Investigations of the Effects of Different Computer Input Methods on Man-Computer Interaction

by

Peter Robert Innocent, B.Sc., M.Sc., MBCS

A Thesis submitted for the Degree of Doctor of Philosophy (Ph.D.), to the Department of Human Sciences in April, 1977.

Supervisor; K.D. Eason, Department of Human Sciences

SUMMARY

The development of interactive man-computer systems is a design process wherein various alternatives must be considered from different points of view. In order to make design decisions, information guidelines are needed. Among the requirements are those for different input methods of computers. This thesis has the objective of providing information and guidelines on how different input methods affect man-computer interaction. The objective is reached through a number of stages: a review of literature; the development of a framework for investigation; deriving and testing experimental hypotheses, and discussing and presenting information for future researchers and designers.

The literature review shows that, despite a large variety in the number of input devices and how they are used, information is fragmented and incomplete and cannot be easily generalised. In particular, no studies were found of comparisons between different input methods when used in problem solving. The thesis proposes a descriptive model of interactive man-computer problem solving which was based on four models. These were: semiotics (the theory of signs), how people use a keyboard, human problem solving and computer processes. The complete model emphasises the role of input methods and was used to produce general hypotheses.

Pragmatic considerations resulted in a series of testable experimental hypotheses which were not systematically related to the general hypotheses. Five experiments are described which are independently reported and discussed in relation to the descriptive model. All the experiments were laboratory rather than field experiments and used the same basic designs. A range of input devices was used including light pen, special function keys, joystick and standard keyboard. The first two experiments used non-problem solving simple input tasks, the others used problem solving tasks. Apart from the input method, the variables examined were different for each experiment. They included the personality, general and specific experience of people and the effects of sub-optimum computer system characteristics such as unreliable long response times and lack of feedback.

The results are discussed in relation to the general hypotheses and to each other and the model is revised accordingly. The main conclusions are that: (i) human problem solving processes are affected by the method of putting information into the computer, (ii) the input method affects the information transfer from man to computer depending on complex interactions between the characteristics of computers, people and problems, (iii) the acceptability of different input methods is based on user judgements of the fastest then least error-prone input that is possible in the particular conditions of computer and task characteristics, (iv) the degree of effect of different factors on the input times of a user of a particular input method is of the same order as the effect due to different input methods, (v) the balance between central cognitive and peripheral sensory/motor processes plays a major role in explaining many of the affects found, and (vi) the approach taken in the thesis is useful in that it leads to a way of generalising and comparing results.

Recommendations are provided for human factors researchers by listing both general and testable research hypotheses. Recommendations are also provided for systems designers as to how to use and interpret the results. This is attempted by exploring the design process and by using an example.

LIST OF CONTENTS

CHAPTER 1

INTRODUCTION

CONTENTS

1. Introduction

The development of sophisticated information processing machines that has taken place over the last 30 or so years has been punctuated by major conceptual milestones. Among these are the concepts of programmable machines, time sharing and distributed intelligence. The advance of technology has kept pace with and partly inspired these developments by providing cheap, reliable and efficient hardware. Correspondingly, the diversity and number of applications has increased until presently there are very few aspects of living unaffected by computer information processing. There are many consequences of this. For example, people who use such machines can no longer be considered as a minor part of the population with particular characteristics. This makes it difficult for the designers of complex computer systems to predict the consequences of their design in terms of both performance and acceptability. The realisation of the visionary dream of Licklider in 1960 of man-computer symbiosis (wherein computers do routine work to free people for more creative thinking) is dependent on such predictivity.

Designers have long recognised that the main problems of arriving at a reliable model for such predictions arise because of the non-deterministic and adaptive behaviour of people. Such problems are particularly important in time sharing systems where the characteristic behaviour of individuals offsets total system performance (Wilkes (1970)).

More generally, the problems are important in the design and operation of real-time systems and are likely to be of even greater importance to the future according to a Department of Industry Report (1975) on the future of real-time technology.

The definition of a real-time system in that report has been used to specify the scope of this thesis. The definition applies to systems in which the computer forms an essential part, receiving stimuli from other parts and responding within the time constraint of the overall system. Thus, in order to design such a system, the time constraints need to be known. One of these is the limits of response of people at a terminal. But people adapt to the system, therefore the establishment of such limits is non-trivial and in the realm of the behavioural rather than the computer scientist.

In recent years, many different types of behavioural scientists have recognised the problems and have taken up the challenge of carrying on research into various aspects. (See, for example, Human Choice and Computers by E. Mumford and H. Sackman (1970).) One particular group, the human factors specialists, have a history of research into man-machine systems (of which man-computer systems are a special class) and their role in systems development has been defined (Meister (1973)).

