G. (1988) Reflections on notecards: Seven issues for the next generation of hypermedia systems. *Communications of the ACM*, 31,7, 836-852.

Hanisch, K.A., Kramer, A.F., Hulin, C.L., & Schumacher, R. (1988) Novice-expert differences in the cognitive representation of system features: Mental models and verbalizable knowledge. In *Proceedings of the Human Factors Society - 32nd Annual Meeting*. pp. 219-223.

Hartley, A.A., Hartley, J.T., & Johnson, S.A. (1984) The older adult as a computer user. In P.K. Robinson, J. Livingstone, & J.E. Birren (eds.) *Aging and Technological Advances*. New York: Plenum Press.

Hasher, L. & Zacks, R.T. (1979) Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 3, 356-388.

Hauptmann, A.G. & Green, B.F. (1983) A comparison of command, menu-selection and natural language computer programs. *Behaviour and Information Technology*, 2, 2, 163-178.

Heppe, D.L., Edmondson, W.H., & Spence, R. (1985) Helping both novice and advanced user in menu-driven information retrieval systems. In P. Johnson & S. Cook (eds.) *People and Computers: Designing the Interface*. Cambridge: Cambridge University Press. pp. 92-101.

Hertzog, C. & Rypma, B. (1991) Age differences in components of mental-rotation task performance. *Bulletin of the Psychonomic Society*, 29, 3, 209-212.

Hinkle, D.E., Wiersma, W., & Jurs, S.G. (1979) *Applied Statistics for the Behavioural Sciences*. Chicago: Rand McNally.

Hiscock, M. (1986) On sex differences in spatial abilities. *American Psychologist*, 41(9), 1011-1012.

Hockey, G.R.J., Briner, R.B., Tattersall, A.J., & Wiethoff, M. (1989) Assessing the impact of computer workload on operator stress: the role of system controllability. *Ergonomics*, 32,11, 1401-1418.

Hogaboam, T.W. & Pellegrino, J.W. (1978) Hunting for individual differences in cognitive processes: Verbal ability and semantic processing of pictures and words. *Memory and Cognition*, 6, 2, 189-193.

Hollands, J.G. & Merikle, P.M. (1989) Menu organization and user expertise in information search tasks. *Human Factors*, 29(5), 577-586.

Horn, J.L. (1985) Remodeling old models of intelligence. In B.B. Wolman (ed.) *Handbook of Intelligence: Theories, Measurements, and Applications.* Chichester: Wiley. pp. 267-300.

Horn, J.L. & Hofer, S.M. (1992) Major abilities and development in the adult period. In R.J. Sternberg & C.A. Berg (eds.) *Intellectual Development.* Cambridge: Cambridge University Press.

Horton, W.K. (1990) *Designing and Writing Online Documentation.* Chichester: Wiley.

Hotson, P.M. & B. Shackel (1990) Screen size and word processing : A study on the Z88 lap-top computer. In E.J. Lovesey (ed.) *Contemporary Ergonomics 1990. Proceedings of the Ergonomics Society's 1990 Annual Conference.* London: Taylor & Francis. pp. 347-352.

Howard, G.S. (1986) *Computer Anxiety and the Use of Microcomputers in Managment.* UMI Research Press.

Hudgens, G.A. & Billingsley, P.A. (1978) Sex: The missing variable in human factors research. *Human Factors,* 20, 2, 245-250.

Hunt, E. (1978) Mechanics of verbal ability. *Psychological Review,* 85, 2, 109-130.

Hunt, E. (1986) Lecture 1: The information processing approach to intelligence. In S.E. Newstead, S.H. Irvine, & P.L. Dann (eds.) *Human Assessment: Cognition and Motivation.* Dordrecht: Martinus Nijoff.

Hunt, E., Lunnborg, C., & Lewis, J. (1975) What does it mean to be high verbal? *Cognitive Psychology,* 7, 194-227.

Hutchins, E.L., Hollan, J.D., & Norman, D.A. (1986) Direct manipulation interfaces. In D.A. Norman & S.W. Draper (Eds.), *User Centered System Design - New Perspectives on Human-Computer Interaction.* Hillsdale, NJ: Lawrence Erlbaum Associates. pp. 87-124.

Igbaria, M. & Chakraba, A. (1990) Computer anxiety and attitudes toward microcomputer use. *Behavior and Information Technology,* 9(3), 229-241.

Institute for Personality and Ability Testing (1973) *Measuring Intelligence With Culture Fair Tests*. Institue for Personality and Ability Testing, USA.

Jeffries, R., Turner, A., Polson, P., & Atwood, M. (1981) The processes involved in designing software. In J. Anderson (Ed.), *Cognitive Skills and their Acquisition*. Hillsdale, N.J.:Erlbaum. pp. 255-283.

Jenkins, J.J., Russell, W.A., & Suci, G.J. (1958) An atlas of semantic profiles for 350 words. *American Journal of Psychology*, 71, 688-699.

Jennings, F, Benyon, D., & Murray, D. (1991) Adapting systems to differences between individuals. *Acta Psychologica*, 78, 243-256.

Johnson-Laird, P.N. (1983) *Mental Models*. Cambridge: Cambridge University Press.

Johnson-Laird, P.N. (1985) Mental models In A.M. Aitkinhead and J.M. Slack (eds.) *Issues in Cognitive Modelling*. Hove: Lawrence Erlbaum Associates.

Jones, P.M. & Mitchell, C.M. (1987) Operator modelling: Conceptual and methodological distinctions. In *Proceedings of the Human Factors Society - 31st Annual Meeting*. pp. 31-35.

Jorgensen, A.H. & Sauer, A. (1990) The personal touch: A study of users' customisation practice. In D. Diaper, D. Gilmore, G. Cockton, & B. Shackel (eds.) *Proceedings of the IFIP TC 13 Third International Conference on Human-Computer Interaction, INTERACT '90*, Cambridge, U.K., 27-31 August. Amsterdam: North Holland. pp. 561-565.

Jung, C.G. (1971) *Psychological Types*. Princeton, NJ: Priceton University Press.

Kacmar, C.J. & Carey, J.M. (1991) Assessing the usability of icons in user interfaces. *Behaviour & Information Technology*, 10, 6, 443-458.

Kagan, D.M. & Douthat, J.M. (1985) Personality and learning FORTRAN. *International Journal of Man-Machine Studies*, 22, 395-402.

Kahneman, D. (1973) *Attention and Effort*. New Jersey: Prentice-Hall.

Kantowitz, B.H. & Weldon, M. (1985) On scaling performance operating characteristics: Caveat emptor. *Human Factors*, 27, 5, 531-547.

Kantorowitz, E. & Sudarsky, O. (1989) The adaptable user interface. *Communications of the ACM*, 32, 11, 1352-1358.

Kass, R. & Finin, T. (1988) A general user modelling facility. In *Conference on Human Factors in Computing Systems, CHI '88*, May, 1988, Washington DC. ACM, SIGCHI. pp. 145-150.

Kennedy, T.C.S. (1975) Some behavioural factors affecting the training of niave users of an interactive computer system. *International Journal of Man-Machine Studies*, 7, 817-834.

Kerr, S.T. (1990) Wayfinding in an electronic database: The relative importance of navigational cues vs. mental models. *Information Processing and Management*, 26, 4, 511-523.

Key, M.R. (1975) *Male/Female Language*. Metuchen, N.J.: Scarecrow Press.

Kieras, D. & Polson, P.G. (1985) An approach to the formal analysis of user complexity. *International Journal of Man-Machine Studies*, 22, 365-394.

Kline, P. (1993) *The Handbook of Psychological Testing*. London: Routledge.

Klix, F. (1986) Memory research and knowledge engineering. *Man-Computer Interaction Research. MACINTER I*. F. Klix & H. Wandke (eds.). North Holland: Elsevier Science. pp. 97-116.

Kogan, N. & Wallach, M.A. (1964) *Risk Taking: A Study in Cognition and Personality*. NY: Rinehart and Winston.

Kolb, D.A. (1971) Individual learning styles and the learning process. Sloan School Working Paper 5-571. Cambridge, MA: Massachusetts Institute of Technology .

Kolb, D.A. (1981) Experiential learning theory and the Learning Style Inventory: A reply to Freedman and Stumpf. *Academy of Mangement Review*, 6, 289-296.

Kolers, P.A., Duchnicky, R.L., & Ferguson, D.C. (1981) Eye movement measurement of readability of CRT displays. *Human Factors*, 23, 5, 517-527.

Kosslyn, S.M. (1980) *Image and Mind*. Cambridge Mass.: Harvard University Press

Koubeck, R.J., LeBold, W.K., & Salvendy, G. (1985) Predicting performance in computer programming courses. *Behaviour and Information Technology*, 4,2, 113-129.

Koubeck, R.J. & Salvendy, G. (1991) Cognitive performance of super-experts on computer program modification tasks. *Ergonimics*, 34, 8, 1095-1112.

Koved, L. & Shneiderman, B. (1986) Embedded menus: Menu selection in context. *Communications of the ACM*, 29, 312-318.

Kunkel, K. & Strothotte, T. (1988) Visualization and direct manipulation in user interfaces: Are we overdoing it? In P. Gorny & M.J. Tauber (eds.) *Visualization in Human-Computer Interaction. 7th Interdisciplinary Workshop on Informatics and Psychology*, Sharding, Austria. 24-27th May, 1988. NY: Springer-Verlag. pp. 181-193.

Kyllonen, P.C. & Christal, R.E. (1990) Reasoning ability is (little more than) working memory capacity?! *Intelligence*, 14, 389-433.

Kyllonen, P.C., Lohman, D.F. & Snow, R.E. (1984) Effects of aptitudes, strategy training, and task facets on spatial task performance. *Journal of Educational Psychology*, 76, 130-145.

Kyllonen, P.C., Lohman, D.F., & Woltz, D.J. (1984) Componential modelling of alternative strategies for performing spatial tasks. *Journal of Educational Psychology*, 76, 1325-1345.

Kyllonen, P.C., Woltz, D.J., & Lohman, D.F. (1981) Models of strategy and strategy-shifting in spatial visualisation performance. (Technical Report No. 17) Aptitude Research Project, School of Education, Stanford University.

Lansdale, M.W. (1988) On the memorability of icons in an information retrieval task. *Behaviour and Information Technology*, 7(2), 131-151.

Lansdale, M.W., Jones, M.R., & Jones, M.A. (1989) Visual search in verbal and iconic interfaces. In E.D. Megaw (Ed.), *Contemporary Ergonomics 1989*. London: Taylor & Francis. pp. 422-429.

Lansdale, M.W., Simpson, M., & Stroud, T. (1987) Comparing words and icons as cue enrichers in an information retrieval task. In *Proceedings of the Conference on Human-Computer Interaction, INTERACT '87*. Amsterdam: Elsevier Science. pp. 911-916.

Lansdale, M.W., Simpson, M. and Stroud, T.R. (1990) A comparison of words and icons as external memory aids in an information retrieval task. *Behaviour and Information Technology*, 9, 2, 111-131.

Lee, E., & MacGregor, J. (1985) Minimising user search time in menu retrieval systems. *Human Factors*, 27, 2, 157-162.

Levinson, E.M. (1986) A review of the Computer Aptitude, Literacy, and Interest. Profile.*Journal of Counselling and Development*, 64 (June), 658-659.

Liebelt, L.S., McDonald, J.E., Stone, J.D., & Karat, J. (1982) The effect of organization on learning menu access. In *Proceedings of the Human Factors Society - 26th Annual Meeting*.

Logan, E. (1990) Cognitive styles and online behavior of novice searchers. *Information Processing and Management*, 26, 4, 503-510.

Lohman, D.F. (1979) Spatial ability: A review and reanalysis of the correlational literature. Stanford University, CA, School of Education: Report on Aptitude Research Project.

Lohman, D.F. & Kyllonen, P.C. (1983) Individual differences in solution strategy on spatial tasks. In R.F. Dillon & R.R. Schmeck (Eds.) *Individual Differences in Cognition Vol. I*. pp. 105-135.

Loo, R. (1978) Individual differences and the perception of traffic signs. *Human Factors*, 20, 65-74.

Loyd, B.H. & Gressard, C. (1984) The effects of sex, age, and computer experience on computer attitudes. *AEDS Journal*, pp. 67-76.

Lyons, M.L. (1985) The dp psyche, *Datamation*, 31, 16, 103-110.

Macaulay, L.A. & Norman, M.A. (1984) Designing interfaces for different types of users - experimental criteria. *INTERACT '84*, First IFIP Conference on Human-Computer Interaction. pp. 643-647.

MacGreggor, J., Lee, E., Lam, N. (1986) Optimising the structure of database menu indexes: A decision model of menu search. *Human Factors*, 28(4), 387-399.

MacGregor, J.N. (1992) A comparison of the effects of icons and descriptors in videotex menu retrieval. *International Journal of Man-Machine Studies*, 37, 767-777.

Mack, R.L., Lewis, C.H., & Carroll, J.M. (1987) Learning to use word processors: Problems and prospects. In R.M. Baecker & W.A.S. Buxton (Eds.), *Readings in Human Computer Interaction - A Multidisciplinary Approach*. San Mateo, Ca.: Morgan Kaufman. pp. 269.

Maguire, M.C. (1990) A review of human factors guidelines and techniques for the design of graphical human-computer interfaces. In J.Preece & L. Keller (Eds.), *Human-Computer Interaction*. Hemel Hempstead: Prentice Hall.

Mani, K. & Johnson-Laird, P.N. (1982) The mental representation of spatial descriptions. *Memory and Cognition*, 10, 2, 181-187.

Marchionini, G. & Schneiderman, B. (1988) Finding facts vs. browsing knowledge in hypertext systems. *Computer*, Jan. 1988, pp. 70-80.

Mariani, J.A. & Lougher, R. (1992) Triplespace: An experiment in a 3D graphical interface to a binary relational database. *Interacting with Computers*, 4, 2, 147-162.

Martin, M.P. & Fuerst, W.L. (1987) Using computer knowledge in the design of interactive systems. *International Journal of Man-Machine Studies*, 26, 333-342.

Martin, M.P. & Fuerst, W.L. (1988) Effect of computer knowledge on user performance over time. *Information and Software Technology*, 30, 9, 561-565.

Matta, K.F. & Kern, G.M. (1991) Interactive videodisc instruction: the influence of personality on learning. *International Journal of Man-Machine Studies*, 35, 541-552.

Mazlack, L.J. (1980) Identifying the potential to aquire programmer skills. *Communications of the ACM*, 23, 14-17.

McDonald, J.E., Stone, J.D., & Liebelt, L.S. (1983) Searching for items in menus: The effects of organization and type of target. *In Proceedings of the Human Factors Society - 27th Annual Meeting*.

McDowd, J.M. & Craik, F.I.M. (1988) Effects of aging and task difficulty on divided attention performance. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 2, 267-280.

McGee, M.G. (1979) Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin*, 86,5,889-918.

McKeithen, K.B., Reitman, J.S., Rueter, H.H., & Hirtle, S.C. (1981) Knowledge organisation and skill differences in computer programmers. *Cognitive Psychology*, 13, 307-325.

McKenna, F.P. (1984) Measures of field dependence: Cognitive style or cognitive ability? *Journal of Personality and Social Psychology*, 47, 593-603.

McKnight, C., Dillon, A., & Richardson, J. (1989) A comparison of linear and hypertext formats in information retrieval. In *HYPERTEXT II, the Second Annual UK Conference on Hypertext*, York, July. pp. 10-19.

McLeod, C.M., Hunt, E.B., & Matthews, N.N. (1978) Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior*, 5, 493-508.

Meade, A.C. (1987) How many spatial abilities are there? A new answer from an old technique, factor analysis. Unpublished doctoral dissertation: University of North Carolina, Chapel Hill.

Meister, D. (1989) Lost in computer space. *International Journal of Human-Computer Interaction*, 1, 1, 5-21.

Merwin, D.H., Dyre, B.P., Humphrey, D.G., Grimes, J., & Larish, J.F. (1990) The impact of icons and visual effects on learning computer databases. In *Proceedings of the Human Factors Society - 34th Annual Meeting*. pp. 424-428.

Messick, S. (1984) The nature of cognitive styles: Problems and promise in educational practice. *Educational Psychologist*, 19, 59-74.

Microsoft (1992) The way we work. *Microsoft Magazine*, 2, 4, 24-27.

Miller, G.A. & Frick, F.C. (1949) Statistical Behavioristics and sequences of responses. *Psychological Review*, 56, 311-324.

Mitchell, D.B. (1989) How many memory systems? Evidence from aging. *Journal of Experimental Psychology: Leaning, Memory, and Cognition*, 15, 1, 31-49.

Mitchell, D.B. & Perlmutter, M. (1986) Semantic activation and episodic memory: Age similarities and differences. *Developmental Psychology*, 22, 1, 86-94.

Mitrushina, M. & Satz, P. (1989) Differential decline of specific memory components in normal aging. *Brain Dysfunction*, 2, 330-335.

Mohageg, M.F. (1992) The influence of hypertext linking structures on the efficiency of information retrieval. *Human Factors*, 34, 3, 351-367.

Monk, A., Walsh, P., & Dix, A. (1988) A comparison of hypertext, scrolling and folding as mechanisms for program browsing. In D. Jones and R. Winder (eds.) *People and Computers IV*. Cambridge: Cambridge University Press.

Moran, T.P. (1981) The Command Language Grammar: A representation for the user interface of interactive computer systems. *International Journal of Man-Machine Studies*, 15, 3-50.

Moray, N. (1967) Where is capacity limited? A survey and a model. *Acta Psychologica*, 27, 84-92.

Morra, S. (1989) Developmental differences in the use of verbatim versus spatial representations in the recall of spatial descriptions: A probabilistic model and an experimental analysis. *Journal of Memory and Language*, 28, 37-55.

Morrill-Tazelar, J. (1990) Object Lessons. *Byte*, 15, 10, 6.

Morrison, P.R. & Noble, G. (1987) Individual differences and ergonomic factors in performance on a videotex-type task. *Behavior and Information Technology*, 6, 1, 69-88.

Morrow, P.C., Prell, E.R., & McElroy, J.C. (1986) Attitudinal and behavioural correlates of computer anxiety. *Psychological Reports*, 59, 1199-1204.

Mumaw, R.J., Pellegrino, J.W., & Glaser, R. (1980) Some puzzling aspects of spatial abiliity. Paper presented at the Annual Meeting of the Psychonomic Society, St Louis, MO. November, 1980.