Human factors researchers have recognised their role from the early days of time sharing (e.g. Licklider (1960), Simon (1966), Sackman (1970)). But this major challenge for human factors research, discussed by Nickerson (1969) has yet to be adequately met. This was pointed out and some reasons for it made clear by Prof. Shackel in his opening address to a conference on Human Factors in Man-Computer Interaction (e.g. Shackel (1976)).

A main criticism was that many reported human factors studies deal, not with the total real-time system but the sub-systems of it on the assumption that the results apply in the total system. Baker (1976) has made this criticism with respect to the design of military systems. Despite the fact that Nickerson (1969) and Carbonell (1967) have pointed a way towards improving the situation by focussing on the man and modelling his behaviour, progress has been slow. Shackel (1969) laid the foundations for the continuation of Carbonells' attempt in 1967 to model man-computer interaction by suggesting a taxonomy wherein various aspects of importance could be recognised and focussed upon. The ideas generated from that work were developed into a philosophy leading to an approach for research into man-computer interaction (MCI). The work of this thesis grew directly out of this approach which is described in Shackel (1969). Many studies have been and are being carried out into different aspects of MCI using the same approach (e.g. Stewart (1974), Eason (1976)). Of the many aspects involved, this thesis deals with the investigation of the usage of input devices for real-time problem solving by individuals.

The objective is to provide information of the time constraints imposed by input devices and behavioural consequences to enable the relative importance of the input sub-system to be made explicit in relation to the total man-computer system. The means of achieving this end is in a number of stages. These are:

(i) to review pertinent literature,
(ii) to propose and develop a framework for investigation,
(iii) to describe and discuss investigations in that framework,
(iv) to provide guidelines for system designers and future researchers.

2. Literature Review

The scope of the review is confined to those papers reporting quantitative data collected in a systematic way about the use of different input devices used with computers. The scope was wide in the sense that task, system, user and environmental differences were allowed. Even so, the number of papers found was small in relation to those differences. Most studies were carried out in the laboratory rather than the field and were comparing different input devices for some simple task. As Siebel (1972) has said, there is no data on data input tasks in problem solving. The review is presented according to a scheme based on simple data entry tasks. These are:-

(i) 'Marker' tasks:

These involve control of a reference mark on a visual display by means of an input device.

(ii) 'Pointer' tasks:

These involve specification of one of a small number displayed options by means of an input device.

(iii) Alphanumeric data entry tasks:

These involve transferring information to the computer by means of alphanumeric data entry devices.

(iv) Numeric data entry tasks:

As (iii) but restricted to numeric data.

(v) Graphic data and symbol entry tasks.

(vi) Voice input tasks.

Within these tasks, general literature rather than reported studies has been referred to in order to identify commonly used devices. In common use the single tasks are not independent. That is, pointer tasks can include marker tasks and alphanumeric data entry tasks can be part of pointer tasks. However, the classification scheme suits the reported literature rather than different user tasks in man-computer interaction.

2.1 Marker Tasks

Commonly used marker devices are:

(i) Joy-stick:

The basic principle is, that movement of a vertical lever is used to produce the mark controlling signal. The device may be operated by hand or knee and be stiff and thick,or thin and easily moved levers.

(ii) Rolling ball:

This is a hemispherical device such that the marker position can be altered according to the angular rotation of the hemisphere about its centre controlled by hand.

(iii) Mouse:

The mouse is moved by hand across a working plane. At any time, its position on the working plane is used to determine the position of the marker.

(iv) Light pen:

Light pens interact directly with a cathode ray tube display so that the position of the mark follows the pen on the display.

The above devices are commonly used in this way because of their ability to provide sufficiently accurate information to control the marker.

However, there is no reason why other devices such as keyboards should not be used as marker devices. The main factors influencing the speed and accuracy of use are the control/display relationship. Operator preferences are also important in the choice of a particular device as well as the task requirements for flexibility of input. No comparative study has been carried out on all these aspects and devices.

Jenkins and Kerr (1954) studied the various aspects of joysticks using a simulated visual display with discrete targets (0.25" wide) and a cursor mark (0.15" wide). They found that the optimum control/display movement ratio was 2.5 (or greater) measured at the top of the stick and at the display. The stick length and starting position were not important. Typical times and error rates for marking a target were:

Joystick entry times

	Mean time (seconds)	Standard Deviation	% Errors
Experienced Subjects	1.58	0.25	4.7
Inexperienced Subjects	1.68	0.30	7.2

Baker (1960, 1961) referred to by Sperrandio and Bisseret (1968) compared the joystick, rolling ball and light pen for plotting data points on a screen with the greatest accuracy. The light pen was fastest allowing 0.80 seconds per plot; the joystick next with 2.4 seconds/plot and finally the rolling ball with 2.9 seconds/plot. According to Thornton (1954) subjects who used the rolling ball before a 'joystick' prefer the rolling ball and vice-versa. All these studies emphasised the importance of control/display compatability in that joystick movement towards the operator must draw the marker down the display and left/right movement must mean left/right on the display.