Mumaw, R.J. & Pellegrino, J.W. (1984) Individual differences in complex spatial processing. *Journal of Educational Psychology*, 76, 920-939.

Murray, D. (1991) Modelling for adaptivity. In M.J. Tauber and D. Ackermann (eds.) *Mental Models in Human-Computer Interaction II*. North Holland: Elsevier Science. pp. 81-95.

Murray, J.B. (1990) Review of research on the Myers-Briggs Type Indicator. *Perceptual and Motor Skills*, 70, 3, 1187-1202.

Muter, P. & Mayson, C. (1986) The role of graphics in item selection from menus. *Behaviour and Information Technology*, 5, 1, 89-95.

Myers, I.B. & McCaulley, M.H. (1985) *Manual: a guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.

Navon, D. & Gopher, G. (1979) On the economy of the human-processing system. *Psychological Review*, 86, 3, 214-255.

Neal, A.S. & Darnell, M.J. (1984) Text-editing performance with partial-line, partial-page, and full-page displays. *Human Factors*, 26, 4, 431-441.

Neal, L.R. (1987) Cognition-sensitive design and user modelling for syntax-directed editors. In *Conference on Human Factors in Computing Systems, CHI '87*. pp. 99-102.

Nelson-Denny Reading Test (1973) Boston: Riverside Pub. Co.

Neuman, G.A. & Nomoto, J.T. (1990) Personnel selection tests for computer professionals and support technicians. *Journal of Business and Psychology*, 5, 2, 165-177.

Newell, A. & Rosenbloom, P.S. (1981) Mechanisms of skill acquisition and the law of practice. In J.R. Anderson (ed.) *Cognitive Skills and Their Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Nielsen, J. (1989) The matters that really matter for hypertext usability. *Hypertext '89 Proceedings*, 5-8 November, 1989. ACM. pp. 239-248.

Nielsen, J. (1990) The art of navigating through hypertext. *Communications of the ACM*, 33, 3, 296-310.

Nielsen, J. (1990) Miniatures versus icons as a visual cache for videotex browsing. *Behaviour and Information Technology*, 9, 6, 441-449.

Nielsen, J. (1992) The usability engineering life cycle. *Computer*, March, 1992.

Norcio, A.F., & Stanley, J. (1989) Adaptive human-computer interfaces: A literature survey and perspective. *IEEE Transactions on Systems, Man, and Cybernetics*. 19, 2, 399-408.

Norman, D.A. (1983) Design principles for human-computer interfaces. In A. Janda (Ed.) *Proceedings of the CHI '83 Conference on Human Factors in Computing Systems*. NY: ACM.

Norman, D.A. (1986) Cognitive engineering. In D.A. Norman & S.W. Draper (Eds.), *User Centered System Design - New Perspectives on Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Norman, D.A. & Bobrow, D.G. (1975) On data-limited and resource-limited processes. *Cognitive Psychology*, 7, 44-64.

Norman, K.L. and Chin, J.P. (1988) The effect of tree structure on search in a hierarchical menu selection system. *Behaviour and Information Technology*, 7(1), 51-65.

Norman, K.L., Weldon, L.J., & Shneiderman, B. (1986) Cognitive layouts of windows and multiple screens for user interfaces. *International Journal of Man-Machine Studies*, 25, 229-248.

Oakhill, J.V. & Johnson-Laird, P. (1984) Representation of spatial descriptions in working memory. *Current Psychological Research and Reviews*, 3, 52-62.

O'Connor, J. (1972) Developmental changes in abstractness and moral reasoning. Dissertation Abstracts International, 32, 4109a.

O'Donnell, R.D. & Eggemeier, F.T. (1986) Workload assessment methodology. In K.R. Boff, L. Kaufman, & J.P. Thomas (eds.) *Handbook of Perception and Human Performance, Vol II*. Chichester: Wiley.

Olfman, L. (1987) A comparison of applications-based and construct-based training methods for DSS generator software. Unpublished Doctoral Dissertation. Indiana University, Graduate School of Business, Bloomington, IN.

Olson, J.R. (1985) CHI Panel Discussion: Expanded design procedures learnable usable interfaces. In L. Borman & B. Curtis (Eds.) *Human Factors In Computing Systems II*. Elsevier Science: North Holland. pp. 142-143.

Oltman, P.K., Ruskin, E., & Witkin, H.A. (1971) *Group Embedded Figures Test*. Palo Alto: Consulting Psychologists Press.

O'Malley, C. & Draper, S. (1992) Representation and interaction: Are mental models all in the mind? In Y. Rogers, A. Rutherford, & P.A. Bibby (eds.) *Models in the Mind*. London: Academic Press.

Paap, K.R. & Roske-Hopstrand, R.J. (1988) Desing of menus. In M. Hellander (ed.) *Handbook of Human-Computer Interaction*. North Holland: Elsevier Science.

Paivio, A. (1971) *Imagery and Verbal Processes*. New York: Holt, Rinehart, and Winston. (Reprinted by Lawrence Erlbaum Associates, Hillsdale, NJ. 1979).

Paivio, A. (1975) Perceptual comparisons through the mind's eye. *Memory & Cognition*, 3, 6, 635-647.

Paivio, A. (1991) *Images in Mind*. Hemel Hempstead: Harvester Wheatsheaf.

Paivio, A., Clark, J.M., & Khan, M. (1988) Effects of concreteness and semantic relatedness on composite imagery ratings and cued recall. *Memory & Cognition*, 16, 5, 422-430.

Parasuraman, S. & Igbaria, M. (1990) An examination of gender differences in the determinants of computer anxiety and attitudes toward microcomputers among managers. *International Journal of Man-Machine Studies*, 32, 3, 327-340.

Parker, J.F. & Fleischman, E.A. (1960) Ability factors and component performance measures as predictors of complex tracking behaviour. *Psychological Monographs*, 74, No. 16, (Whole No. 503): 1-36.

Parton, D., Huffman, Pridgen, Norman, & Shniederman (1985) Learning a menu selection tree: Training methods compared. *Behaviour and Information Technology*, 4, 2, 87-91.

Pask, G. (1977) *Learning Styles, Educational Strategies, and Representation of Knowledge: Methods and Applications*. Richmond: Systems Research Ltd.

Paxton, A.L. & Turner, E.J. (1984) The application of human factors to the needs of the novice computer user. *International Journal of Man-Machine Studies*, 20, 137-156.

Payne, S.J. & Green, T.R.G. (1986) Task-action grammar: A model of the mental representation of task languages. *Human-Computer Interaction*, 2, 93-133.

Pedhazur, E.J. (1982) *Multiple Regression in Behavioral Research.* Fort Worth, TX: Harcourt Brace Jovanovich, Inc.

Peio, K.J. & Klein, H.A. (1984) Skilled performance: Computer software. In *Proceedings of the Human Factors Society - 28th Annual Meeting.*

Pellegrino, J.W., Rosinski, R.R., Chiesi, H.L., & Siegel, A. (1977) Picture-word differrences in decision latency: An analysis of single and dual memory models. *Memory & Cognition*, 5, 4, 383-396.

Pepin, M., Beaulieu, R., Matte, R., & Leroux, Y. (1985) Microcomputer games and sex-related differences: Spatial, verbal, and mathematical abilities. *Psychological Reports*, 56(3), 783-786.

Peverett, T. (1992) The friendlier face of computing. *Informatics*, Feb, 1992.

Polson, P.G. & Keiras, D.E. (1985) A quantitive model of the learning and performance of text editing knowledge. In L. Borman & B. Curtis (Eds.), *Human Factors in Computing Systems II.* North Holland: Elsevier Science. pp. 207-212.

Poplin, M.S., Drew, D.E., & Gable, R.S. (1984) *Computer Aptitude, Literacy, & Interest Profile.* Austin, Texas: Pro-Ed.

Posner, M.I., Boies, S.J., Eichelman, W.H., & Taylor, R.L. (1969) Retention of visual and name codes of single letters. *Journal of Experimental Psychology (Monograph)*, 79, 1-16.

Potosnak, K.M. (1984) *Choice of interface modes by empirically derived categories of users.* Unpublished Doctoral Dissertation, The John Hopkins University. Baltimore, MD. (Can be obtained from University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106.

Potosnak, K.M., Hayes, P.J., Rosson, M.B., Scheider, M.L. & Whiteside, J.A. (1986) Classifying users: A hard look at some controversial issues (Panel). In *Proceedings of CHI '86 Conference.* pp. 84-88.

Prumper, J., Frese, M., Zapf, D., & Brodbeck, F.C. (1991) Errors in computerized office work: Differences between novice and expert users. *SIGCHI Bulletin*, 23, 2, pp. 63-66.

Prumper, J., Zapf, D., Brodbeck, F.C., & Frese, M. (1992) Some surprising differences between novice and expert errors in computerised office work. *Behaviour and Information Technology*, 11, 6, 319-328.

Pylyshyn, Z. (1973) What the mind's eye tells the mind's brain. *Psychological Bulletin*, 80, 1-24.

Ramamurthy, K., King, W.R., & Premkumar, G. (1992) User characteristics - DSS effectiveness linkage: an empirical assessment. *International Journal of Man-Machine Studies*, 36, 469-505.

Rasmussen, J. (1990) Mental models and the control of actions in complex environments. In D. Ackermann & M.J. Tauber (Eds.) *Mental Models in Human-Computer Interaction, Vol. I.* Amsterdam: North Holland. pp. 41-72.

Raub, A.C. (1981) Correlates of computer anxiety in college students. Unpublished PhD. Thesis. University of Pennsylvania

Rauterberg, M. (1992) An empirical comparison of menu-selection (CUI) and desktop (GUI) computer programs carried out by beginers and experts. *Behaviour & Information Technology*, 11, 4, 227-236.

Ray, N.M. & Minch, R.P. (1990) Computer anxiety and alienation: Toward a definitive and parsimonious measure. *Human Factors*, 32, 4, 477-491.

Ribeaux, P. & Poppleton, S.E. (1978) *Psychology and Work.* London: Macmillan.

Rich, E. (1983) Users are individuals: Individualizing user models. *International Journal of Man-Machine Studies*, 18, 199-214.

Roberts, T.L. (1980) Evaluation of computer text editors. Unpublished PhD. Dissertation, Stanford University.

Roberts, T.L. & Moran, T.P. (1983) The evaluation of text editors: Methodology and empirical results. *Communications of the ACM*, 26, 4, 265-283.

Robertson, I.T. (1985) Human information-processing strategy and style. *Behaviour and Information Technology*, 18, 199-214.

Rogers, Y. (1989) Icon design for the user interface. *International Review of Ergonomics 2*. pp. 129-154.

Rogers, Y. & Osborne, D.J. (1987) Pictorial communication of abstract verbs in relation to human-computer interaction. *British Journal of Psychology*, 78, 99-112.

Rohr, G. (1986) Using visual concepts. In Chang, S-K., Ichikawa, T., & Ligomenides, P.A. (Eds.) *Visual Languages*. New York: Plenum Press. pp. 325-348.

Romary, L., Carbonell, N., & Pierrel, J-M. (1991) Multi-modal human-computer dialogue. In Y. Queinnec & F. Daniellou (Eds.). *Designing for Everyone. Proceedings of the Eleventh Congress of the International Ergonomics Association*. Paris. 1991. London: Taylor & Francis. pp. 739-741.

Rosenbaum, S. (1989) Selecting appropriate subjects for documentation usability testing. In M.J. Smith & G. Salvendy (Eds.) *Work with Computers: Organizational, Management, Stress and Health Aspects*. Amsterdam: Elsevier Science. pp. 620-627.

Ross, B.H. & Moran, T.P. (1983) Remindings and their effects in learning a text editor. In A. Janda (Ed.), *Human Factors in Computing Systems*. New York: ACM. pp. 222-225.

Rosson, M.B. (1983) Patterns of experience in text editing. In A. Janda (Ed.), *Human Factors in Computing Systems*. New York: ACM. pp. 171-175.

Rosson, M.B. (1984a) Effects of experience on learning, using, and evaluating a text editor. *Human Factors*, 26,4, 463-475.

Rosson, M.B. (1984b) The role of experience in editing. In B. Shackel (Ed.), *Human Computer Interaction - INTERACT '84*. North Holland: Elsevier Science. pp. 45-50.

Roth, P.L. & Campion, J.E. (1992) An analysis of the predictive power of the panel interview and pre-employment tests. *Journal of Occupational Psychology*, 65, 51-60.

Rotter, J.B. (1966) Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80, 1, Whole No. 609.

Rouse, W.B. & Morris, N.M. (1986) On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, 100, 3, 349-363.

Ruch, F.L. & Ruch, W.L. (1957) *The Employee Aptitude Survey*. Los Angeles, CA: Psychological Services Inc..

Rutherford, A. & Wilson, J.R. (1991) Models of mental models: An ergonomist - psychologist dialogue. In M.J. Tauber & D. Ackermann (eds.) *Mental Models and Human-Computer Interaction II*. North Holland: Elsevier Science.

Salthouse, T.A. (1984) Effects of age and skill in typing. *Journal of Experimental Psychology: General*, 113, 3, 345-371.

Salthouse, T.A. (1987) The role of representations in age differences in analogical reasoning. *Psychology and Aging*, 2, 4, 357-362.

Salthouse, T.A. (1991a) Expertise as the circumvention of human processing limitations. In K.A. Ericsson & J. Smith (eds.) *Towards a General Theory of Expertise*. Cambridge: Cambridge University Press. pp. 286-300.

Salthouse, T.A. (1991b) Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, 2, 3, 179-183.

Salthouse, T.A. (1991c) *Theoretical Perspectives on Cogntive Aging*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Salthouse, T.A. (1992) Working memory mediation of adult age differences in integrative reaoning. *Memory & Cognition*, 20, 4, 413-423.

Salthouse, T.A. & Babcock, R.L. (1991) Decomposing adult age differences in working memory. *Developmental Psychology*, 27, 5, 763-776.

Salthouse, T.A., Babcock, R.L., & Shaw, R.J. (1991) Effects of adult age on structural and operational capacities in working memory. *Psychology and Aging*, 6, 1, 118-127.

Salthouse, T.A., Babcock, R.L., Skovronek, E., Mitchell, D.R.D., & Palmon, R. (1990) Age and experience effects in spatial visualization. *Developmental Psychology*, 26,1,128-136.

Salthouse, T.A., Davenport-Rogan, J. & Prill, K. (1984) Division of attention: Age differences on a visually presented memory task. *Memory and Cognition*, 12, 6, 613-620.

Salthouse, T.A., Kausler, D.H., & Saults, J.S. (1988) Utilization of path-analytic procedures to investigate the role of processing resources in cognitive aging. *Psychology and Aging*, 3, 158-166.

Salthouse, T.A. & Mitchell, D.R.D. (1990) Effects of age and naturally occuring experience on spatial visualization performance. *Developmental Psychology*, 26, 5, 845-854.

Salthouse, T.A., Mitchell, D.R.D., Skovronek, E., & Babcock, R.L. (1989) Effects of adult age and working memory on reasoning and spatial abilities. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 15, 507-516.

Salthouse, T.A. & Prill, K.A. (1987) Inferences about age impairments in inferential reasoning. *Psychology and Aging*, 2, 1, 43-51.

Salthouse, T.A. & Skovronek, E. (1992) Within-context assessments of age differences in working memory. *Journal of Gerontology*, 47, 3, 110-120.

Salthouse, T.A. & Somberg, B.L. (1982) Skilled performance: Effects of adult age and experience on elementary processes. *Journal of Experimental Psychology: General*, 111, 176-207.

Sanders, B. Cohen, M.R. & Soares, M.P. (1986) The sex difference in spatial ability: A rejoinder. *American Psychologist*, 41, 9, 1015-1016.

Sandman, R.S. (1979) *Mathematics Attitude Inventory*. Minnesota: Research and Evaluation Center, University of Minnesota.

Santhanam, R. & Wiedenbeck, S. (1993) Neither novice nor expert: The discretionary user of software. *International Journal of Man-Machine Studies*, 38, 201-229.

Saracevic, T. & Kantor, P. (1988) A study of information seeking and retrieving. III Searches, searchers, and overlap. *Journal of the American Society for Information Science*, 39, 197-216.

Schneider, W. & Shiffrin, R.M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1, 1-66.

Schvaneveldt, R.W., Durso, F.T., Goldsmith, T.E., Breen, T.J., & Cooke, N.M. (1985) Measuring the structure of expertise. *International Journal of Man-Machine Studies*, 23, 699-728.

Schwarz, E., Beldie, I.P., & Pastoor, S. (1983) A comparison of paging and scrolling for changing screen contents by inexperienced users. *Human Factors*, 25, 3, 279-282.

Scott, D. & Findlay, J.M. (1991) Visual search and VDUs. In D. Brogan (Ed.), *Visual Search II*. London: Taylor & Francis.

Seagull, F.J. & Walker, N. (1992) The effects of hierarchical structure and visualization ability on computerized information retrieval. *International Journal of Human-Computer Interaction*, 4, 4, 369-385.

Sebrechts, M.M., Deck, J.G., & Black, J.B. (1983) A diagrammatic approach to computer instruction for the niave user. *Behavior Research Methods & Instrumentation*, 15, 2, 200-207.

Sebrechts, M.M., Deck, J.G., Wagner, R.K., & Black, J.B. (1984) How human abilities affect component skills in word processing. *Behavior Research Methods, Instruments and Computers*, 16, 2, 234-237.

Sein, M.K. & Bostrom, P. (1989) Individual differences and conceptual models in training novice users. *Human-Computer Interaction*, 4(3), 197-229.

Sein, M.K., Bostrom, R.P. & Olfman, L. (1987) Conceptual models in training novice users. In H-J. Bullinger & B. Shackel (eds.). *Human-Computer Interaction - INTERACT '87*. Amsterdam: Elsevier Science. pp. 861-867.

Shackel, B (1990) Human Factors and Usability. In J.Preece & L. Keller (Eds.), *Human-Computer Interaction*. Hemel Hempstead: Prentice Hall.

Shackleton, V. (1989) *How to Pick People for Jobs*. Glasgow: Collins & Sons.

Shackleton, V. & Newell, S. (1991) Management selection: A comparative survey of methods used in top British and French companies. *Journal of Occupational Psychology*, 64, 23-36.

Shepard, R. (1967) Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behaviour*, 6, 156-163.

Sheppard, B. (1987) High-performance software analysis on the IBM PC - Examine routine execution times with this high-resolution timer. *Byte*, January, pp. 157-164.

Shneiderman, B. (1987a) *Designing the User Interface*. Reading, MA: Addison-Wesley.

Shneiderman, B. (1987b) Direct manipulation: A step beyond programming languages (excerpt). In R.M. Baecker & W.A.S. Buxton (Eds.), *Readings in Human Computer Interaction - A Multidisciplinary Approach*. San Mateo, Ca.: Morgan Kaufman. pp. 461-468.

Shneiderman, B. (1987c) User interface design and evaluation for an electronic encyclopedia. In G. Salvendy (Ed.) *Cognitive Engineering in the Design of Human Computer Interaction and Expert Systems*. Amsterdam: Elsevier Science. pp. 207-223.

Shneiderman, B. (1988) We can design better user interfaces: A review of human-computer interaction styles. *Ergonomics*, 31, 5, 699-710.

Shu, N.C. (1986) Visual programming languages: A perspective and a dimensional analysis In S-K. Chang, T. Ichikawa, & P.A. Ligomenides (eds.) *Visual Languages*. NY: Plenum Press.

Shurtleff, M.S., Jenkins, J.A., & Sams, M.R. (1988) Deriving menu structures through modal block clustering: A promising alternative to hierarchical techniques. In *Proceedings of the Human Factors Society. 32nd Annual Meeting*. pp. 347-351.

Shute, V.J. (1991) Who is likely to acquire programming skills? *Journal of Educational Computing Research*, 7, 1, 1-24.

Simon, H.A. & Chase, W.G. (1973) Skill in chess. *American Scientist*, 61, 394-403.

Simpson, A. (1990) The interface to a hypertext journal. In D. Daiper, D. Gilmore, G. Cockton, & B. Shackel (eds.) *INTERACT '90. Proceedings of the IFIP TC 13 Third International Conferernce on Human-Computer Interaction*, 27-31 August. Cambridge, UK. North Hollan: Amsterdam. pp. 869-874.

Singley, M.E. & Anderson, J.R. (1985) The transfer of text-editing skill. *International Journal of Man-Machine Studies*, 22, 403-423.

Sitton, S. & Chmelir, G. (1984) The intuitive computer programmer. *Datamation*, 30, 15, 137-140.

Sloboda, J. (1991) Musical expertise. In K.A. Ericsson and J. Smith (eds.) *Toward a General Theory of Expertise*. Cambridge: Cambridge University Press.

Smith, D.C., Irby, C., Kimball, R., Verplank, B., & Harslem, E. (1990) Designing the star user interface. In J. Preece & L. Keller (eds.) *Human-Computer Interaction*. Hemel Hempstead: Prentice Hall. (first published in Byte, 1982) pp. 237-259.

Smith, J.B. & Weiss, S.F. (1988) Hypertext. *Communications of the ACM*, 31, 7, 816-819.

Smyth, M.M. & Pelky, P.L. (1992) Short-term retention of spatial information *British Journal of Psychology*, 83, 3, 359-376.

Snowberry, K., Parkinson, S., & Sisson, N. (1983) Computer display menus. *Ergonomics*, 26, 699-712.

Soderston, C. (1986) A study of spatial models and human navigation within complex computer interfaces. Unpublished doctoral dissertation: Rensselaer Polytechnic Institute.

Somberg, B.L. & Picardi, M.C. (1983) Locus of the information familiarity effect in the search of computer menus. In *Proceedings of the Human Factors Society - 27th Annual Meeting.*

Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R., & Jacobs, G.A. (1983) *Manual for the State-Trait Anxiety Inventory*. Palo Alto: Consulting Psychologists Press.

Spinas, P. (1989) User oriented software development and dialogue design.In M.J. Smith & G. Salvendy (Eds.) *Work with Computers: Organizational, Management, Stress and Health Aspects*. Amsterdam: Elsevier Science. pp. 200-207.

Stammers, R.B., George, D.A., & Carey, M.S. (1989) An evaluation of abstract and concrete icons for a CAD package. In E.D. Megaw (ed.) *Contemporary Ergonomics*. London: Taylor and Francis. pp. 416-421.

Stanton, N.A. & Stammers, R.B. (1989) Learning styles in a non-linear training environment. In *HYPERTEXT II*, the Second Annual UK Conference on Hypertext, York, July.pp. 114-120.

Sternberg, R.J. (1977) *Intelligence, Information processing, and Analogical Reasoning: The Componential Analysis of Human Abilities*. Hillsdale, NJ: Lawrence Erlbaum.

Sternberg, R.J. (1985) Cognitive approaches to intelligence. In B.B. Wolman (ed.) *Handbook of Intelligence: Theories, Measurements, and Applications*. Chichester: Wiley. pp. 59-118.

Sternberg, R.J. & Frensch, P.A. (1990) Intelligence and Cognition. In M.W. Eysenk (Ed.) *Cognitive Psychology: An International Review*. Chichester: Wiley. pp. 57-103.

Sternberg, R.J. & Gardner, M.K. (1983) Unities in inductive reasoning. *Journal of Experimental Psychology: General*, 112, 80-116.

Stevens, D.J. (1983) Cognitive processes and success of students in instructional computer courses. *AEDS Journal*, 16, 228-233.

Streeter, L.A. & Vitello, D. (1986) A profile of drivers' map reading abilities. *Human Factors*, 28, 2, 223-239.

Streeter, L.A., Vitello, D., & Wonsiewicz, S.A. (1985) How to tell people where to go: Comparing navigational aids. *International Journal of Man-Machine Studies*, 22, 549-562.

Streitz, N.A. (1987) Mental models and metaphors: Implications for the design of adaptive user-system interfaces. In H. Mandl & A. Lesgold (Eds) *Learning Issues for Intelligent Tutoring Systems. Cognitive Science Series*. New York: Springer. pp. 164-186.

Streitz, N.A., Spijkers, W.A.C., & van Duren, L.L. (1987) From novice to expert user: A transfer of learning experiment on different interaction modes. In H-J. Bullinger & B. Shackel (eds.) *Human-Computer Interaction - INTERACT '87*. North Holland: Elsevier Science. pp. 841-846.

Stuart-Hamilton, I. (1991) *The Psychology of Aging*. London: Jessica Kingsley.

Stumpf, S.A. & Freedman, R.D. (1981) The Learning Style Inventory: Still less than meets the eye. *Academy of Management Review*, 6, 297-299.

Swierenga, S.J. (1990) Menuing and scolling as alternative information access techniques. In *Proceedings of the Human Factors Society - 34th Annual Meeting*. pp. 356-359.

Taylor, R.G. & Hinson, N. (1988) Individual differences in the use of a pointing device. In E.D. Megaw (ed.) *Contemporary Ergonomics 1988: Proceedings of the Ergonomics Society 1988 Annual Conference*, Manchester, England, 11-15 April, 1988. London : Taylor & Francis, pp. 521-525.

Teitelbaum, R.C. & Granda, R. (1983) The effects of positional constancy on searching menus for information. *Proceedings of CHI '83. Human Factors in Computing Systems.* Baltimore, MD: ACM. pp. 150-153.

Thimbleby, H. (1990) *User Interface Design.* NY: ACM Press.

Thorndyke, E.L. & Lorge, I. (1972) *The Teacher's Word Book of 30,000 Words.* NY: Teachers College.

Thorndyke, P.W. (1981) Spatial cognition and reasoning.In J.H. Harvey (ed.) *Cognition, Social Behaviour, and the Environment.* Hillsdale NJ: Erlbaum. pp 111-136.

Thorndyke, P.W. & Hayes-Roth, B. (1982) Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology,* 14, 560-589.

Thorndyke, P.W. & Stasz, C. (1980) Individual differences in procedures for knowledge aquisition from maps. *Cognitive Psychology,* 12, 137-175.

Thurstone, T.G. (1958) *Manual for the SRA Priamry Mental Abilities.* Chicago: Science Research Associates.

Tiedeman, J. (1989) Measures of cognitive styles: A critical review. *Educational Psychologist,* 24, 261-275.

Tischer, M. (1989) *PC System Programming.* Grand Rapids, MI: ABACUS.

Tolman, E.C. (1948) Cognitive maps in rats and men. *Psychological Review,* 55, 189-208.

Treat, N.J. & Reese, H.W. (1976) Age, pacing, and imagery in paired associate learning. *Developmental Psychology,* 12, 2, 119-124.

Trevallyan, R. & Browne, D.P. (1987) A self-regulating adaptive system. In *Conference on Human Factors in Computing Systems, CHI '87.* pp. 103-107.

Trumbly, J.E., Arnett, K.P., & Martin, M.P. (1993) Performance effect of matching computer interface characteristics and user skill level. *International Journal of Man-Machine Studies,* 38, 713-724.

Turner, M.L. & Engle, R.W. (1989) Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.

Ukelson, J. & Makowsky, J. (1992) Formal interactive menu design. *Interacting with Computers*, 4, 1, 83-101.

Valdez, F., Chignell, M., & Glenn, B. (1988) Browsing models for hypermedia databases. In *Proceedings of the Human Factors Society, 32nd Annual Meeting*. pp. 318-322.

Valentine, D.W. (1990) Adapting software to the user's cognitive style: Field dependence-field independence. In *Symposium on Smanll Systems. Proceedings of the 1990 ACM Symposium / PC Symposium*. 28-30th March, 1990. New York: ACM Press. pp. 61-67.

Van der Veer, G.C. (1989a) Individual differences and the user interface. *Ergonomics*, 32, 11, 1431-1449.

Van der Veer, G.C. (1989b) Users' representations of system-variety as function of user interface, culture, and individual style. In F. Klix, N.A. Streitz, Y. Waern, & H. Wandke (eds.) *Man-Computer Interaction Research - MACINTER II*. North Holland: Elsevier Science. pp. 55-74.

Van der Veer, G.C. & Felt, A.M. (1988) Development of mental models of an office system: A field study on an introductory course. In G.C. van der Veer & G. Mulder (eds.) *Human-Computer Interaction: Psychonomic Aspects*. NY: Springer-Verlag. pp. 250-272.

Van der Veer, G.C., Tauber, M.J., Waern, Y & Van Muylwijk, B. (1985) On the interaction between system and user characteristics. *Behaviour and Information Technology*, 4, 289-308.

Van der Veer, G.C. & Van der Wolde, J. (1982) Psychological aspects of problem solving with the help of computer languages. *Computers and Education*, 6, 229-234.

Van der Veer, G.C., Van Muylwijk, B., & Van de Wolde, G.J.E. (1978) Cognitive styles and differential learning. In A.M. Lesgold, J.W. Pellegrino, S.D.

Fokkema, & R. Glaser (eds.) *Cognitive Psychology and Instruction*. NY: Plenum Press.

Van der Veer, G.C. & Wijk, R. (1988) Teaching a spreadsheet application - visual-spatial metaphors in relation to spatial ability, and the effect on mental models.In P. Gorny & M.J. Tauber (eds.) *Visualization in Human-Computer Interaction. 7th Interdisciplinary Workshop on Informatics and Psychology*, Sharding, Austria. 24-27th May, 1988. NY: Springer-Verlag. pp. 194-208.

Van Hoe, R., Poupeye, K., Vandierendonck, A., & De Soete, G. (1990) Some effects of menu characteristics and user personality on performance with a menu-driven interface. *Behaviour and Information Technology*, 9, 1, 17-29.

Vaubel, K.P. & Gettys, C.F. (1990) Inferring user expertise for adaptive interfaces. *Human Computer Interaction*, 5, 95-117.

Vernon, P.E. (1961) *The Structurs of Human Abilities*. London: Methuen.

Vicente, K.J., Hayes, B.C., & Williges, R.C. (1987) Assaying and isolating individual differences in searching a hierarchical file system. *Human Factors*, 29,3,349-359.

Vicente, K.J. & Williges, R.C. (1988) Accomodating individual differences in searching a hierarchical file system. *International Journal of Man-Machine Studies*, 29(6), 647-668.

Vidulich, M.A. & Tsang, P.S. (1986) Collecting NASA workload ratings: A paper and pencil package. Working Paper. Moffett Field, CA: NASA Ames Research Center.

Waern, Y. (1989) *Cognitive Aspects of Computer Supported Tasks*. Chichester: Wiley & Sons.

Wetherell, A. (1979) Short term memory for verbal and graphic route information. In *Proceedings of the Human Factors Society 23rd. Annual Meeting*. NY: HFS. pp. 464-468.

Whiteside, J., Jones, S., Levy, P.S., & Wixon, D. (1985) User performance with command, menu, and iconic interfaces. In L. Borman & B. Curtis (Eds.), *Human Factors in Computing Systems II*. North Holland: Elsevier Science. pp. 185-191.

Wickens, C.D. (1984a) *Engineering Psychology and Human Performance*. Illinois: Scott, Foresman and Co.

Wickens, C.D. (1984b) Processing resources in attention. In D.R. Davies and R. Parasuraman (eds.) *Varieties of Attention*. London: Academic Press.

Wickens, C.D. (1992) *Engineering Psychology and Human Performance*, 2nd Edition. NY: HarperCollins.

Wickens, C.D. & Liu, Y. (1988) Codes and modalities in multiple resources: A success and a qualification. *Human Factors*, 30, 5, 599-616.

Wickens, C.D., Sandry, D.L., & Vidulich, M. (1983) Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors*, 25, 2, 227-248.

Wickens, C.D. & Weingartner, A. (1985) Process control monitoring: The effects of spatial and verbal ability and concurrent task demand. In R.E. Eberts and C.G. Eberts (eds.) *Trends in Ergonomics/Human Factors II*. North Holland: Elsevier Science.

Wiedenbeck, S. (1985) Novice/expert differences in programming skills. *International Journal of Man-Machine Studies*, 23, 383-390.

Will, R.P. (1992) Individual differences in the performance and use of an expert system. *International Journal of Man-Machine Studies*, 37, 173-190.

Witkin, H.A. & Goodenough, D.R. (1981) *Cognitive Styles: Essence and Origins*. Int. University Press.

Woelfl, N.N. (1984) Individual differences in online search behaviour: The effect of learning styles and cognitive abilities on process and outcome. Unpublished Doctoral Dissertation, School of Information and Library Science, Case Western Reserve University.

Wogalter, M.S. & Frei, R.L. (1990) Social influence and preference of direct-manipulation and keyboard-command computer interfaces. In *Proceedings of the Human Factors Society - 34th Annual Meeting*. NY: HFS. pp. 907-911.

Wolf, J.M. (1970) A new look at programming aptitudes. *Business Automation*, 17, 36-45.

Woltz, D.J. (1988) An investigation of the role of working memory in procedural skill acquisition. *Journal of Experimental Psychology : General*, 117, 3, 319-331.

Wong, G.K.T. (1987) Effects of learner background characteristics, learning methods, and their interactions, on the leaning of a managerial application software system. Unpublished Doctoral Dissertation, University of California, Los Angeles.

Woods, D.D. (1984) Visual momentum: a concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies*, 21, 229-244.

Woods, D.D. & Roth, E.M. (1988) Cognitive systems engineering. In M. Helander (ed.) *Handbook of Human-Computer Interaction*. Amsterdam: North Holland.

Wright, P. & Lickorish, A. (1989) An empirical comparison of two navigation systems for two hypertexts. In *HYPERTEXT II, the Second Annual UK Conference on Hypertext*, York, July. pp. 84-93.

Young, R.M. & MacLean, A. (1988) Choosing between methods: Analysing the users decision space in terms of schemas and linear models. *In Conference on Human Factors in Computing Systems, CHI '88*, May, 1988, Washington DC. ACM, SIGCHI. pp. 139-143.

Zandri, E. & Charness, N. (1989) Training older and younger adults to use software. *Educational Gerontology*, 15, 6, 615-631.

Ziegler, J.E. & Fahnrich, K-P. (1988) Direct manipulation In M. Hellander (ed.) *Handbook of Human-Computer Interaction*. Amsterdam: North Holland. pp. 123-133.

Zipf, G.K. (1949) *Human Behavior and the Principle of Least Effort*. Cambridge Mass.: Addison-Wesley.

Appendix - Chapter 2

Regressions analysing interactions between age and cognitive ability for each component task response time and for molar task completion time.