In some marker tasks the mark may be a cursor to indicate a text area or character. Such a task was used by English, Englebart and Bremman (1967) to compare the joystick, mouse, grafacon and light pen. Their results are difficult to interpret since their analysis of the data did not include statistical tests. A selected summary of their findings is as follows.

	Mouse	Light Pen	Grafacon	Joystick	
All subjects, i.e. Experienced Subjects and Inexperienced Subjects	1.93s	2.13s	2.43s	2.87s	Characters select time (seconds)
Experienced Subjects only	0.93	0.201	0.208	0.278	Error rates

Subjects found the mouse was not as tiring as the light pen (whose accuracy was a function of screen luminosity). However, the light pen was easier to learn to use than the mouse. The experiment involved giving a penalty for error which slowed subjects down on the grafacon and joystick by ~9%, light pen by 4% and on the mouse by ~2%.

The accuracy required by the task was also varied by requiring characters rather than words (of 5 characters) to be selected. This again slowed down the subjects between by 20-40% depending on the device type.

Goodwin (1975) compared light pen, light gun (similar to light pen) and keyboard for marking different places in a text display. Three tasks were used; arbitrary cursor positioning, sequential cursor positioning and check reading. For each of these tasks the subject moved the cursor to a target character then overtyped it. In the 'arbitrary' task, 10 targets were randomly placed on the screen, and these could be overtyped in a random order; in the sequential task, the order had to be from top to bottom of the screen; in the check reading task, 10 substitute errors in text had to be corrected. A summary of their results was as follows:

Mean time to reach a target (seconds) in the arbitrary task

Light Pen	Light Gun	Keyboard
2.59s	3.21s	13.48s

The reason for long keyboard time was that the marker had to be moved along rows of the display by tabs or spaces and carriage return used to shift down the display; cursor keys were not used.

In a study of the effects of different cursor forms used for indicating the marked position and moved by cursor keys, Vartabedian (1970) found the cursor movement time achieved by subjects depended on the cursor type (box, underline, cross, channel) and its blink rate. The average times for operating the cursor keys are not given. These may be inferred if it is assumed that each subject took the shortest route (50 keypresses on average) and the movement time was about 0.25s (optimum) to 0.35s (sub-optimum). From Goodwin's work, we can infer that about 20-30 cursor keypresses were necessary to reach a target. The corresponding key-in time would be about 4-9 seconds. This is slower than the light pen or light gun.

Earl and Goff (1965) compared performances of the standard keyboard and the light pen used for entering words of 3 to 7 letters. Each letter had to be typed or pointed at. There was no significant difference in speed but the error rate of selection in pointing (~0.75%) was less than in typing (5.1%).

2.2 Pointer Tasks

Pointer tasks are used to provide the computer with one of a small number of options which are displayed during interaction. The display may be formatted in such a way that the options are presented as a vertical list (a 'menu'), horizontal lines or according to the relationships between the functions represented by the symbols. Symbols may change during interaction ('soft' symbols) or remain fixed ('hard' symbols).

Commonly used devices for pointer tasks using 'soft' symbols on a cathode ray tube display are:

(i) Light pens (see description and discussion on marker tasks).

(ii) Touch displays:

The display is modified by placing sensors on it which detect when a finger (or wand) is in contact and indicates to the computer which sensor has been activated. Symbols are allocated and displayed next to the positions of the sensors on the display by the computer. Johnson (1967) describes the use of this device in detail for air traffic control tasks. On the basis of experiments comparing touch displays with conventional keyboards, he concluded that touch displays were faster and more accurate than keyboards (no supportive data was supplied). However, there was a difference in coding such that one touch was equivalent to several keypresses. Hence the result is not unexpected.

Devices used for pointer tasks for hard symbols are:

(i) Touch boards or pads:

These consist of a number of small sensors mounted in some way and operated by touching with a finger (or wand). The sensor may be positioned according to the relationships between the symbols they represent. Overlays or masks may be used to change the meaning of each sensor. Usually operation of the device gives no proximal feedback although in some types, touch pad sensors areas light up on activation.

No comparative data exists on the devices used in pointer tasks although the results of work on marker devices and keyboards have some relevance.

(ii) Special function keyboards:

These are similar to touch boards and pads but use keys as sensors which normally have printed symbols on them.