Page task reponse time - Interaction between age group and spatial visualisation				
Variance	R square	SS	df	p
Age group	.114	34196545.55	1	
SV	.129	38757398.45	1	
Age x SV	.132	39642334.87	1	ns
Residual		260830523.94	60	

Page task reponse time - Interaction between age group and spatial memory				
Variance	R square	SS	df	p
Age group	.114	34196545.55	1	
SM	.152	45643389.01	1	
Age x SM	.154	46174905.13	1	ns
Residual		254297953.68	60	

Page task reponse time - Interaction between age group and Nelson-Denny vocabulary test scores				
Variance	R square	SS	df	p
Age group	.114	34196545.55	1	
ND	.196	58812713.09	1	
Age x ND	.255	76493739.94	1	<.05
Residual		223979118.87	60	

Page task reponse time - Interaction between age group and computer literacy				
Variance	R square	SS	df	p
Age group	.114	34196545.55	1	
LIT	.224	67359756.13	1	
Age x LIT	.277	83292746.69	1	<.05
Residual		217180112.12	60	

Page task reponse time - Interaction between age group and logical reasoning				
Variance	R square	SS	df	p
Age group	.114	34196545.55	1	
LR	.249	74715239.90	1	
Age x LR	.283	85177223.41	1	ns
Residual		215295635.40	60	

Scroll task reponse time - Interaction between age group and spatial visualisation				
Variance	R square	SS	df	p
Age group	.168	83427101.65	1	
SV	.185	92161008.67	1	
Age x SV	.188	93734153.38	1	ns
Residual		404191953.19	60	

Scroll task reponse time - Interaction between age group and spatial memory				
Variance	R square	SS	df	p
Age group	.168	83427101.65	1	
SM	.231	115171357.02	1	
Age x SM	.246	122728068.20	1	ns
Residual		375198038.38	60	

Scroll task reponse time - Interaction between age group and Nelson-Denny vocabulary test scores				
Variance	R square	SS	df	p
Age group	.168	83427101.65	1	
ND	.251	124876650.90	1	
Age x ND	.344	171408357.93	1	<.01
Residual		326517748.65	60	

Scroll task reponse time - Interaction between age group and computer literacy				
Variance	R square	SS	df	p
Age group	.168	83427101.65	1	
LIT	.263	130913746.74	1	
Age x LIT	.279	138819577.76	1	ns
Residual		359106528.82	60	

Scroll task reponse time - Interaction between age group and logical reasoning				
Variance	R square	SS	df	p
Age group	.168	83427101.65	1	
LR	.312	155328006.16	1	
Age x LR	.315	157045835.50	1	ns
Residual		340880271.07	60	

Fine search task reponse time - Interaction between age group and spatial visualisation				
Variance	R square	SS	df	p
Age group	..221	24322774.54	1	
SV	.239	26312653.27	1	
Age x SV	.239	26384366.81	1	ns
Residual		83829345.47	60	

Fine search task reponse time - Interaction between age group and spatial memory				
Variance	R square	SS	df	p
Age group	..221	24322774.54	1	
SM	.305	33662281.35	1	
Age x SM	.319	35167436.73	1	ns
Residual		75046275.55	60	

Fine search task reponse time - Interaction between age group and Nelson-Denny vocabulary test scores				
Variance	R square	SS	df	p
Age group	..221	24322774.54	1	
ND	.471	24496324.16	1	
Age x ND	.250	27545743.10	1	ns
Residual		82667969.18	60	

Fine search task reponse time - Interaction between age group and computer literacy				
Variance	R square	SS	df	p
Age group	..221	24322774.54	1	
LIT	.308	33904152.97	1	
Age x LIT	.327	36032179.29	1	ns
Residual		74181532.99	60	

Fine search task reponse time - Interaction between age group and logical reasoning				
Variance	R square	SS	df	p
Age group	..221	24322774.54	1	
LR	.268	29540923.17	1	
Age x LR	.294	32454682.40	1	ns
Residual		77759029.88	60	

Decision task reponse time - Interaction between age group and spatial visualisation				
Variance	R square	SS	df	p
Age group	.347	9555177.41	1	
SV	.373	10254185.89	1	
Age x SV	.392	10780523.83	1	ns
Residual		16747793.93	60	

Decision task reponse time - Interaction between age group and spatial memory				
Variance	R square	SS	df	p
Age group	.347	9555177.41	1	
SM	.469	12902871.38	1	
Age x SM	.469	12905288.26	1	ns
Residual		14623029.50	60	

Decision task reponse time - Interaction between age group and Nelson-Denny vocabulary test scores				
Variance	R square	SS	df	p
Age group	.347	9555177.41	1	
ND	.358	9846006.44	1	
Age x ND	.371	10235988.05	1	ns
Residual		17292329.71	60	

Decision task reponse time - Interaction between age group and computer literacy				
Variance	R square	SS	df	p
Age group	.347	9555177.41	1	
LIT	.417	11466733.12	1	
Age x LIT	.426	11733727.59	1	ns
Residual		15794590.17	60	

Decision task reponse time - Interaction between age group and logical reasoning				
Variance	R square	SS	df	p
Age group	.347	9555177.41	1	
LR	.416	11462263.99	1	
Age x LR	.434	11956425.64	1	ns
Residual		15571892.12	60	

Insert task response time - Interaction between age group and spatial visualisation				
Variance	R square	SS	df	p
Age group	.207	37138862.28	1	
SV	.243	43631171.76	1	
Age x SV	.243	43715958.33	1	ns
Residual		136110651.74	60	

Insert task response time - Interaction between age group and spatial memory				
Variance	R square	SS	df	p
Age group	.207	37138862.28	1	
SM	.265	47705440.87	1	
Age x SM	.277	49731802.87	1	ns
Residual		130094807.20	60	

Insert task response time - Interaction between age group and Nelson-Denny vocabulary test score				
Variance	R square	SS	df	p
Age group	.207	37138862.28	1	
ND	.218	39231315.91	1	
Age x ND	.219	39333747.45	1	ns
Residual		140492862.62	60	

Insert task response time - Interaction between age group and computer literacy				
Variance	R square	SS	df	p
Age group	.207	37138862.28	1	
LIT	.230	41367784.30	1	
Age x LIT	.251	45222054.62	1	ns
Residual		134604555.46	60	

Insert task response time - Interaction between age group and logical reasoning				
Variance	R square	SS	df	p
Age group	.207	37138862.28	1	
LR	.280	50350539.73	1	
Age x LR	.285	51267654.42	1	ns
Residual		128558955.65	60	

Backspace task response time - Interaction between age group and spatial visualisation				
Variance	R square	SS	df	p
Age group	.315	128011104.90	1	
SV	.379	153730116.63	1	
Age x SV	.384	155761059.00	1	ns
Residual		250142445.52	60	

Backspace task response time - Interaction between age group and spatial memory				
Variance	R square	SS	df	p
Age group	.315	128011104.90	1	
SM	.381	154528411.20	1	
Age x SM	.400	162470392.13	1	ns
Residual		243433112.39	60	

Backspace task response time - Interaction between age group and Nelson-Denny vocabulary test scores				
Variance	R square	SS	df	p
Age group	.315	128011104.90	1	
ND	.317	128592355.20	1	
Age x ND	.320	130069618.16	1	ns
Residual		275833886.36	60	

Backspace task response time - Interaction between age group and computer literacy				
Variance	R square	SS	df	p
Age group	.315	128011104.90	1	
LIT	.360	146101250.31	1	
Age x LIT	.361	146367769.79	1	ns
Residual		259535734.73	60	

Backspace task response time - Interaction between age group and logical reasoning				
Variance	R square	SS	df	p
Age group	.315	128011104.90	1	
LR	.378	153266485.61	1	
Age x LR	.386	156655854.15	1	ns
Residual		249247650.38	60	

Delete task response time - Interaction between age group and spatial visualisation				
Variance	R square	SS	df	p
Age group	.324	111195838.69	1	
SV	.347	119171424.64	1	
Age x SV	.349	119608375.71	1	ns
Residual		223385753.80	60	

Delete task response time - Interaction between age group and spatial memory				
Variance	R square	SS	df	p
Age group	.324	111195838.69	1	
SM	.370	126997414.91	1	
Age x SM	.370	127004022.97	1	ns
Residual		215990106.55	60	

Delete task response time - Interaction between age group and Nelson-Denny vocabulary test scores				
Variance	R square	SS	df	p
Age group	.324	111195838.69	1	
ND	.325	111502126.69	1	
Age x ND	.325	111502965.35	1	ns
Residual		231491164.16	60	

Delete task response time - Interaction between age group and computer literacy				
Variance	R square	SS	df	p
Age group	.324	111195838.69	1	
LIT	.340	116724739.12	1	
Age x LIT	.354	121519421.39	1	ns
Residual		221474708.13	60	

Delete task response time - Interaction between age group and logical reasoning				
Variance	R square	SS	df	p
Age group	.324	111195838.69	1	
LR	.367	125942393.80	1	
Age x LR	.369	126537445.05	1	ns
Residual		216456684.47	60	

Molar task completion time - Interaction between age group and spatial visualisation				
Variance	R square	SS	df	p
Age group	.278	690975.11	1	
SV	.286	709538.53	1	
Age x SV	.288	713540.68	1	ns
Residual		1767989.19	60	

Molar task completion time - Interaction between age group and spatial memory				
Variance	R square	SS	df	p
Age group	.278	690975.11	1	
SM	.308	758769.13	1	
Age x SM	.307	761999.97	1	ns
Residual		1719529.91	60	

Molar task completion time - Interaction between age group and Nelson-Denny vocabulary test scores				
Variance	R square	SS	df	p
Age group	.278	690975.11	1	
ND	.354	879549.65	1	
Age x ND	.469	1164432.96	1	<.001
Residual		1317096.91	60	

Molar task completion time - Interaction between age group and computer literacy				
Variance	R square	SS	df	p
Age group	.278	690975.11	1	
LIT	.291	722520.24	1	
Age x LIT	.291	723011.75	1	ns
Residual		1758518.13	60	

Molar task completion time - Interaction between age group and logical reasoning				
Variance	R square	SS	df	p
Age group	.278	690975.11	1	
LR	.351	871914.25	1	
Age x LR	.374	928238.85	1	ns
Residual		1553291.02	60	

Finding composite response time - Interaction between age group and spatial visualisation				
Variance	R square	SS	df	p
Age group	.234	10.34	1	
SV	.258	11.41	1	
Age x SV	.258	11.41	1	ns
Residual		32.85	60	

Finding composite response time - Interaction between age group and spatial memory				
Variance	R square	SS	df	p
Age group	.234	10.34	1	
SM	.320	14.18	1	
Age x SM	.320	14.18	1	ns
Residual		30.08	60	

Finding composite response time - Interaction between age group and Nelson-Denny vocabulary test scores				
Variance	R square	SS	df	p
Age group	.234	10.34	1	
ND	.293	12.98	1	
Age x ND	.374	16.55	1	<.01
Residual		27.72	60	

Finding composite response time - Interaction between age group and computer literacy				
Variance	R square	SS	df	p
Age group	.234	10.34	1	
LIT	.372	16.46	1	
Age x LIT	.379	16.79	1	ns
Residual		27.47	60	

Finding composite response time - Interaction between age group and logical reasoning				
Variance	R square	SS	df	p
Age group	.234	10.34	1	
LR	.381	16.87	1	
Age x LR	.407	18.02	1	ns
Residual		26.24	60	

Generating composite response time - Interaction between age group and spatial visualisation				
Variance	R square	SS	df	p
Age group	.402	18.10	1	
SV	.451	20.28	1	
Age x SV	.453	20.36	1	ns
Residual		24.58	60	

Generating composite response time - Interaction between age group and spatial memory				
Variance	R square	SS	df	p
Age group	.402	18.10	1	
SM	.500	22.49	1	
Age x SM	.506	22.72	1	ns
Residual		22.22	60	

Generating composite response time - Interaction between age group and Nelson-Denny vocabulary test scores				
Variance	R square	SS	df	p
Age group	.402	18.10	1	
ND	.408	18.35	1	
Age x ND	.409	18.38	1	ns
Residual		26.56	60	

Generating composite response time - Interaction between age group and computer literacy				
Variance	R square	SS	df	p
Age group	.402	18.10	1	
LIT	.451	20.27	1	
Age x LIT	.454	20.41	1	ns
Residual		24.53	60	

Generating composite response time - Interaction between age group and logical reasoning				
Variance	R square	SS	df	p
Age group	.402	18.10	1	
LR	.488	21.91	1	
Age x LR	.498	22.40	1	ns
Residual		22.54	60	

Appendix - Chapter 3

ANOVA summary tables

ANOVA summary table for response times (not including variables relating to individual differences).

Variance	df	MS	F Ratio	p
Within cells	63	1966563.5		
Command (A)	1	1639600.2	.83	.365
Within cells	63	4069229.3		
Window size (B)	1	702655840	172.68	<.001
Within cells	63	1531455.5		
Target location (C)	1	60269506	39.35	<.001
Within cells	63	2975860.3		
Target distance (D)	1	3.643E + 09	1224.10	<.001
Within cells	63	968208.29		
A x B	1	3068956.5	3.17	.08
Within cells	63	1112853.4		
A x C	1	707438.70	.64	.428
Within cells	63	1506231.5		
A x D	1	5464344.5	3.63	.061
Within cells	63	834612.49		
B x C	1	10633917	12.74	.001
Within cells	63	2051683.0		
B x D	1	72516401	35.34	<.001
Within cells	63	1315903.7		
C x D	1	8589845.1	6.53	.013
Within cells	63	1180220.8		
A x B x C	1	31141.22	.04	.845
Within cells	63	1180220.8		
A x B x D	1	1358900.2	1.15	.287
Within cells	63	2259246.1		
A x C x D	1	54741.38	.02	.877
Within cells	63	963167.51		
B x C x D	1	3697808.8	3.84	.054
Within cells	63	963167.51		
A x B x C x D	1	7240.95	.01	.942

•

ANOVA summary table for response times for interactive effects of expertise

Variance	df	MS	F Ratio	p
Within cells	62	12518609		
Expertise (E)	1	231571.49	.02	.892
Within cells	62	1976343.3		
E x A	1	1360212.0	.69	.410
Within cells	62	4132977.9		
E x B	1	116814.42	.03	.867
Within cells	62	1551438.4		
E x C	1	292511.96	.19	.666
Within cells	62	3023135.5		
E x D	1	44798.37	.01	.904
Within cells	62	980827.98		
E x A x B	1	185787.94	.19	.665
Within cells	62	1078491.9		
E x A x C	1	3243263.3	3.01	.088
Within cells	62	1523264.3		
E x A x D	1	450199.06	.30	.589
Within cells	62	848043.94		
E x B x C	1	1862.46	.00	.963
Within cells	62	2084774.5		
E x B x D	1	11.60	.00	.998
Within cells	62	1294102.0		
E x C x D	1	2667607.6	2.06	.156
Within cells	62	658837.88		
E x A x B x C	1	10334015	15.69	<.001
Within cells	62	1199250.9		
E x A x B x D	1	350.39	.00	.986
Within cells	62	2164341.2		
E x A x C x D	1	8143354.0	3.76	.057
Within cells	62	955008.60		
E x B x C x D	1	1469019.8	1.54	.220
Within cells	62	1337233.2		
E x A x B x C x D	1	2830490.8	2.12	.151

ANOVA summary table for response times for interactive effects of computer literacy				
Variance	df	MS	F Ratio	p
Within cells	62	12406190		
Computer literacy (LIT)	1	7201545.0	.58	.449
Within cells	62	1970812.9		
LIT x A	1	1703100.6	.86	.356
Within cells	62	4111744.9		
LIT x B	1	1433262.0	.35	.557
Within cells	62	1543484.9		
LIT x C	1	785633.44	.51	.478
Within cells	62	3011725.4		
LIT x D	1	752222.10	.25	.619
Within cells	62	978183.79		
LIT x A x B	1	349727.58	.36	.552
Within cells	62	1111739.5		
LIT x A x C	1	1181914.5	1.06	.307
Within cells	62	1521040.6		
LIT x A x D	1	588068.09	.39	.536
Within cells	62	847621.69		
LIT x B x C	1	28042.10	.03	.856
Within cells	62	2063158.6		
LIT x B x D	1	1340199.6	.65	.423
Within cells	62	1336841.5		
LIT x C x D	1	17757.55	.01	.909
Within cells	62	773388.82		
LIT x A x B x C	1	3231856.2	4.18	.045
Within cells	62	1195558.5		
LIT x A x B x D	1	229284.30	.19	.663
Within cells	62	2241098.2		
LIT x A x C x D	1	3384418.0	1.51	.224
Within cells	62	961019.22		
LIT x B x C x D	1	1096361.7	1.14	
Within cells	62	1296782.4		
LIT x A x B x C x D	1	5338439.7	4.12	.047

•

ANOVA summary table for response times for interactive effects of spatial visualisation				
Variance	df	MS	F Ratio	p
Within cells	60	12691840		
Spatial visualisation (SV)	1	58541.45	.00	.946
SV x E	1	14573376	1.15	.288
Within cells	60	2007007.9		
SV x A	1	1561154.0	.78	.381
SV x E x A	1	540714.33	.27	.606
Within cells	60	4122811.9		
SV x B	1	7542128.7	1.83	.181
SV x E x B	1	1296470.6	.31	.577
Within cells	60	1590509.4		
SV x C	1	685035.60	.43	.514
SV x E x C	1	76259.98	.05	.827
Within cells	60	3115970.5		
SV x D	1	285659.68	.09	.763
SV x E x D	1	187752.48	.06	.807
Within cells	60	990981.26		
SV x A x B	1	299507.95	.30	.585
SV x E x A x B	1	1046278.9	1.06	.308
Within cells	60	1081942.4		
SV x A x C	1	1720625.4	1.59	.212
SV x E x A x C	1	236813.08	.22	.642
Within cells	60	1539141.1		
SV x A x D	1	304651.90	.20	.658
SV x E x A x D	1	1797953.6	1.17	.284
Within cells	60	809411.53		
SV x B x C	1	3850811.3	4.76	.033
SV x E x B x C	1	172732.74	.21	.646
Within cells	60	2004380.6		
SV x B x D	1	6867620.9	3.43	.069
SV x E x B x D	1	2080515.1	1.04	.312
Within cells	60	1316732.4		
SV x C x D	1	219181.36	.17	.685
SV x E x C x D	1	1016743.2	.77	.383
Within cells	60	663469.03		
SV x A x B x C	1	158077.42	.24	.627
SV x E x A x B x C	1	886122.12	1.34	.252
Within cells	60	1207728.0		
SV x A x B x D	1	907881.12	.75	.389
SV x E x A x B x D	1	970821.16	.80	.374
Within cells	60	2147688.3		
SV x A x C x D	1	3114306.5	1.45	.233
SV x E x A x C x D	1	2244651.9	1.05	.311
Within cells	60	923307.06		
SV x B x C x D	1	3728745.9	4.04	.049
SV x E x B x C x D	1	90089.23	.10	.756
Within cells	60	1329431.9		
SV x A x B x C x D	1	1067235.0	.80	.374
SV x E x A x B x C x D	1	2057659.3	1.55	.218

ANOVA summary table for response times for interactive effects of spatial memory				
Variance	df	MS	F Ratio	p
Within cells	60	12758166		
Spatial memory (SM)	1	2674693.4	.21	.649
SM x E	1	7895906.9	.62	.435
Within cells	60	1876126.3		
SM x A	1	7735759.5	4.12	.047
SM x E x A	1	2313728.5	1.23	.271
Within cells	60	4190071.4		
SM x B	1	4835805.8	1.15	.287
SM x E x B	1	7996.91	.00	.965
Within cells	60	1488165.1		
SM x C	1	6205751.4	4.17	.046
SM x E x C	1	652508.40	.44	.510
Within cells	60	3096022.0		
SM x D	1	631252.46	.20	.653
SM x E x D	1	1025535.6	.33	.567
Within cells	60	1001475.8		
SM x A x B	1	665807.12	.66	.418
SM x E x A x B	1	53135.88	.05	.819
Within cells	60	1080899.2		
SM x A x C	1	419618.81	.39	.536
SM x E x A x C	1	1609192.9	1.49	.227
Within cells	60	1488632.9		
SM x A x D	1	5123627.7	3.44	.068
SM x E x A x D	1	2572.85	.00	.967
Within cells	60	862866.75		
SM x B x C	1	798935.02	.93	.340
SM x E x B x C	1	9444.46	.01	.917
Within cells	60	2133070.0		
SM x B x D	1	166926.25	.08	.781
SM x E x B x D	1	1113405.7	.52	.473
Within cells	60	1173514.2		
SM x C x D	1	8438357.2	7.19	.009
SM x E x C x D	1	1317316.8	1.12	.294
Within cells	60	619605.76		
SM x A x B x C	1	658005.93	1.06	.307
SM x E x A x B x C	1	2985141.7	4.82	.032
Within cells	60	1233844.8		
SM x A x B x D	1	125753.96	.10	.751
SM x E x A x B x D	1	200261.96	.16	.688
Within cells	60	2131018.1		
SM x A x C x D	1	3287832.1	1.54	.219
SM x E x A x C x D	1	2976907.1	1.40	.242
Within cells	60	923166.77		
SM x B x C x D	1	2530442.3	2.74	.103
SM x E x B x C x D	1	1254002.8	1.36	.248
Within cells	60	1370734.8		
SM x A x B x C x D	1	253812.72	.19	.669
SM x E x A x B x C x D	1	417010.17	.30	.583