(iii) Chord Keyboards:

A chord keyboard allows simultaneous operation of 2 or more keys to which can be allocated symbols. Seibel (1972) discusses the history and characteristics of use of chord keyboards at length. Among his findings were (a) that practised subjects can strike chords within 0.3 and 0.4 seconds after being shown which chord to strike; (b) there are relative differences in speed and accuracy of the chords of one hand; (c) that 'simultaneously' can mean less than 30 mS and that the number of different chords has little or no effect on speed of response for practised subjects. In essence chord keyboards may be used in a similar way to special function keyboards. There are many studies by Seibel (1962, 1964) which show the advantages, practice curves and so on of chord keyboards, compared with QWERTY standard keyboards. Most of these differences arise because of the allocations of the symbols to the chords (i.e. coding).

2.3 Alphanumeric Data Entry Tasks

The most commonly used device is the standard keyboard such as that presently used on office typewriters and sometimes referred to as 'QWERTY'. A great deal of research has been conducted into the characteristics of standard keyboards which effect their acceptability and performance in simple data entry tasks. The research has been organised around three main and complementary variables. These are the layout and physical characteristics of the keys, the key coding and the training of operators. The effects of these are reviewed by, for example, Stewart (1974), Sperrandio and Bisseret (1968), and Seibel (1972). It is not intended to provide a full review in this thesis, since most of this work has been carried out outside of the context of man-computer interactive problem solving.

A short summary is given as follows.

(i) Physical Characteristics and Layout:
The physical characteristics are such things as the size, shape, key displacement, spacing between keys, and provision of proximal feedback. These characteristics are discussed and reviewed by Stewart (1974) and the main conclusion was that most current light-action computer terminal keyboards have characteristics within the recommended range. Proximal feedback (e.g. key clicks on activation) did not affect the performance of experienced typists but affected the rate of learning by typists.

Keyboard layout is discussed by Seibel (1972) who points out that 'despite demonstrated advantages for other arrangements, the overall economics and re-training aspects strongly suggest that the QWERTY arrangement is the standard'. Other keyboard arrangements have been designed and evaluated (e.g. Dvorak (1943), Griffith (1949)) and their relative efficiency is a function of learning of the operator. Stewart (1974) points out that, although the 35 standard alphabetic characters have been accepted, the location of the other keys is still disputed, and computer terminals can have different standards. In such cases there may be negative transfer of training effects when different layouts are used and this could effect the evaluation of such layouts. This is particularly true when such evaluations take into account (a) the case of self-detected error corrections, (b) verification procedures, (c) ability to insert or delete parts of a message. Seibel (1972) again says there are no data for guidance in estimating the trade off functions involved.

(ii) Coding of keys:

With a standard keyboard, the data are input one character at a time. Coding is an attempt to increase the quantity of information input per elementary input (i.e. keypress). This may be achieved by either extracting the determining characters from the words in the data (extraction coding) or by replacing the words by other shorter words which can be easily memorised (mnemonic coding). The effect of coding on the rate of information transfer is therefore very substantial as shown by the comparative studies of Seibel (1964). The performance and acceptability of use of a particular device compared with another should therefore take coding differences into account. The rate of data entry decreases as the amount of information per entry increases. But with sufficient practice the effect disappears until the fastest rate is again reached (Conrad, (1962)). The slope of the learning curve depends on the coding compatability.

(iii) Training of Operators:

Operators almost always encode some data into 'chunks' (Miller (1956)) which are stored in short term memory. Whenever short term memory is involved some form of encoding is used and the chunk size increases with the training of the operator in that particular task. Effectively, this makes the task of data entry easier. Leonard and Newman (1965) have demonstrated this with typing tasks. This therefore affects performance measures of rates of input of information.

Operators must also be trained to use keyboards whose keys are encoded in a particular way. This involves learning the coding system and the special motor responses that go with it.

The effects of training are large and dependent on particular aspects such as the coding of keys, their layout and encoding of source data.

2.4 Numeric Data Entry Tasks

The most common form of numeric data entry is through keyboards. Though many different digit layouts have been investigated, there are four basic layouts:

(i) "1 2 3" or telephone layout

In this layout, the digits are arranged in rows with 1 2 3 on the top and 7 8 9 on the bottom.

(ii) "7 8 9" or add listing layout

Opposite to 1 2 3

(iii) Typewriter layout:

The digits appear in ascending order left to right above the top row of the alphabetic keys.

(iv) Adding Machine Layout:

The digits are replicated in a matrix such that rows have particular significance (e.g. x 10) and columns a multiplying factor.

Conrad and Hull (1968) compared 123 and 789 layouts with naive subjects and found that the 123 conformed more to subjective expectations and was used significantly more accurately than the 789 layout. Entry rates were about 1 digit/second with an error rate of $\lesssim$ 1%. Other studies such as Conrad (1967) report 0.67 seconds per digit with an error rate of 0.55% for 123 as opposed to 0.74 secs/digit with 1.16% error for 789.

No comparative data exists for the typewriter or adding machine layouts, although some research is currently being undertaken.