ANOVA summary table for response times for interactive effects of Nelson-Denny vocabulary test scores				
Variance	df	MS	F Ratio	p
Within cells	60	12898459		
Nelson-Denny (ND)	1	2244013.4	.17	.678
ND x E	1	2202.85	.00	.990
Within cells	60	2034666.6		
ND x A	1	169783.72	.08	.774
ND x E x A	1	283506.74	.14	.710
Within cells	60	3882710.1		
ND x B	1	6926216.3	1.78	.187
ND x E x B	1	16355810	4.21	.044
Within cells	60	1456540.9		
ND x C	1	8006031.9	5.50	.022
ND x E x C	1	790699.06	.54	.464
Within cells	60	2947058.7		
ND x D	1	10579566	3.59	.063
ND x E x D	1	31310.04	.01	.918
Within cells	60	900653.35		
ND x A x B	1	700829.96	.78	.381
ND x E x A x B	1	6071303.8	6.74	.012
Within cells	60	1026507.8		
ND x A x C	1	122218.38	.12	.731
ND x E x A x C	1	5153809.4	5.02	.029
Within cells	60	1569007.3		
ND x A x D	1	259806.99	.17	.686
ND x E x A x D	1	42139.06	.03	.870
Within cells	60	874906.88		
ND x B x C	1	12113.75	.01	.907
ND x E x B x C	1	72198.24	.08	.775
Within cells	60	2018819.4		
ND x B x D	1	4191806.8	2.08	.155
ND x E x B x D	1	3935047.3	1.95	.168
Within cells	60	1327128.4		
ND x C x D	1	323221.24	.24	.623
ND x E x C x D	1	283397.75	.21	.646
Within cells	60	649279.00		
ND x A x B x C	1	1840847.3	2.84	.097
ND x E x A x B x C	1	50360.96	.08	.782
Within cells	60	1161011.6		
ND x A x B x D	1	2053266.5	1.77	.189
ND x E x A x B x D	1	2639598.8	2.27	.137
Within cells	60	2199044.5		
ND x A x C x D	1	230304.69	.10	.747
ND x E x A x C x D	1	2016178.4	.92	.342
Within cells	60	951598.64		
ND x B x C x D	1	1392625.7	1.46	.231
ND x E x B x C x D	1	721989.36	.76	.387
Within cells	60	1319446.5		
ND x A x B x C x D	1	3644341.6	2.76	.102
ND x E x A x B x C x D	1	97323.01	.07	.787

ANOVA summary table for errors (not including variables relating to individual differences).				
Variance	df	MS	F Ratio	p
Within cells	63	.01		
Command (A)	1	.09	16.29	<.001
Within cells	63	.02		
Window size (B)	1	.02	1.17	.283
Within cells	63	.00		
Target location (C)	1	.00	.86	.359
Within cells	63	.01		
Target distance (D)	1	.14	19.74	<.001
Within cells	63	.01		
A x B	1	.00	.49	.488
Within cells	63	.00		
A x C	1	.01	2.19	.144
Within cells	63	.01		
A x D	1	.00	.00	.971
Within cells	63	.01		
B x C	1	.00	.27	.605
Within cells	63	.01		
B x D	1	.03	4.02	.049
Within cells	63	.01		
C x D	1	.01	2.00	.162
Within cells	63	.00		
A x B x C	1	.01	1.32	.255
Within cells	63	.00		
A x B x D	1	.00	.01	.909
Within cells	63	.01		
A x C x D	1	.01	1.56	.217
Within cells	63	.00		
B x C x D	1	.03	5.30	.025
Within cells	63	.00		
A x B x C x D	1	.01	3.00	.088

ANOVA summary table for errors for interactive effects of expertise				
Variance	df	MS	F Ratio	p
Within cells	62	.02		
Expertise (E)	1	.14	6.30	.015
Within cells	62	.01		
E x A	1	.00	.20	.657
Within cells	62	.02		
E x B	1	.01	.43	.514
Within cells	62	.00		
E x C	1	.01	2.01	.161
Within cells	62	.01		
E x D	1	.01	2.11	.151
Within cells	62	.01		
E x A x B	1	.00	.25	.622
Within cells	62	.00		
E x A x C	1	.01	2.49	.120
Within cells	62	.01		
E x A x D	1	.00	.37	.544
Within cells	62	.01		
E x B x C	1	.00	.75	.391
Within cells	62	.01		
E x B x D	1	.01	2.09	.153
Within cells	62	.00		
E x C x D	1	.02	4.85	.031
Within cells	62	.00		
E x A x B x C	1	.01	2.12	.150
Within cells	62	.00		
E x A x B x D	1	.00	.07	.791
Within cells	62	.01		
E x A x C x D	1	.03	.488	.031
Within cells	62	.00		
E x B x C x D	1	.00	.04	.852
Within cells	62	.00		
E x A x B x C x D	1	.00	1.12	.294

ANOVA summary table for errors for interactive effects of computer literacy				
Variance	df	MS	F Ratio	p
Within cells	62	.02		
Computer literacy (LIT)	1	.02	.83	.366
Within cells	62	.01		
LIT x A	1	.00	.10	.750
Within cells	62	.01		
LIT x B	1	.07	4.94	.03
Within cells	62	.00		
LIT x C	1	.00	.76	.386
Within cells	62	.01		
LIT x D	1	.00	.05	.820
Within cells	62	.01		
LIT x A x B	1	.00	.06	.801
Within cells	62	.00		
LIT x A x C	1	.00	.00	.995
Within cells	62	.01		
LIT x A x D	1	.00	.22	.644
Within cells	62	.01		
LIT x B x C	1	.00	.00	.960
Within cells	62	.01		
LIT x B x D	1	.01	1.90	.173
Within cells	62	.01		
LIT x C x D	1	.00	.25	.621
Within cells	62	.00		
LIT x A x B x C	1	.03	5.67	.02
Within cells	62	.00		
LIT x A x B x D	1	.00	.17	.679
Within cells	62	.01		
LIT x A x C x D	1	.00	.11	.746
Within cells	62	.00		
LIT x B x C x D	1	.00	.07	.791
Within cells	62	.00		
LIT x A x B x C x D	1	.01	1.38	.244

ANOVA summary table for errors for interactive effects of spatial visualisation

Variance	df	MS	F Ratio	p
Within cells	60	.02		
Spatial visualisation (SV)	1	.00	.01	.903
SV x E	1	.05	2.51	.118
Within cells	60	.01		
SV x A	1	.00	.64	.427
SV x E x A	1	.02	3.29	.071
Within cells	60	.02		
SV x B	1	.00	.01	.942
SV x E x B	1	.01	.61	.438
Within cells	60	.00		
SV x C	1	.00	.01	.928
SV x E x C	1	.01	1.68	.200
Within cells	60	.01		
SV x D	1	.00	.13	.718
SV x E x D	1	.00	.13	.718
Within cells	60	.01		
SV x A x B	1	.00	.03	.870
SV x E x A x B	1	.00	.02	.883
Within cells	60	.00		
SV x A x C	1	.00	.25	.618
SV x E x A x C	1	.01	1.52	.222
Within cells	60	.01		
SV x A x D	1	.00	.42	.518
SV x E x A x D	1	.02	3.73	.058
Within cells	60	.01		
SV x B x C	1	.01	1.60	.211
SV x E x B x C	1	.00	.64	.426
Within cells	60	.01		
SV x B x D	1	.01	1.62	.210
SV x E x B x D	1	.00	.36	.552
Within cells	60	.00		
SV x C x D	1	.01	2.29	.135
SV x E x C x D	1	.01	2.28	.136
Within cells	60	.00		
SV x A x B x C	1	.01	3.12	.082
SV x E x A x B x C	1	.01	1.41	.239
Within cells	60	.00		
SV x A x B x D	1	.00	.03	.863
SV x E x A x B x D	1	.00	.77	.385
Within cells	60	.01		
SV x A x C x D	1	.01	1.13	.293
SV x E x A x C x D	1	.00	.67	.416
Within cells	60	.00		
SV x B x C x D	1	.01	2.87	.095
SV x E x B x C x D	1	.00	.51	.477
Within cells	60	.00		
SV x A x B x C x D	1	.00	.01	.935
SV x E x A x B x C x D	1	.00	.71	.402

ANOVA summary table for errors for interactive effects of spatial memory				
Variance	df	MS	F Ratio	p
Within cells	60	.02		
Spatial memory (SM)	1	.14	7.07	.01
SM x E	1	.05	2.69	.106
Within cells	60	.01		
SM x A	1	.00	.41	.525
SM x E x A	1	.00	.02	.888
Within cells	60	.02		
SM x B	1	.01	.51	.477
SM x E x B	1	.01	.36	.551
Within cells	60	.00		
SM x C	1	.00	.49	.486
SM x E x C	1	.00	.10	.751
Within cells	60	.01		
SM x D	1	.07	11.87	.001
SM x E x D	1	.00	.19	.667
Within cells	60	.01		
SM x A x B	1	.00	.06	.809
SM x E x A x B	1	.00	.02	.899
Within cells	60	.00		
SM x A x C	1	.04	10.38	.002
SM x E x A x C	1	.01	4.23	.044
Within cells	60	.01		
SM x A x D	1	.00	.72	.399
SM x E x A x D	1	.01	1.19	.279
Within cells	60	.01		
SM x B x C	1	.00	.16	.692
SM x E x B x C	1	.00	.01	.909
Within cells	60	.01		
SM x B x D	1	.01	1.51	.224
SM x E x B x D	1	.00	.02	.893
Within cells	60	.00		
SM x C x D	1	.00	.46	.500
SM x E x C x D	1	.01	2.21	.142
Within cells	60	.00		
SM x A x B x C	1	.01	2.58	.113
SM x E x A x B x C	1	.00	.26	.609
Within cells	60	.00		
SM x A x B x D	1	.01	1.15	.288
SM x E x A x B x D	1	.00	.23	.631
Within cells	60	.01		
SM x A x C x D	1	.00	.58	.448
SM x E x A x C x D	1	.01	.83	.366
Within cells	60	.00		
SM x B x C x D	1	.01	1.83	.181
SM x E x B x C x D	1	.00	.91	.345
Within cells	60	.00		
SM x A x B x C x D	1	.00	.35	.554
SM x E x A x B x C x D	1	.00	.70	.406

ANOVA summary table for errors for interactive effects of Nelson-Denny vocabulary test scores				
Variance	df	MS	F Ratio	p
Within cells	60	.02		
Nelson-Denny (ND)	1	.07	3.29	.75
ND x E	1	.01	.40	.531
Within cells	60	.01		
ND x A	1	.01	.88	.352
ND x E x A	1	.00	.11	.743
Within cells	60	.02		
ND x B	1	.00	.01	.918
ND x E x B	1	.02	1.18	.281
Within cells	60	.00		
ND x C	1	.00	.00	.956
ND x E x C	1	.00	.02	.877
Within cells	60	.01		
ND x D	1	.01	.81	.372
ND x E x D	1	.00	.16	.687
Within cells	60	.01		
ND x A x B	1	.00	.03	.867
ND x E x A x B	1	.00	.03	.854
Within cells	60	.00		
ND x A x C	1	.00	.01	.932
ND x E x A x C	1	.00	.91	.343
Within cells	60	.01		
ND x A x D	1	.00	.01	.919
ND x E x A x D	1	.00	.21	.648
Within cells	60	.01		
ND x B x C	1	.00	.12	.729
ND x E x B x C	1	.00	.14	.710
Within cells	60	.01		
ND x B x D	1	.00	.39	.537
ND x E x B x D	1	.00	.38	.542
Within cells	60	.00		
ND x C x D	1	.00	.10	.756
ND x E x C x D	1	.00	.93	.338
Within cells	60	.01		
ND x A x B x C	1	.00	.01	.937
ND x E x A x B x C	1	.00	.23	.635
Within cells	60	.00		
ND x A x B x D	1	.02	3.87	.054
ND x E x A x B x D	1	.00	.99	.323
Within cells	60	.01		
ND x A x C x D	1	.00	.32	.571
ND x E x A x C x D	1	.00	.10	.747
Within cells	60	.00		
ND x B x C x D	1	.00	.10	.755
ND x E x B x C x D	1	.00	.72	.401
Within cells	60	.00		
ND x A x B x C x D	1	.00	1.26	.266
ND x E x A x B x C x D	1	.01	3.04	.086

ANOVA summary table for self-report workload				
Variance	df	MS	F Ratio	p
Within cells	63	65.62		
A	1	2.07	.03	.860
Within cells	63	95.99		
C	1	1658.87	17.28	<.001
Within cells	63	45.27		
A x C	1	132.73	2.93	.092

ANOVA summary table for self-report workload for interactive effects of expertise				
Variance	df	MS	F Ratio	p
Within cells	62	898.64		
E	1	11.53	.01	.910
Within cells	62	65.37		
E x A	1	81.38	1.24	.269
Within cells	62	97.44		
E x C	1	5.94	.06	.806
Within cells	62	45.99		
E x A x C	1	.96	.02	.886

ANOVA summary table for self-report workload for interactive effects of computer literacy				
Variance	df	MS	F Ratio	p
Within cells	62	898.57		
LIT	1	15.76	.02	.895
Within cells	62	65.90		
LIT x A	1	48.22	.73	.396
Within cells	62	97.53		
LIT x C	1	.63	.01	.936
Within cells	62	45.47		
LIT x A x C	1	33.03	.73	.397

ANOVA summary table for self-report workload for interactive effects of spatial visualisation				
Variance	df	MS	F Ratio	p
Within cells	60	882.90		
SV	1	904.33	1.02	.316
SV x E	1	1852.72	2.10	.153
Within cells	60	66.85		
SV x A	1	5.36	.08	.778
SV x E x A	1	36.44	.55	.463
Within cells	60	100.37		
SV x B	1	19.19	.19	.664
SV x E x B	1	.11	.00	.974
Within cells	60	47.38		
SV x A x B	1	6.05	.13	.722
SV x E x A x B	1	2.27	.05	.828

ANOVA summary table for self-report workload for interactive effects of spatial memory				
Variance	df	MS	F Ratio	p
Within cells	60	888.33		
SM	1	618.96	.70	.407
SM x E	1	1775.71	2.00	.163
Within cells	60	67.48		
SM x A	1	1.26	.02	.892
SM x E x A	1	2.59	.04	.845
Within cells	60	90.08		
SM x B	1	633.63	7.03	.010
SM x E x B	1	2.45	.03	.870
Within cells	60	47.50		
SM x A x B	1	.97	.02	.887
SM x E x A x B	1	.50	.01	.918

ANOVA summary table for self-report workload for interactive effects of Nelson-Denny vocabulary test scores				
Variance	df	MS	F Ratio	p
Within cells	60	893.06		
ND	1	953.59	1.07	.306
ND x E	1	1178.49	1.32	.255
Within cells	60	66.76		
ND x A	1	46.61	.70	.407
ND x E x A	1	.72	.01	.918
Within cells	60	99.35		
ND x B	1	6.18	.06	.804
ND x E x B	1	74.34	.75	.390
Within cells	60	47.40		
ND x A x B	1	2.36	.05	.824
ND x E x A x B	1	4.76	.10	.752

ANOVA summary table for command strategy (not including variables relating to individual differences).				
Variance	df	MS	F Ratio	p
Within cells	63	78237.09		
Command (A)	1	146976.39	1.88	.175
Within cells	63	663.57		
Window size (B)	1	8579.39	12.93	.001
Within cells	63	524.86		
Target location (C)	1	888.79	1.69	.198
Within cells	63	914.22		
Target distance (D)	1	1951958.3	2135.10	<.001
Within cells	63	28279.90		
A x B	1	129690.02	4.59	.036
Within cells	63	561.86		
A x C	1	725.63	1.29	.260
Within cells	63	20688.47		
A x D	1	54580.64	2.64	.109
Within cells	63	690.77		
B x C	1	911.29	1.32	.255
Within cells	63	861.40		
B x D	1	907.52	1.05	.309
Within cells	63	764.99		
C x D	1	561.10	.73	.395
Within cells	63	608.22		
A x B x C	1	82.13	.14	.715
Within cells	63	8113.98		
A x B x D	1	47360.64	5.84	.019
Within cells	63	699.54		
A x C x D	1	2030.63	2.90	.093
Within cells	63	875.87		
B x C x D	1	1336.82	1.53	.221
Within cells	63	854.72		
A x B x C x D	1	312.85	.37	.547

ANOVA summary table for command strategy for interactive effects of expertise				
Variance	df	MS	F Ratio	p
Within cells	62	874.20		
Expertise (E)	1	372.97	.43	.516
Within cells	62	79490.86		
E x A	1	503.44	.01	.937
Within cells	62	659.08		
E x B	1	941.72	1.43	.237
Within cells	62	533.32		
E x C	1	.02	.00	.996
Within cells	62	928.65		
E x D	1	19.69	.02	.885
Within cells	62	27780.20		
E x A x B	1	59261.82	2.13	.149
Within cells	62	570.61		
E x A x C	1	19.14	.03	.855
Within cells	62	21016.75		
E x A x D	1	335.35	.02	.900
Within cells	62	701.80		
E x B x C	1	6.89	.01	.921
Within cells	62	861.19		
E x B x D	1	873.94	1.01	.318
Within cells	62	738.80		
E x C x D	1	2388.77	3.23	.077
Within cells	62	617.79		
E x A x B x C	1	15.02	.02	.877
Within cells	62	7830.33		
E x A x B x D	1	25700.10	3.28	.075
Within cells	62	710.17		
E x A x C x D	1	40.64	.06	.812
Within cells	62	887.00		
E x B x C x D	1	185.64	.21	.649
Within cells	62	864.82		
E x A x B x C x D	1	228.77	.26	.609