Other devices used for numeric data entry are:

(i) levers for each digit
(ii) rotary knobs for each digit
(iii) thumb wheel switch
(iv) rotary telephone dial

Deininger (1967) compared the rotary dial and thumb wheel devices and found that, although error rates were similar ($\sim$ 2%) there was a difference in entry rates such that the thumb wheel took 20-60% longer than the rotary dial for entering successively different 10 digit numbers. (3.3 secs/digit vs 1.8 secs/digit.) Plath and Kolesnick (1967) found that the average time per digit with a thumb wheel device was about 2.74s (error ~ 2%); the number of digits was 8. Hence, it may be that the devices are comparable in rates of digit input and errors for numbers of 5 or less digits. Conrad (1958) found the rotary telephone dial less accurate than a ten-key pushbutton set when the number had to be held in memory during the operation.

Minor and Revesman (1962) compared a L10 key keyboard, a lever device, a matrix keyboard (10 x 10 digits) and a rotary knob device. They found the ten by ten keyboard was best in terms of accuracy and preference. It was faster to enter a 10 digit number with a keyboard and the matrix device (~ 1.3 seconds/digit) than with the lever or rotary knob device (~1.8 seconds/digit). The median error rate for the 10 key device was only 0.6% compared with 1-2% for the other devices.

2.5 Graphic Data and Symbol Input Tasks

Graphic data may consist of the co-ordinates of geometric information or geometric symbols. Systems have been developed which allow direct input of symbols traced on a pad or other visual display. For example, symbols may be flow-chart boxes, Roman capital letters etc., (Ellis and Sibley (1969)). Commonly used devices are those described in the "Marker Task" section of the chapter. No comparative studies have been found which allow the devices to be judged at this level.

2.6 Voice Input Tasks

General purpose computer systems controlled by human voice have recently become widely available (e.g. Cox and Martin (1975)). In these systems, the operator normally uses a previously established set of about 30-50 multi-syllable words but this can be easily extended. The operator trains the system by repeating each word a number of times and error rates of < 2% have been reported by Martin and Cox for such vocabularies over a range of applications. The advantages of voice data entry are that operators can be using their hands and/or eyes for manipulating/monitoring other aspects of their environment.

A comparison of voice recognition with keyboards for inputting digital data was carried out by Braunstein and Anderson (1959). The voice recognition system used was that of the experimenter who measured the speed and accuracy of naive subjects reading digits aloud. Subjects spoke digits at about twice the speed at which they could type them but the typing task was preferred and judged to be easier than speaking the digits. The inference of this is that preferences are not simply associated with the performance aspects of input methods, but with ease of use. Since voice input was not preferred, it may be assumed that the underlying processes of voice communication are, in the situation being discussed, significantly different from those of keyboard input. The talking rate was about 2.5-3 digits per second and the typing rate about 1.5 digits/second. For experienced typists the typing rate reached 2.8 digits per sec. Since error rates were comparable, the authors concluded that voice input offers advantage over typing in this task. However, if alphabetic words rather than digits had been used, the conclusion may be that voice data entry also has speed and accuracy advantages over keyboard. But no comparative data has been published.

3. Conclusion

The attention given to the study of the effective use of input devices is negligible compared with the variety between and within devices, tasks, users and working environments. Most attention has been given to the use of keyboards by typists and not by 'occasional' users solving problems. Furthermore, there is no clear framework within which to carry on investigations of the effects of using different input devices in man-computer problem solving.

The implicit assumption behind general recommendations based on this experimental work is that the performance characteristics of the input sub-system do not depend on what it is used for. This is an important assumption since there are many different systems in which the same input sub-system can be used. For example, Martin (1973) describes twenty-three techniques of using alpha-numeric keyboards combined with a visual display; Newman and Sproull (1973) describe ten techniques for interactive graphics. There are no data presently available to test the generality of results comparing input devices using simple laboratory tasks. Some indication of lack of generality is provided by statistics on time-sharing computer use such as those discussed by, for example, Boies (1972) and Yule (1972). Boies points out that the time taken to input a command to a computer is a function of the command complexity and that this time is also related to the system response time (SRT). As SRT increases from 1 to 10 seconds, the user's time to input a message using a teletype keyboard increases from 15 to 24 seconds. Yule (1972) collected statistics on the use of a mixed input device system and there is a wide range of the user's time to input a command. Other studies have been concerned with aspects of interactive problem-solving, particularly response time (Grossberg et al. (1976)) and keyboard lock-out effects (Boehm et al (1971)). The role of different input devices in these tasks was not a major aspect of investigation, but it was assumed that the input sub-system had a constant effect.

In order to proceed with the fruitful investigation of the area, a systematic approach is required which establishes a suitable framework for the development of an experimental strategy.

CHAPTER 2

SCOPE AND PHILOSOPHY

CONTENTS

1. Introduction - The General Aim of Research

The previous chapter has stated the scope of the thesis to be about the performance and acceptability of alternative input devices used for man/computer interaction. The literature review, which was based on a simple task taxonomy, showed that there were very few reported studies in relation to the variety of input devices and how they are used. The purpose of this thesis is to extend the range of knowledge within this scope by considering the use of input devices for problem solving. This objective may be reached in a number of ways. This chapter describes the strategy of the thesis and lists the specific objectives of the research in the form of testable experimental hypotheses.

2. The Research Strategy

The need for a research strategy arises because of the large number of different input devices, that can exist in a variety of man-computer systems, and how they are used. For example, measurements may be made of the use of a light pen by a practising designer with a sophisticated multi-input computer-aided design system. Without a suitable framework derived from a research strategy, such measurements may not be usefully comparable with measuresments about, for example, the use of a graphic tablet in different circumstances. A first step in the development of a strategy is the identification and naming of the critical parameters within a framework. This requires two conditions: the existence of a suitable language for description, and a descriptive model of the processes involved which allows parameters of importance and their relationships to be identified.

Once such a language and a model have been derived, this chapter continues by developing the tactics of the thesis. That is, how the specific research described in this thesis is related to the general strategy.

3. Language for Research

The literature review was organised around a taxonomy of input tasks. A main basis of the taxonomy was the different types of information being input to the computer, i.e. graphic symbols, numeric data, etc. The information being referred to is the source information for input to the computer; this will be referred to, therefore, as source information to distinguish it from information which is operated on by the computer. The latter is referred to as receiver information and may be different from source information depending on the input device and its coding. Both source and receiver information are at different levels and there are corresponding differences in the processes which may operate on them.

Many authors (e.g. Cherry (1957)), have examined the concept of different levels of receiver information and processing. Table 1 shows the names being given to these levels which derive from semiotics (the study of signs and systems). The processes are briefly described as follows:

(i) Syntactic processing acts on primitives which may be marks, alphabetic, numeric or other forms of information. They are decoded according to some specified rules but no further interpretation of the message (string of primitives) is allowed.

(ii) Semantic processing acts on the total message checking it for consistency with respect to some pre-defined grammar. This process may act on individual primitives and/or groups of primitives (words).

(iii) Pragmatic processing is where the message is interpreted in terms of the internal state and goals of the receiver.

Level of Processing	Information Type	Examples
1 Syntactic	Primitives	Marks, characters
2 Semantic	Groups of primitives (words)	"calculate"
3 Pragmatic	Strings of words and primitives (messages)	"calculate phi"

Table 1: Levels and Type of Receiver Information

The information type described refers to that which is operated on during the appropriate processing. This is referred to as descriptive information. Another class of information types may be called "prescriptive" and refers to that information which causes the different levels of processing to be carried out. This idea derives from the work of Stamper (1973). Prescriptive information for the receiver may also be of the three different types shown in Table 1.

The use of the above definition is not restricted to the man being the source information and computers being receivers. Corley and Allan (1976) have used the language to describe an approach which treats the computer as the information source and the man as the receiver. In interactive man-computer dialogue, both man and computer are alternately sources and receivers of information from each other. Hence the language should be appropriate for describing a model of man-computer interaction.

4. The Development of a Model of Man-Computer Interaction

The situation to be described in a model is that of a man interacting with a computer through a terminal of some kind in order to solve a problem. Such a situation has been modelled by J.R. Carbonell (1967) with the same intent as that here, namely for the purpose of deriving a framework for talking, thinking and carrying on investigations. However, Carbonell's model was constructed by considering the man as a decision maker and taking an information theory point of view which did not emphasise the role of input devices. The development of a model of the man as an interactive problem solver, emphasising the role of input devices in man-computer interaction, is the goal of this section.

The assumptions are:-

(i) The model is descriptive:
The model is not intended to be a normative or rigorous formulation of man-computer interaction. It is developed as far as necessary for the purpose of the thesis (Chapanis (1961)).

(ii) Man-computer interaction consists of goal-directed transactions:
This is based on Miller's (1969) observation of archetypal tasks in man-computer interaction. A transaction consists of exchanges of information between man and computer until a particular state is reached.

(iii) There are different levels of information and processing:
as implied by the semiotic approach.