ANOVA summary table for command strategy for interactive effects of computer literacy				
Variance	df	MS	F Ratio	p
Within cells	62	870.19		
Computer literacy (LIT)	1	621.51	.71	.401
Within cells	62	79448.09		
LIT x A	1	3155.29	.04	.843
Within cells	62	668.88		
LIT x B	1	334.10	.50	.428
Within cells	62	532.99		
LIT x C	1	20.32	.04	.864
Within cells	62	909.01		
LIT x D	1	1237.45	1.36	.248
Within cells	62	28654.94		
LIT x A x B	1	5027.62	.18	.677
Within cells	62	570.82		
LIT x A x C	1	6.11	.01	.918
Within cells	62	21020.22		
LIT x A x D	1	120.00	.01	.940
Within cells	62	699.73		
LIT x B x C	1	135.41	.19	.662
Within cells	62	872.84		
LIT x B x D	1	151.65	.17	.678
Within cells	62	777.26		
LIT x C x D	1	4.83	.01	.937
Within cells	62	595.51		
LIT x A x B x C	1	1396.06	2.34	.131
Within cells	62	8242.25		
LIT x A x B x D	1	161.43	.02	.889
Within cells	62	695.71		
LIT x A x C x D	1	937.01	1.35	.250
Within cells	62	886.79		
LIT x B x C x D	1	198.44	.22	.638
Within cells	62	861.14		
LIT x A x B x C x D	1	456.92	.53	.469

ANOVA summary table for command strategy for interactive effects of spatial visualisation				
Variance	df	MS	F Ratio	p
Within cells	60	889.26		
Spatial visualisation (SV)	1	368.84	.41	.522
SV x E	1	481.05	.54	.465
Within cells	60	81899.11		
SV x A	1	9994.68	.12	.728
SV x E x A	1	4413.29	.05	.817
Within cells	60	660.76		
SV x B	1	906.32	1.37	.246
SV x E x B	1	317.16	.48	.491
Within cells	60	548.08		
SV x C	1	112.01	.20	.653
SV x E x C	1	70.46	.13	.721
Within cells	60	920.51		
SV x D	1	1792.73	1.95	.168
SV x E x D	1	564.99	.61	.436
Within cells	60	27955.85		
SV x A x B	1	16209.93	.58	.449
SV x E x A x B	1	29066.65	1.04	.312
Within cells	60	571.75		
SV x A x C	1	33.38	.06	.810
SV x E x A x C	1	1037.55	1.81	.183
Within cells	60	21631.55		
SV x A x D	1	3146.80	.15	.704
SV x E x A x D	1	1968.56	.09	.764
Within cells	60	714.25		
SV x B x C	1	230.98	.32	.572
SV x E x B x C	1	429.64	.60	.441
Within cells	60	875.79		
SV x B x D	1	37.33	.04	.837
SV x E x B x D	1	811.64	.93	.340
Within cells	60	717.31		
SV x C x D	1	2439.95	3.40	.070
SV x E x C x D	1	316.77	.44	.509
Within cells	60	581.89		
SV x A x B x C	1	67.85	.12	.734
SV x E x A x B x C	1	3327.37	5.72	.020
Within cells	60	7850.90		
SV x A x B x D	1	6284.41	.80	.375
SV x E x A x B x D	1	8227.24	1.05	.310
Within cells	60	721.29		
SV x A x C x D	1	523.64	.73	.398
SV x E x A x C x D	1	233.65	.32	.571
Within cells	60	879.34		
SV x B x C x D	1	4.61	.01	.943
SV x E x B x C x D	1	2229.87	2.54	.117
Within cells	60	844.07		
SV x A x B x C x D	1	1.97	.00	.962
SV x E x A x B x C x D	1	2971.72	3.52	.065

ANOVA summary table for command strategy for interactive effects of spatial memory				
Variance	df	MS	F Ratio	p
Within cells	60	884.74		
Spatial memory (SM)	1	1024.45	1.16	.286
SM x E	1	97.49	.11	.741
Within cells	60	79937.19		
SM x A	1	14214.82	.18	.675
SM x E x A	1	118797.20	1.49	.228
Within cells	60	679.06		
SM x B	1	73.77	.11	.743
SM x E x B	1	46.38	.07	.795
Within cells	60	544.78		
SM x C	1	26.79	.05	.825
SM x E x C	1	354.04	.65	.423
Within cells	60	892.17		
SM x D	1	2820.83	3.16	.080
SM x E x D	1	1262.97	1.42	.239
Within cells	60	28445.82		
SM x A x B	1	11085.62	.39	.535
SM x E x A x B	1	4679.97	.16	.686
Within cells	60	579.00		
SM x A x C	1	437.28	.76	.388
SM x E x A x C	1	194.91	.34	.564
Within cells	60	21421.74		
SM x A x D	1	4672.67	.22	.642
SM x E x A x D	1	13217.09	.62	.435
Within cells	60	678.22		
SM x B x C	1	1025.61	1.51	.224
SM x E x B x C	1	1765.74	2.60	.112
Within cells	60	875.91		
SM x B x D	1	422.38	.48	.490
SM x E x B x D	1	408.53	.47	.497
Within cells	60	696.69		
SM x C x D	1	831.41	1.19	.279
SM x E x C x D	1	3140.41	4.51	.038
Within cells	60	633.13		
SM x A x B x C	1	216.96	.34	.560
SM x E x A x B x C	1	95.14	.15	.700
Within cells	60	7895.71		
SM x A x B x D	1	7455.07	.94	.335
SM x E x A x B x D	1	4396.94	.56	.458
Within cells	60	732.85		
SM x A x C x D	1	7.16	.01	.922
SM x E x A x C x D	1	52.12	.07	.791
Within cells	60	916.50		
SM x B x C x D	1	1.02	.00	.974
SM x E x B x C x D	1	3.20	.00	.953
Within cells	60	893.02		
SM x A x B x C x D	1	32.00	.04	.850
SM x E x A x B x C x D	1	6.17	.01	.934

ANOVA summary table for command strategy for interactive effects of Nelson-Denny vocabulary test scores				
Variance	df	MS	F Ratio	p
Within cells	60	863.65		
Nelson-Denny (ND)	1	1831.55	2.12	.151
ND x E	1	549.88	.64	.428
Within cells	60	78406.56		
ND x A	1	213493.62	2.72	.104
ND x E x A	1	10546.49	.13	.715
Within cells	60	663.22		
ND x B	1	164.13	.25	.621
ND x E x B	1	905.73	1.37	.247
Within cells	60	482.87		
ND x C	1	3291.53	6.82	.011
ND x E x C	1	802.23	1.66	.202
Within cells	60	950.93		
ND x D	1	274.74	.29	.593
ND x E x D	1	245.54	.26	.613
Within cells	60	25944.78		
ND x A x B	1	100765.00	3.88	.053
ND x E x A x B	1	64920.65	2.50	.119
Within cells	60	516.44		
ND x A x C	1	.75	.00	.970
ND x E x A x C	1	4390.64	8.50	.005
Within cells	60	20579.79		
ND x A x D	1	63004.91	3.06	.085
ND x E x A x D	1	5245.73	.25	.615
Within cells	60	722.10		
ND x B x C	1	3.51	.00	.945
ND x E x B x C	1	182.37	.25	.617
Within cells	60	871.60		
ND x B x D	1	516.79	.59	.444
ND x E x B x D	1	581.34	.67	.417
Within cells	60	752.76		
ND x C x D	1	116.38	.15	.696
ND x E x C x D	1	523.97	.70	.407
Within cells	60	634.54		
ND x A x B x C	1	183.86	.29	.592
ND x E x A x B x C	1	46.39	.07	.788
Within cells	60	7244.49		
ND x A x B x D	1	28149.07	3.89	.053
ND x E x A x B x D	1	22662.32	3.13	.082
Within cells	60	717.66		
ND x A x C x D	1	17.42	.02	.877
ND x E x A x C x D	1	953.26	1.33	.254
Within cells	60	901.29		
ND x B x C x D	1	913.33	1.01	.318
ND x E x B x C x D	1	3.37	.00	.951
Within cells	60	870.36		
ND x A x B x C x D	1	22.17	.03	.874
ND x E x A x B x C x D	1	1374.92	1.58	.214

Appendix - Chapter 4

Experiment 1 : ANOVA summary tables

ANOVA summary table for response times for 'model development' blocks.				
Variance	df	MS	F Ratio	p
Within cells	30	2116.42		
Ability group	1	12471.10	5.89	.021
Within cells	90	410.94		
Block	3	15380.52	37.43	<.001
Ability group x Block	3	2688.48	6.54	<.001

ANOVA summary table for accuracy for 'model development' blocks.				
Variance	df	MS	F Ratio	p
Within cells	30	386.70		
Ability group	1	1262.53	3.26	.081
Within cells	90	89.06		
Block	3	2100.86	23.59	<.001
Ability group x Block	3	610.36	6.85	<.001

ANOVA summary table for response times: pre vs post disruption.				
Variance	df	MS	F Ratio	p
Within cells	28	597.90		
Ability group (A)	1	593.78	.99	.328
Experimental vs control (E)	1	569.22	.95	.338
A x E	1	1031.10	1.72	.200
Within cells	28	139.25		
Pre vs post disruption (P)	1	659.61	4.74	.038
A x P	1	13.53	.10	.758
E x P	1	698.76	5.02	.033
A x E x P	1	199.70	1.43	.241

ANOVA summary table for accuracy: pre vs post disruption.				
Variance	df	MS	F Ratio	p
Within cells	28	103.23		
Ability group (A)	1	78.77	.76	.390
Experimental vs control (E)	1	31.64	.31	.584
A x E	1	656.64	6.36	.018
Within cells	28	31.58		
Pre vs post disruption (P)	1	129.39	4.10	.053
A x P	1	31.64	1.00	.325
E x P	1	47.27	1.50	.231
A x E x P	1	1.89	.06	.808

ANOVA summary table for workload: pre vs post disruption.				
Variance	df	MS	F Ratio	p
Within cells	28	407.86		
Ability group (A)	1	411.75	1.01	.324
Experimental vs control (E)	1	359.42	.88	.356
A x E	1	11.39	.03	.868
Within cells	28	84.58		
Pre vs post disruption (P)	1	388.42	4.59	.041
A x P	1	11.96	.14	.710
E x P	1	257.34	3.04	.092
A x E x P	1	8.75	.10	.750

ANOVA summary table for response times: pre vs post disruption for experimental groups only.				
Variance	df	MS	F Ratio	p
Within cells	14	476.26		
Ability group (A)	1	1594.90	3.35	.089
Within cells	14	152.80		
Pre vs post disruption (P)	1	.28	.00	.966
A x P	1	158.60	1.04	.326

ANOVA summary table for accuracy: pre vs post disruption for experimental groups only.				
Variance	df	MS	F Ratio	p
Within cells	14	82.10		
Ability group (A)	1	595.13	7.25	.018
Within cells	14	28.74		
Pre vs post disruption (P)	1	10.13	.35	.562
A x P	1	24.50	.85	.372

Experiment 2 : ANOVA summary tables and correlation matrices

ANOVA summary table for response times				
Variance	df	MS	F Ratio	p
Within cells	44	14.82		
Ability group (A)	1	2.62	.18	.676
Semantic content (S)	1	379.16	25.59	<.001
A x S	1	11.49	.78	.383
Within cells	88	.97		
Block (B)	2	21.32	22.07	<.001
A x B	2	3.90	4.04	.021
S x B	2	5.93	6.14	.003
A x S x B	2	4.50	4.66	.012

ANOVA summary table for navigational efficiency				
Variance	df	MS	F Ratio	p
Within cells	44	.05		
Ability group (A)	1	.14	2.74	.105
Semantic content (S)	1	4.50	90.33	<.001
A x S	1	.09	1.83	.183
Within cells	88	.01		
Block (B)	2	.10	9.95	<.001
A x B	2	.06	5.90	.004
S x B	2	.11	11.02	<.001
A x S x B	2	.03	2.77	.068

Network recollection test: Correlation matrix for survey and route question response times and accuracy (n=48).			
	Survey quest. RT	Route quest. RT	Survey quest. acc.
Route question RT	.74 ***		
Survey question acc.	.14	.16	
Route question acc.	.34 *	.32 *	.48 ***

Network recollection test: Correlation matrix for survey and route question response times and accuracy : Low semantic condition only (n=24).			
	Survey quest. RT	Route quest. RT	Survey quest. acc.
Route question RT	.63 ***		
Survey question acc.	.35	.39	
Route question acc.	.59 **	.59 **	.50 *

Network recollection test: Correlation matrix for survey and route question response times and accuracy : High semantic condition only (n=24).			
	Survey quest. RT	Route quest. RT	Survey quest. acc.
Route question RT	.87 ***		
Survey question acc.	-.02	-.01	
Route question acc.	.16	.21	.36

Brief report 1 : ANOVA summary tables

ANOVA summary table for response times for 'model development' blocks.				
Variance	df	MS	F Ratio	p
Within cells	37	98.03		
Age group	1	682.61	6.96	.012
Within cells	111	3.86		
Block	3	44.63	11.56	<.001
Age group x Block	3	1.28	.33	.803

ANOVA summary table for accuracy for 'model development' blocks.				
Variance	df	MS	F Ratio	p
Within cells	37	21.58		
Ability group	1	8.71	.40	.529
Within cells	111	1.25		
Block	3	6.82	5.43	.002
Ability group x Block	3	2.32	1.85	.142

ANOVA summary table for response times: pre vs post disruption.				
Variance	df	MS	F Ratio	p
Within cells	35	34.07		
Age group (A)	1	290.09	8.51	.006
Experimental vs control (E)	1	108.18	3.18	.083
A x E	1	32.64	.96	.334
Within cells	35	1.38		
Pre vs post disruption (P)	1	11.52	8.33	.007
A x P	1	2.21	1.59	.215
E x P	1	6.90	4.99	.032
A x E x P	1	.46	.33	.567

ANOVA summary table for accuracy: pre vs post disruption.				
Variance	df	MS	F Ratio	p
Within cells	35	8.00		
Ability group (A)	1	12.68	1.59	.216
Experimental vs control (E)	1	18.31	2.29	.139
A x E	1	5.80	.73	.400
Within cells	35	.42		
Pre vs post disruption (P)	1	3.86	9.16	.005
A x P	1	.00	.00	.984
E x P	1	.68	1.62	.212
A x E x P	1	.30	.71	.405

ANOVA summary table for response times: pre vs post disruption for experimental groups only.				
Variance	df	MS	F Ratio	p
Within cells	17	46.85		
Age group (A)	1	253.11	5.40	.033
Within cells	17	1.98		
Pre vs post disruption (P)	1	.29	.15	.707
A x P	1	2.29	1.16	.296

ANOVA summary table for accuracy: pre vs post disruption for experimental groups only.				
Variance	df	MS	F Ratio	p
Within cells	17	10.61		
Age group (A)	1	17.43	1.64	.217
Within cells	17	.59		
Pre vs post disruption (P)	1	.64	1.07	.315
A x P	1	.15	.26	.617

Brief report 2 : ANOVA summary tables

ANOVA summary table for response times during model development blocks.				
Variance	df	MS	F Ratio	p
Within cells	12	41018.94		
Age group (A)	1	81645.05	1.99	.184
Within cells	24	6977.37		
Block (B)	2	22776.66	3.26	.056
A x B	2	2877.38	.41	.667

ANOVA summary table for navigational efficiency during model development blocks.				
Variance	df	MS	F Ratio	p
Within cells	12	.19		
Age group (A)	1	.00	.01	.905
Within cells	24	.06		
Block (B)	2	.06	1.10	.350
A x B	2	.01	.09	.917

ANOVA summary table for response times for each primary task condition.				
Variance	df	MS	F Ratio	p
Within cells	12	3975.56		
Age group (A)	1	17555.33	4.42	.057
Within cells	24	196.08		
Primary task condition (C)	2	2603.77	13.28	<.001
A x B	2	1278.89	6.52	.005

ANOVA summary table for accuracy for each primary task condition.				
Variance	df	MS	F Ratio	p
Within cells	12	.02		
Age group (A)	1	.00	.12	.730
Within cells	24	.01		
Primary task condition (C)	2	.00	.68	.518
A x B	2	.01	1.59	.225

ANOVA summary table for response times for each secondary task condition.				
Variance	df	MS	F Ratio	p
Within cells	12	385106.95		
Age group (A)	1	896267.10	2.33	.153
Within cells	12	19005.87		
Spatial vs Verbal task (SV)	1	116117.91	6.11	.029
A x SV	1	3012.13	.16	.698
Within cells	12	33043.42		
Single vs Dual task (SD)	1	762591.19	23.08	<.001
A x SD	1	45729.51	1.38	.262
Within cells	12	22544.66		
SV x SD	1	165.81	.01	.933
A x SV x SD	1	15417.41	.68	.424

ANOVA summary table for accuracy for each secondary task condition.				
Variance	df	MS	F Ratio	p
Within cells	12	.03		
Age group (A)	1	.00	.08	.787
Within cells	12	.02		
Spatial vs Verbal task (SV)	1	.00	.27	.611
A x SV	1	.03	1.85	.199
Within cells	12	.01		
Single vs Dual task (SD)	1	.90	61.35	<.001
A x SD	1	.01	.34	.569
Within cells	12	.01		
SV x SD	1	.01	.54	.476
A x SV x SD	1	.00	.12	.734

Brief report 3 : ANOVA summary tables

ANOVA summary table for response times				
Variance	df	MS	F Ratio	p
Within cells	34	11237.53		
Semantic content (S)	1	204138.27	18.17	<.001
Within Cells	34	7818.83		
Target difficulty (D)	1	65463.07	8.37	.007
S x D	1	20548.77	2.63	.114
Within cells	68	897.67		
Block (B)	2	9182.02	10.23	<.001
S x B	2	4259.70	4.75	.012
Within cells	68	425.04		
D x B	2	3634.83	8.55	<.001
S x D x B	2	2267.21	5.33	.007

ANOVA summary table for navigational efficiency				
Variance	df	MS	F Ratio	p
Within cells	34	.08		
Semantic content (S)	1	11.84	144.22	<.001
Within Cells	34	.05		
Target difficulty (D)	1	.04	.95	.337
S x D	1	.09	1.87	.180
Within cells	68	.02		
Block (B)	2	.09	5.34	.007
S x B	2	.05	2.61	.081
Within cells	68	.01		
D x B	2	.05	4.96	.010
S x D x B	2	.02	2.23	.115