Figure 1 shows a simple descriptive model of man-computer interaction; it is not original but was implied in discussion by Corley and Allen (1976). It is the starting point for the development of a more detailed model, and uses the language of semiotics described in the previous section. A brief description of the operation of the model is as follows:-

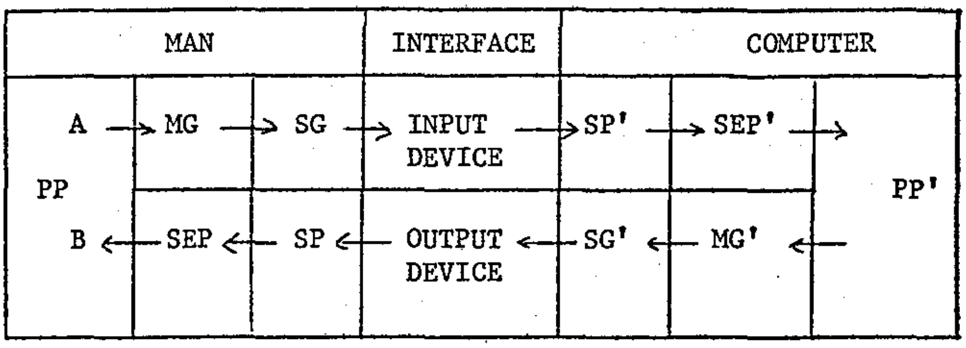
(i) The human problem solver develops a solution involving man-computer transactions.

(ii) Each description of a transaction consists of the specification of operation(s) to be performed by the computer and expected result(s). The operations must be transformed into a suitable language for the computer, (message generation), and then input to the computer as a string of primitives (string generation).

(iii) The computer carries out appropriate syntactic and semantic processing on received primitives which results in a verified message to be acted upon. This action is taken (computer pragmatic processing) and the results presented to the man after suitable message and string generation by the computer.

(iv) The human problem solver carries out syntactic and semantic processing on computer displayed information, resulting in a verified message which may correspond to an expected result of a transaction.

(v) The cycle (ii)-(iv) repeats until all transactions are accomplished. If the solution is acceptable, the process is complete. Otherwise, the complete process is re-entered.



KEY:	
——→	signifies normal information flow
A→B	signifies a transaction
MG	signifies message generation
SG	signifies string generation
SP	signifies syntactic information processing on symbols
SEP	signifies semantic information processing on messages
PP	signifies pragmatic inforation processing on verified messages
'	signifies that process and primitives etc. are different from those of the man

Figure 1: A Simplified Model of Man-Computer Interaction Showing Basic Processes

It is apparent, even in this simple description that human interactive problem solving cannot easily be described as sequential information processes. The total process consists of a hierarchy of sub-processes whose structure is complex. Miller, Gallanter and Pribram (1960) in their book "Plans and the Structure of Behaviour" provide a means for the description of complex processes as units of "Test-Operate-Test-Exit" or TOTE processes. Any given behaviour can be represented by a hierarchy of TOTE processes. It is assumed that such a representation is suitable for man-computer problem solving. The approach taken is to consider TOTE models of component processes and fabricate a complete model from these. The component processes have been chosen so as to emphasise the role of input devices and information flow through them. The critical factors affecting the latter are:-

(i) the rate and quantity at which source primitives can be generated and put into the input device by problem solvers (string generation and coding).

(ii) the rate and quantity of messages generated by human problem solvers (pragmatic processing).

(iii) the rate and type of information provided by the computer in transactions about input of information by problem solvers (all computer processes).

Therefore, the component processes to be modelled in greater detail are:- the human output process corresponding to (i) above; human problem solving for (ii) above; all computer processes for (iii) above .

4.1 A Model of Human Input to a Computer

Hillix and Coburn (1961) derived the model shown in Figure 2 of information flow for people operating a keyboard inputting simple messages to a computer. It is assumed that it is valid, in the context of this chapter, for all input devices and message sources. The operation is briefly described as follows:-

(i) The receptor enables the person to perceive what message is to be input and this is held in the input store.

(ii) The connector associates the stored information with the responses to be made by the effector. These associations are held in a connection store.

(iii) The connection store is interrogated and updated by the connector according to a learning process.

(iv) The output store keeps orders ready for the effectors which control movement of the computer input devices.

(v) The checker and re-arranger check that the information has been effectively input in the right order. If not, the output store is appropriately modified and information is put in again.