ANOVA summary table for response times for high and low WM activation groups				
Variance	df	MS	F Ratio	p
Within cells	32	11933.22		
Activation (A)	1	192.76	.02	.900
A x S	1	13.49	.00	.973
Within Cells	32	8300.30		
A x D	1	26.94	.00	.955
A x S x D	1	211.63	.03	.874
Within cells	64	848.00		
A x B	2	1629.46	1.92	.155
A x S x B	2	1953.14	2.30	.108
Within cells	64	398.03		
A x D x B	2	837.42	2.10	.130
A x S x D x B	2	977.39	2.46	.094

ANOVA summary table for navigational efficiency for high and low WM activation groups				
Variance	df	MS	F Ratio	p
Within cells	32	.09		
WM Activation (A)	1	.00	.02	.902
A x S	1	.02	.19	.669
Within Cells	32	.05		
A x D	1	.01	.13	.721
A x S x D	1	.02	.45	.506
Within cells	64	.02		
A x B	2	.04	2.29	.110
A x S x B	2	.04	2.37	.102
Within cells	64	.01		
A x D x B	2	.00	.25	.780
A x S x D x B	2	.02	1.52	.226

ANOVA summary table for response times for high and low WM efficiency groups				
Variance	df	MS	F Ratio	p
Within cells	32	12228.81		
WM Efficiency (E)	1	692.32	.06	.813
E x S	1	2328.98	.19	.666
Within Cells	32	8555.29		
E x D	1	10.87	.00	.972
E x S x D	1	576.07	.07	.797
Within cells	64	868.99		
E x B	2	1721.00	1.98	.147
E x S x B	2	1911.34	2.20	.119
Within cells	64	363.09		
E x D x B	2	1767.70	4.87	.011
E x S x D x B	2	1483.76	4.09	.022

ANOVA summary table for navigational efficiency for high and low WM efficiency groups				
Variance	df	MS	F Ratio	p
Within cells	32	.09		
WM Efficiency (E)	1	.04	.50	.483
E x S	1	.03	.35	.556
Within Cells	32	.05		
E x D	1	.01	.21	.653
E x S x D	1	.07	1.56	.221
Within cells	64	.02		
E x B	2	.00	.11	.895
E x S x B	2	.01	.67	.517
Within cells	64	.01		
E x D x B	2	.02	1.52	.226
E x S x D x B	2	.02	1.81	.172

Appendix - Chapter 5

ANOVA summary table: The effects of command type upon response times for the command recognition task.				
Variance	df	MS	F Ratio	p
Within cells	189	21690.68		
Command type	3	3910207.1	180.27	<.001

ANOVA summary table: The effects of command type upon correct responses for the command recognition task.				
Variance	df	MS	F Ratio	p
Within cells	189	2.65		
Command type	3	49.67	18.74	<.001

ANOVA summary table: The effects of command type upon false positives for the command recognition task.				
Variance	df	MS	F Ratio	p
Within cells	189	16.96		
Command type	3	223.33	13.16	<.001

ANOVA summary table: The effects of parameter condition upon response times for the command line interface.				
Variance	df	MS	F Ratio	p
Within cells	189	1873852.7		
No. of Parameters (P)	3	435808676	232.57	<.001

ANOVA summary table: The effects of parameter condition upon response times for the menu interface.				
Variance	df	MS	F Ratio	p
Within cells	189	1106118.2		
No. of Parameters (P)	3	345045626	311.94	<.001

ANOVA summary table: The effects of parameter condition upon response times for the command line interface adjusted for typing speed.				
Variance	df	MS	F Ratio	p
Within cells	189	1300456.0		
No. of Parameters (P)	3	112388756	86.42	<.001

ANOVA summary table: The main and interactive effects of command method upon response times.				
Variance	df	MS	F Ratio	p
Within cells	63	4673068.7		
Command (C)	1	119323249	25.53	<.001
Within cells	189	890782.82		
C x P	3	5068383.8	5.69	.001

ANOVA summary table: The effects of expetise upon response times for the command line interface.				
Variance	df	MS	F Ratio	p
Within cells	62	18124987		
Expertise (E)	1	212136857	11.70	.001
Within cells	186	1715712.2		
E x P	3	11678563	6.81	<.001

ANOVA summary table: The effects of expertise upon response times for the menu interface.				
Variance	df	MS	F Ratio	p
Within cells	62	13794908		
Expertise (E)	1	16826568	12.20	.001
Within cells	186	1010059.1		
E x P	3	7061782.8	6.99	<.001

ANOVA summary table: The effects of expetise upon response times for the command line interface adjusted for typing speed.				
Variance	df	MS	F Ratio	p
Within cells	62	6621019.2		
Expertise (E)	1	25833249	3.90	.053
Within cells	186	1267530.8		
E x P	3	3341815.4	2.64	.051

ANOVA summary table: The interactive effects of expertise and command method upon response times.				
Variance	df	MS	F Ratio	p
Within cells	62	4727967.8		
E x C	1	1269322.4	.27	.606
Within cells	186	888652.74		
E x C x P	3	1022848.1	1.15	.330

ANOVA summary table: The effects of computer literacy upon response times for the command line interface.				
Variance	df	MS	F Ratio	p
Within cells	62	20821207		
Compiter literacy (LIT)	1	44971168	2.16	.147
Within cells	186	1846582.7		
LIT x P	3	3564592.2	1.93	.126

ANOVA summary table: The effects of computer literacy upon response times for the menu interface.				
Variance	df	MS	F Ratio	p
Within cells	62	15435397		
Computer literacy (LIT)	1	66552264	4.31	.042
Within cells	186	1021762.3		
LIT x P	3	6336183.5	6.20	<.001

ANOVA summary table: The effects of computer literacy upon response times for the command line interface adjusted for typing speed.				
Variance	df	MS	F Ratio	p
Within cells	62	7022852.2		
Compiter literacy (LIT)	1	919602.77	.13	.719
Within cells	186	1311343.9		
LIT x P	3	625402.06	.48	.699

ANOVA summary table: The interactive effects of computer literacy and command method upon response times.				
Variance	df	MS	F Ratio	p
Within cells	62	4731440.6		
LIT x C	1	1054010.4	.22	.639
Within cells	186	893537.01		
LIT x C x P	3	720023.01	.81	.492

ANOVA summary table: The effects of spatial visualisation upon response times for the command line interface.				
Variance	df	MS	F Ratio	p
Within cells	60	14873258		
Spatial visualisation (SV)	1	223371236	15.02	<.001
SV x E	1	7157927.8	.48	.491
Within cells	180	1546811.0		
SV x P	3	11468092	7.41	<.001
SV x E x P	3	2029493.1	1.31	.272

ANOVA summary table: The effects of spatial visualisation upon response times for the menu interface.				
Variance	df	MS	F Ratio	p
Within cells	60	12094633		
Spatial visualisation (SV)	1	127883525	10.57	.002
SV x E	1	1438013.1	.12	.731
Within cells	180	983397.30		
SV x P	3	2977980.4	3.03	.031
SV x E x P	3	615048.62	.63	.599

ANOVA summary table: The effects of spatial visualisation upon response times for the command line interface adjusted for typing speed.				
Variance	df	MS	F Ratio	p
Within cells	60	6122960.6		
Spatial visualisation (SV)	1	38947969	6.36	.014
SV x E	1	4436685.0	.72	.398
Within cells	180	1247576.3		
SV x P	3	2527921.0	2.03	.112
SV x E x P	3	1201793.9	.96	.411

ANOVA summary table: The interactive effects of spatial visualisation and command method upon response times.				
Variance	df	MS	F Ratio	p
Within cells	60	4756261.5		
SV x C	1	6614064.4	1.39	.243
SV x E x C	1	1089669.7	.23	.634
Within cells	180	874705.90		
SV x C x P	3	2010945.2	2.30	.079
SV x E x C x P	3	598891.12	.68	.562

ANOVA summary table: The effects of spatial memory upon response times for the command line interface.				
Variance	df	MS	F Ratio	p
Within cells	60	15763010		
Spatial memory (SM)	1	104121310	6.61	.013
SM x E	1	73847256	4.68	.034
Within cells	180	1504581.3		
SM x P	3	6934509.9	4.61	.004
SM x E x P	3	9164772.3	6.09	.001

ANOVA summary table: The effects of spatial memory upon response times for the menu interface.				
Variance	df	MS	F Ratio	p
Within cells	60	12114101		
Spatial memory (SM)	1	77194998	6.37	.014
SM x E	1	51243233	4.23	.044
Within cells	180	942444.19		
SM x P	3	3619911.1	3.84	.011
SM x E x P	3	2457101.6	2.61	.053

ANOVA summary table: The effects of spatial memory upon response times for the command line interface adjusted for typing speed.				
Variance	df	MS	F Ratio	p
Within cells	60	6150944.4		
spatial memory (SM)	1	21685090	3.53	.065
SM x E	1	19761432	3.21	.078
Within cells	180	1192178.1		
SM x P	3	2316361.1	1.94	.124
SM x E x P	3	4739863.0	3.98	.009

ANOVA summary table: The interactive effects of spatial memory and command method upon response times.				
Variance	df	MS	F Ratio	p
Within cells	60	4851652.3		
SM x C	1	1005243.8	.21	.651
SM x E x C	1	1029619.3	.21	.647
Within cells	180	886879.21		
SM x C x P	3	679147.06	.77	.515
SM x E x C x P	3	1204569.9	1.36	.257

ANOVA summary table: The effects of verbal ability (Nelson-Denny vocabulary test) upon response times for the command line interface.				
Variance	df	MS	F Ratio	p
Within cells	60	18203166		
Nelson-Denny (ND)	1	29484528	1.62	.208
ND x E	1	1920696.0	.11	.746
Within cells	180	1709526.8		
ND x P	3	1162422.5	.68	.565
ND x E x P	3	2614614.8	1.53	.208

ANOVA summary table: The effects of verbal ability (Nelson-Denny vocabulary test) upon response times for the menu interface.				
Variance	df	MS	F Ratio	p
Within cells	60	13963682		
Nelson-Denny (ND)	1	14919184	1.07	.305
ND x E	1	2668865.4	.19	.664
Within cells	180	1009149.8		
ND x P	3	1651252.9	1.64	.183
ND x E x P	3	437257.08	.43	.729

ANOVA summary table: The effects of verbal ability (Nelson-Denny vocabulary test) upon response times for the command line interface adjusted for typing speed.				
Variance	df	MS	F Ratio	p
Within cells	60	6724289.5		
Nelson-Denny (ND)	1	2547852.7	.38	.541
ND x E	1	4429928.7	.66	.420
Within cells	180	1257709.3		
ND x P	3	97070.24	.08	.972
ND x E x P	3	3021587.4	2.40	.069

ANOVA summary table: The interactive effects of verbal ability (Nelson-Denny vocabulary test) and command method upon response times.				
Variance	df	MS	F Ratio	p
Within cells	60	4788311.8		
ND x C	1	1228418.1	.26	.614
ND x E x C	1	4558865.4	.95	.333
Within cells	180	875448.96		
ND x C x P	3	397685.14	.45	.715
ND x E x C x P	3	2160043.5	2.47	.064

ANOVA summary table: The effects of logical reasoning upon response times for the command line interface.				
Variance	df	MS	F Ratio	p
Within cells	60	18596113		
Logical reasoning (LR)	1	1153209.6	.06	.804
LR x E	1	6908648.0	.37	.544
Within cells	180	1760985.2		
LR x P	3	212298.38	.12	.948
LR x E x P	3	511646.73	.29	.832

ANOVA summary table: The effects of logical reasoning upon response times for the menu interface.				
Variance	df	MS	F Ratio	p
Within cells	60	14032802		
Logical reasoning (LR)	1	996689.06	.07	.791
LR x E	1	12216407	.87	.355
Within cells	180	1029881.2		
LR x P	3	460294.68	.45	.720
LR x E x P	3	376234.87	.37	.778

ANOVA summary table: The effects of logical reasoning upon response times for the command line interface adjusted for typing speed.				
Variance	df	MS	F Ratio	p
Within cells	60	6682442.3		
Logical reasoning (LR)	1	2288941.7	.34	.561
LR x E	1	7383920.9	1.10	.297
Within cells	180	1298760.3		
LR x P	3	174937.37	.13	.939
LR x E x P	3	493895.14	.38	.767

ANOVA summary table: The interactive effects of logical reasoning and command method upon response times.				
Variance	df	MS	F Ratio	p
Within cells	60	4843082.9		
LR x C	1	2147046.1	.44	.508
LR x E x C	1	375642.56	.08	.782
Within cells	180	898638.67		
LR x C x P	3	458361.86	.51	.676
LR x E x C x P	3	731060.90	.81	.488

ANOVA summary table: The effects of associative memory upon response times for the command line interface.				
Variance	df	MS	F Ratio	p
Within cells	60	17385578		
Associative memory (MA)	1	80612186	4.64	.035
MA x E	1	36483.68	.00	.964
Within cells	180	1632460.3		
MA x P	3	8208825.3	5.03	.002
MA x E x P	3	235127.18	.14	.933

ANOVA summary table: The effects of associative memory upon response times for the menu interface.				
Variance	df	MS	F Ratio	p
Within cells	60	12876632		
Associative memory (MA)	1	26454976	2.05	.157
MA x E	1	54994690	4.27	.043
Within cells	180	990037.37		
MA x P	3	1694281.7	1.71	.166
MA x E x P	3	1478573.2	1.49	.218

ANOVA summary table: The effects of associative memory upon response times for the command line interface adjusted for typing speed.				
Variance	df	MS	F Ratio	p
Within cells	60	6165048.5		
Associative memory (MA)	1	29227685	4.74	.033
MA x E	1	11958141	1.94	.196
Within cells	180	1217530.1		
MA x P	3	4561218.8	3.75	.012
MA x E x P	3	1031322.6	.85	.470

ANOVA summary table: The interactive effects of associative memory and command method upon response times.				
Variance	df	MS	F Ratio	p
Within cells	60	4288402.8		
MA x C	1	7353569.7	1.71	.195
MA x E x C	1	28932064	6.75	.012
Within cells	180	879575.66		
MA x C x P	3	1343371.8	1.53	.209
MA x E x C x P	3	1013248.0	1.15	.330

ANOVA summary table: The interactive effects of parameters required and command method upon errors.				
Variance	df	MS	F Ratio	p
Within cells	63	.01		
C	1	.12	20.32	<.001
Within cells	189	.01		
P	3	.02	4.79	.003
Within cells	189	.00		
C x P	3	.01	3.02	.031

ANOVA summary table: The effects of self-report workload.				
Variance	df	MS	F Ratio	p
Within cells	63	82.03		
C	1	111.63	1.35	.250

ANOVA summary table: The effects of expertise upon self-report workload.				
Variance	df	MS	F Ratio	p
Within cells	62	432.38		
E	1	2.44	.01	.940
Within cells	62	82.56		
E x C	1	49.58	.60	.441

ANOVA summary table: The effects of computer literacy upon self-report workload.				
Variance	df	MS	F Ratio	p
Within cells	62	420.74		
LIT	1	724.14	1.72	.194
Within cells	62	81.70		
LIT x C	1	102.73	1.26	.266

ANOVA summary table: The interactive effects of spatial visualisation upon self-report workload.				
Variance	df	MS	F Ratio	p
Within cells	60	446.00		
SV	1	34.36	.08	.782
SV x E	1	13.29	.03	.864
Within cells	60	80.93		
SV x C	1	260.98	3.23	.078
SV x E x C	1	2.48	.03	.862

ANOVA summary table: The interactive effects of spatial memory upon self-report workload.				
Variance	df	MS	F Ratio	p
Within cells	60	427.73		
SM	1	833.00	1.95	.168
SM x E	1	310.79	.73	.397
Within cells	60	80.25		
SM x C	1	301.24	3.75	.057
SM x E x C	1	2.24	.03	.868

ANOVA summary table: The interactive effects of verbal ability (Nelson-Denny vocabulary test) upon self-report workload.				
Variance	df	MS	F Ratio	p
Within cells	60	436.62		
ND	1	136.40	.31	.578
ND x E	1	468.37	1.07	.304
Within cells	60	84.62		
ND x C	1	28.64	.34	.563
ND x E x C	1	12.50	.15	.702

ANOVA summary table: The interactive effects of logical reasoning upon self-report workload.				
Variance	df	MS	F Ratio	p
Within cells	60	384.71		
LR	1	3561.33	9.62	.003
LR x E	1	142.30	.37	.545
Within cells	60	81.84		
LR x C	1	11.91	.15	.704
LR x E x C	1	197.36	2.41	.126

ANOVA summary table: The interactive effects of associative memory upon self-report workload.				
Variance	df	MS	F Ratio	p
Within cells	60	445.43		
MA	1	53.33	.12	.731
MA x E	1	29.31	.07	.798
Within cells	60	80.62		
MA x C	1	270.98	3.36	.072
MA x E x C	1	12.22	.15	.698

ANOVA summary table: Command selection strategy.				
Variance	df	MS	F Ratio	p
Within cells	63	119.35		
C	1	1424.45	11.94	.001
Within cells	189	.00		
P	3	.00		
Within cells	189	3.16		
C x P	3	7.65	2.42	.067

ANOVA summary table: The effects of expertise upon command selection strategy.				
Variance	df	MS	F Ratio	p
Within cells	62	58.53		
E	1	130.82	2.24	.140
Within cells	186	1.54		
E x P	3	4.11	2.67	.049

ANOVA summary table: The effects of computer literacy upon command selection strategy.				
Variance	df	MS	F Ratio	p
Within cells	62	60.35		
LIT	1	17.90	.30	.588
Within cells	186	1.54		
LIT x P	3	3.85	2.50	.061

ANOVA summary table: The effects of spatial visualisation upon command selection strategy.				
Variance	df	MS	F Ratio	p
Within cells	60	59.19		
SV	1	21.32	.36	.551
SV x E	1	56.50	.95	.333
Within cells	180	1.57		
SV x P	3	.26	.17	.918
SV x E x P	3	.77	.49	.692

ANOVA summary table: The effects of spatial memory upon command selection strategy.				
Variance	df	MS	F Ratio	p
Within cells	60	57.83		
SM	1	157.55	2.72	.104
SM x E	1	1.52	.03	.872
Within cells	180	1.57		
SM x P	3	1.20	.77	.513
SM x E x P	3	.28	.18	.909

ANOVA summary table: The effects of spatial verbal ability (Nelson-Denny vocabulary test) upon command selection strategy.				
Variance	df	MS	F Ratio	p
Within cells	60	58.40		
ND	1	83.17	1.42	.237
ND x E	1	40.07	.69	.411
Within cells	180	1.56		
ND x P	3	.82	.53	.663
ND x E x P	3	1.15	.74	.529

ANOVA summary table: The effects of logical reasoning upon command selection strategy.				
Variance	df	MS	F Ratio	p
Within cells	60	59.94		
LR	1	29.50	.49	.486
LR x E	1	3.07	.05	.822
Within cells	180	1.54		
LR x P	3	1.22	.79	.498
LR x E x P	3	1.84	1.20	.313

ANOVA summary table: The effects of associative memory upon command selection strategy.				
Variance	df	MS	F Ratio	p
Within cells	60	53.83		
MA	1	93.56	1.74	.192
MA x E	1	310.89	5.78	.019
Within cells	180	1.57		
MA x P	3	1.27	.81	.489
MA x E x P	3	.13	.08	.971

Appendix - Chapter 6

ANOVA summary tables

ANOVA summary table: The effects of field type and stimulus type upon response times.

Variance	df	MS	F Ratio	p
Within cells	47	216.25		
Field type (F)	1	19805.18	91.58	<.001
Within cells	47	77.86		
Stimulus type (S)	1	39.69	.51	.479
Within cells	47	98.41		
F x S	1	861.55	8.75	.005

ANOVA summary table: The effects of field type and stimulus type upon self-report workload.

Variance	df	MS	F Ratio	p
Within cells	47	59.29		
Field type (F)	1	1284.78	21.67	<.001
Within cells	47	67.81		
Stimulus type (S)	1	17.52	.26	.614
Within cells	47	49.63		
F x S	1	498.37	10.04	.003

ANOVA summary table: The effects of field type and stimulus type upon errors.

Variance	df	MS	F Ratio	p
Within cells	47	.00		
Field type (F)	1	.00	3.45	.070
Within cells	47	.00		
Stimulus type (S)	1	.00	3.07	.086
Within cells	47	.00		
F x S	1	.00	.00	1.000

Regression Equation Summaries

Regression of response time upon stimulus and field factors, and spatial visualisation scores, controlling for mouse performance.

Variance	R square	SS	df	p
Mouse	.203	17759.46	1	
SV	.241	21069.96	1	
Between Subs	.552	48172.15	47	
Stimulus (S)	.552	48211.83	1	
SV x S	.553	48264.68	1	ns
subs (S)	.594	51871.90	47	
Field (F)	.821	71676.18	1	
SV x F	.821	71711.63	1	ns
subs (F)	.937	81839.36	47	
S x F	.947	82700.93	1	
SV x S x F	.947	82732.92	1	ns
Residual		4592.59	47	

Regression of response time upon stimulus and field factors, and verbal ability (Nelson-Denny vocabulary test scores), controlling for mouse performance.				
Variance	R square	SS	df	p
Mouse	.203	17759.46	1	
ND	.215	18795.61	1	
Between Subs	.552	48172.15	47	
Stimulus (S)	.552	48211.83	1	
ND x S	.552	48215.36	1	ns
subs (S)	.594	51871.90	47	
Field (F)	.821	71676.18	1	
ND x F	.826	72124.88	1	ns
subs (F)	.937	81839.36	47	
S x F	.973	82700.93	1	
ND x S x F	.947	82704.03	1	ns
Residual		4621.48	47	

Regression of self-report workload upon stimulus and field factors, and spatial visualisation scores, controlling for mouse performance.				
Variance	R square	SS	df	p
Mouse	.411	20063.78	1	
SV	.423	20628.44	1	
Between Subs	.793	38700.12	47	
Stimulus (S)	.793	38717.64	1	
SV x S	.793	38722.17	1	ns
subs (S)	.859	41904.85	47	
Field (F)	.885	43189.29	1	
SV x F	.886	43225.32	1	ns
subs (F)	.942	45975.88	47	
S x F	.952	46474.34	1	
SV x S x F	952	46481.40	1	ns
Residual		2324.84	47	

Regression of self-report workload upon stimulus and field factors, and verbal ability (Nelson-Denny vocabulary test scores), controlling for mouse performance.				
Variance	R square	SS	df	p
Mouse	.411	20063.78	1	
ND	.422	20601.82	1	
Between Subs	.792	38700.12	47	
Stimulus (S)	.793	38717.64	1	
ND x S	.794	38735.79	1	ns
subs (S)	.859	41904.85	47	
Field (F)	.885	43189.29	1	
ND x F	.885	43212.09	1	ns
subs (F)	.942	45975.88	47	
S x F	.952	46474.34	1	
ND x S x F	.952	46485.13	1	ns
Residual		2321.11	47	

Appendix - Chapter 7

Regression Equation Summaries for Component Tasks

Regression of response time for the explicit menu component task upon age group and spatial visualisation.				
Variance	R square	SS	df	p
Age (A)	.414	9.70	1	
SV	.432	10.12	1	
A x SV	.432	10.13	1	ns
Residual		13.30	39	

Regression of response time for the explicit menu component task upon age group and spatial memory.				
Variance	R square	SS	df	p
Age (A)	.414	9.70	1	
SM	.522	12.22	1	
A x SM	.566	13.26	1	ns
Residual		10.17	39	

Regression of response time for the explicit menu component task upon age group and verbal ability (Nelson-Denny vocabulary test).				
Variance	R square	SS	df	p
Age (A)	.414	9.70	1	
ND	.442	10.35	1	
A x ND	.448	10.50	1	ns
Residual		12.93	39	

Regression of response time for the explicit menu component task upon age group and logical reasoning.				
Variance	R square	SS	df	p
Age (A)	.414	9.70	1	
LR	.467	10.95	1	
A x LR	.469	10.99	1	ns
Residual		12.44	39	

Regression of response time for the explicit menu component task upon age group and associative memory.				
Variance	R square	SS	df	p
Age (A)	.414	9.70	1	
MA	.423	9.91	1	
A x MA	.472	11.06	1	ns
Residual		12.37	39	

Regression of response time for the embedded menu component task upon age group and spatial visualisation.				
Variance	R square	SS	df	p
Age (A)	.384	290.77	1	
SV	.397	300.55	1	
A x SV	.463	349.93	1	<.05
Residual		406.35	41	

Regression of response time for the embedded menu component task upon age group and spatial memory.				
Variance	R square	SS	df	p
Age (A)	.384	290.77	1	
SM	.499	377.12	1	
A x SM	.636	480.81	1	<.001
Residual		275.47	41	

Regression of response time for the embedded menu component task upon age group and verbal ability (Nelson-Denny reading test).				
Variance	R square	SS	df	p
Age (A)	.384	290.77	1	
ND	.404	305.63	1	
A x ND	.417	315.63	1	ns
Residual		440.64	41	

Regression of response time for the embedded menu component task upon age group and logical reasoning.				
Variance	R square	SS	df	p
Age (A)	.384	290.77	1	
LR	.399	301.88	1	
A x LR	.399	301.88	1	ns
Residual		454.40	41	

Regression of response time for the embedded menu component task upon age group and associative memory.				
Variance	R square	SS	df	p
Age (A)	.384	290.77	1	
MA	.402	303.89	1	
A x MA	.471	356.05	1	<.05
Residual		400.22	41	

Anova Summary Tables for Information Retrieval Task

ANOVA summary table: The effects of interface condition and age group upon response times.				
Variance	df	MS	F Ratio	p
Main Effects	4	109240.802	11.826	<.001
Age (A)	1	127757.618	13.831	.001
Interface (I)	3	103068.531	11.158	<.001
2-Way Interactions	3	6514.294	.705	.555
A x I	3	6514.294	.705	.555
Explained	7	65215.156	7.060	<.001
Residual	37	9237.093		
Total	44	18142.694		

ANOVA summary table: The effects of interface condition and age group upon navigational efficiency.				
Variance	df	MS	F Ratio	p
Main Effects	4	.127	2.415	.066
Age (A)	1	.005	.096	.759
Interface (I)	3	.168	3.188	.035
2-Way Interactions	3	.035	.669	.576
A x I	3	.035	.669	.576
Explained	7	.088	1.667	.148
Residual	37	.053		
Total	44	.058		

ANOVA summary table: The effects of performance block and age group upon response times.				
Variance	df	MS	F Ratio	p
Within cells	43	51620.65		
Age (A)	1	668658.51	12.95	.001
Within cells	172	13508.28		
Block (B)	4	146330.63	10.83	<.001
A x B	4	62058.51	4.59	.002

ANOVA summary table: The effects of performance block and age group upon navigational efficiency.				
Variance	df	MS	F Ratio	p
Within cells	43	.18		
Age (A)	1	.07	.41	.526
Within cells	172	.13		
Block (B)	4	.57	4.27	.003
A x B	4	.02	.14	.967

ANOVA summary table: The effects of performance block and interface condition upon response times.				
Variance	df	MS	F Ratio	p
Within cells	41	47148.69		
I	3	318416.59	6.75	.001
Within cells	164	15166.23		
B	4	145186.78	9.57	<.001
I x B	12	7033.04	.46	.933

ANOVA summary table: The effects of performance block and interface condition upon navigational efficiency.				
Variance	df	MS	F Ratio	p
Within cells	41	.15		
I	3	.44	2.83	.050
Within cells	164	.14		
B	4	.59	4.34	.002
I x B	12	.08	.57	.863

ANOVA summary table: The effects of performance block, interface condition and age group upon response accuracy.				
Variance	df	MS	F Ratio	p
Within cells	37	.18		
Age group (A)	1	.01	.05	.816
Interface condition (I)	3	.28	1.55	.218
A x I	3	.05	.27	.846
Within cells	148	.21		
Block (B)	4	.65	3.13	.017
A x B	4	.24	1.17	.325
I x B	12	.21	1.03	.425
A x I x B	12	.32	1.55	.111

Regression Equation Summaries for Information Retrieval Task

Regression of response time upon linear, quadratic, and cubic, terms for spatial visualisation.				
Variance	R square	SS	df	p
SV	.055	43900.05	1	ns
SV^2	.069	55177.58	1	ns
SV^3	.087	69088.33	1	ns
Residual		729193.49	41	

Regression of response time upon linear, quadratic, and cubic, terms for spatial memory.				
Variance	R square	SS	df	p
SM	.100	79550.11	1	<.05
SM^2	.331	87506.12	1	
SM^3	368	108070.00	1	
Residual		690211.82	41	

Regression of response time upon linear, quadratic, and cubic, terms for verbal ability (Nelson-Denny vocabulary test).				
Variance	R square	SS	df	p
ND	.00	879.38	1	ns
ND^2	.014	11198.70	1	ns
ND^3	.02	15942.92	1	ns
Residual		782338.89	41	

Regression of response time upon linear, quadratic, and cubic, terms for logical reasoning.				
Variance	R square	SS	df	p
LR	.062	49510.66	1	ns
LR^2	.070	55733.42	1	ns
LR^3	.077	61391.43	1	ns
Residual		736890.39	41	

Regression of response time upon linear, quadratic, and cubic, terms for associative memory.				
Variance	R square	SS	df	p
MA	.031	24615.65	1	ns
MA^2	.031	24627.01	1	ns
MA^3	.032	25781.35	1	ns
Residual		772500.46	41	

Regression of response time upon interface condition and spatial visualisation.				
Variance	R square	SS	df	p
SV	.055	43900.05	1	
Interface (I)	.412	328966.16	3	
SV x I	.422	337017.25	3	ns
Residual		461264.56	37	

Regression of navigational efficiency upon interface condition and spatial visualisation.				
Variance	R square	SS	df	p
SV	.000	.00	1	
Interface (I)	.197	.51	3	
SV x I	.271	.70	3	ns
Residual		1.88	37	

Regression of response accuracy upon interface condition and spatial visualisation.				
Variance	R square	SS	df	p
SV	.001	.029	1	
Interface (I)	.111	4.27	3	
SV x I	.301	11.60	3	<.05
Residual		26.98	37	

Regression of response time upon interface condition and spatial memory.				
Variance	R square	SS	df	p
SM	.100	79550.12	1	
Interface (I)	.431	344446.18	3	
SM x I	.453	361698.42	3	ns
Residual		436583.39	37	

Regression of navigational efficiency upon interface condition and spatial memory.				
Variance	R square	SS	df	p
SM	.00	.01	1	
Interface (I)	.198	.51	3	
SM x I	.217	.56	3	ns
Residual		2.01	37	

Regression of response accuracy upon interface condition and spatial memory.				
Variance	R square	SS	df	p
SM	.006	.23	1	
Interface (I)	.111	4.27	3	
SM x I	.166	6.41	3	ns
Residual		32.17	37	

Regression of response time upon interface condition and verbal ability (Nelson-Denny vocabulary test).				
Variance	R square	SS	df	p
ND	.001	879.38	1	
Interface (I)	.382	304608.65	3	
ND x I	.403	321348.74	3	ns
Residual		476933.07	37	

Regression of navigational efficiency upon interface condition and verbal ability (Nelson-Denny vocabulary test).				
Variance	R square	SS	df	p
ND	.012	.032	1	
Interface (I)	.226	.582	3	
ND x I	.259	.666	3	ns
Residual		1.907	37	

Regression of response accuracy upon interface condition and verbal ability (Nelson-Denny vocabulary test).				
Variance	R square	SS	df	p
ND	.031	1.19	1	
Interface (I)	.129	4.99	3	
ND x I	.258	9.96	3	ns
Residual		28.62	37	

Regression of response time upon interface condition and logical reasoning.				
Variance	R square	SS	df	p
LR	.062	49510.66	1	
Interface (I)	.420	335442.13	3	
LR x I	.429	342119.26	3	ns
Residual		456162.59	37	

Regression of navigational efficiency upon interface condition and logical reasoning.				
Variance	R square	SS	df	p
LR	.149	.057	1	
Interface (I)	.208	.534	3	
LR x I	.281	.724	3	ns
Residual		1.849	37	

Regression of response accuracy upon interface condition and logical reasoning.				
Variance	R square	SS	df	p
LR	.050	.095	1	
Interface (I)	.118	4.54	3	
LR x I	.219	8.45	3	ns
Residual		30.13	37	

Regression of response time upon interface condition and associative memory.				
Variance	R square	SS	df	p
MA	.031	24615.65	1	
Interface (I)	.391	311843.43	3	
MA x I	.447	357195.39	3	ns
Residual		441086.42	37	

Regression of navigational efficiency upon interface condition and associative memory.				
Variance	R square	SS	df	p
MA	.006	.02	1	
Interface (I)	.221	.568	3	
MA x I	.276	.71	3	ns
Residual		1.86	37	

Regression of response accuracy upon interface condition and associative memory.				
Variance	R square	SS	df	p
MA	.005	.203	1	
Interface (I)	.112	4.316	3	
MA x I	.189	7.28	3	ns
Residual		31.29	37	

Regression of response time upon age group and spatial visualisation.				
Variance	R square	SS	df	p
Age (A)	.160	127759.96	1	
SV	.181	144549.30	1	
A x SV	.182	144968.63	1	ns
Residual		653313.17	41	

Regression of navigational efficiency upon age group and spatial visualisation.				
Variance	R square	SS	df	p
Age (A)	.002	.00	1	
SV	.003	.01	1	
A x SV	.073	.19	1	ns
Residual		2.39	41	

Regression of response accuracy upon age group and spatial visualisation.				
Variance	R square	SS	df	p
Age (A)	.002	.07	1	
SV	.003	.13	1	
A x SV	.010	.39	1	ns
Residual		38.19	41	

Regression of response time upon age group and spatial memory.				
Variance	R square	SS	df	p
Age (A)	.160	127759.96	1	
SM	.231	184350.64	1	
A x SM	.246	196155.30	1	ns
Residual		602126.52	41	

Regression of navigational efficiency upon age group and spatial memory.				
Variance	R square	SS	df	p
Age (A)	.002	.00	1	
SM	.006	.02	1	
A x SM	.109	.28	1	<.05
Residual		2.29	41	

Regression of response accuracy upon age group and spatial memory.				
Variance	R square	SS	df	p
Age (A)	.002	.07	1	
SM	.009	.35	1	
A x SM	.019	.74	1	ns
Residual		37.84	41	

Regression of response time upon age group and verbal ability (Nelson-Denny vocabulary test).				
Variance	R square	SS	df	p
Age (A)	.160	127759.96	1	
ND	.228	182367.89	1	
A x ND	.233	185725.92	1	ns
Residual		612555.90	41	

Regression of navigational efficiency upon age group and verbal ability (Nelson-Denny vocabulary test).				
Variance	R square	SS	df	p
Age (A)	.002	.00	1	
ND	.025	.06	1	
A x ND	.025	.06	1	ns
Residual		2.51	41	

Regression of response accuracy upon age group and verbal ability (Nelson-Denny vocabulary test).				
Variance	R square	SS	df	p
Age (A)	.002	.07	1	
ND	.033	1.27	1	
A x ND	.036	1.39	1	ns
Residual		37.19	41	

Regression of response time upon age group and logical reasoning.				
Variance	R square	SS	df	p
Age (A)	.160	127759.96	1	
LR	.204	162600.27	1	
A x LR	.204	163071.29	1	ns
Residual		635210.52	41	

Regression of navigational efficiency upon age group and logical reasoning.				
Variance	R square	SS	df	p
Age (A)	.002	.00	1	
LR	.023	.06	1	
A x LR	.06	.16	1	ns
Residual		2.41	41	

Regression of response accuracy upon age group and logical reasoning.				
Variance	R square	SS	df	p
Age (A)	.002	.07	1	
LR	.005	.19	1	
A x LR	.09	3.30	1	ns
Residual		35.28	41	

Regression of response time upon age group and associative memory.				
Variance	R square	SS	df	p
Age (A)	.160	127759.96	1	
MA	.172	137198.72	1	
A x MA	.189	150559.10	1	ns
Residual		647722.72	41	

Regression of navigational efficiency upon age group and associative memory.				
Variance	R square	SS	df	p
Age (A)	.002	.00	1	
MA	.007	.02	1	
A x MA	.011	.03	1	ns
Residual		2.54	41	

Regression of response accuracy upon age group and associative memory.				
Variance	R square	SS	df	p
Age (A)	.002	.07	1	
MA	.009	.33	1	
A x MA	.043	1.65	1	ns
Residual		36.93	41	