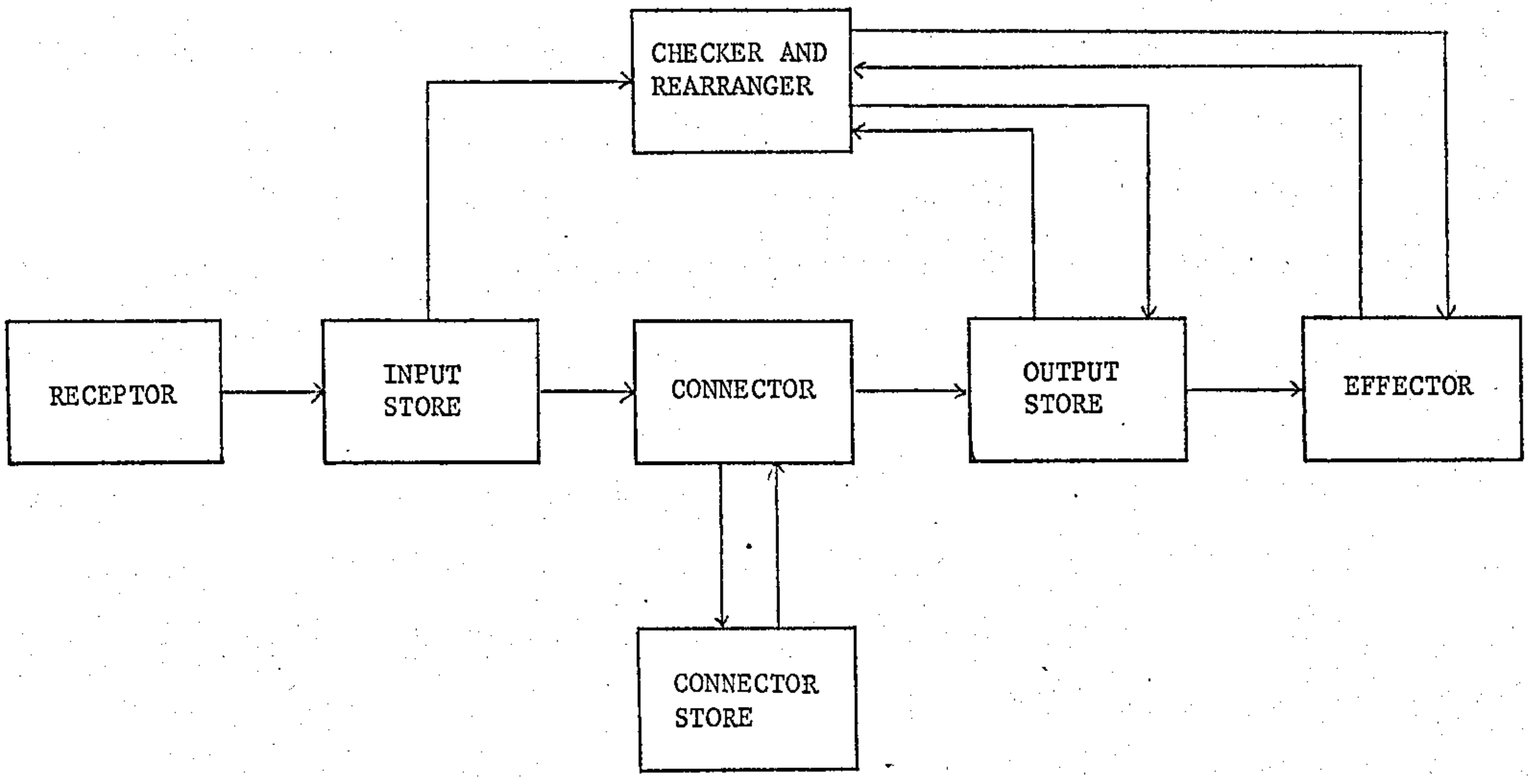


Figure 2: A Model of Human Information Processes in Simple Input
(From Hillix and Coburn (1961))

As in similar models (e.g. Van Geffen (1966)), the exploration of this level of detail has exposed some of the main factors affecting the rate and quantity of information that is transmitted by the input device. Seibel (1972) and Sperandio and Bisseret (1968) point out that a most important factor is the parallel processing which is afforded by effector and central processing. Thus, short term memory may act as a 'buffer' store between the central and effector processes. This store may be under- or over-loaded depending on the rates of input to and output from it. The consequences of over-loading may be loss of information at most, and increased error rates at least.

The quantity of and rate at which information is put into the output store depends on the compatability between the form of the message in the input store and the required form at the output store. If the syntactic and semantic rules for the message in the output store (defined by the computer language) are incompatible with those of the input store, then the connection and checking processes may be slow and complex (Carlisle (1974)).

The quantity and rate at which information may be emptied from the output store depends, amongst other things, on the input device and how it is used.

It is apparent from this model that, ideally, language compatability and the computer input characteristics (input device and how it is used) should be such that there is optimal use of the human input and output stores for a given level of knowledge (connector store content). This implies a balanced view of man-computer dialogue design which takes into account the context of man-computer interaction, and the level and type of information transferred.

4.2 A Model of Human Problem Solving

The particular view of problem solving being taken is summarised by the following conditions:

(i) A well defined set of initial conditions exists and may be described.

(ii) A well defined goal may be specified.

(iii) A set of rules which must be followed in reaching the goal may be specified.

(iv) For the problem solver, there exists some means of manipulating or expressing himself in an environment.

Figure 3 shows a model of human problem solving based on the work of Newell and Simon (1972), Maier (1931) and Guetzkow (1951). At the level of description used here, there is no disagreement between these workers on the problem solving process.

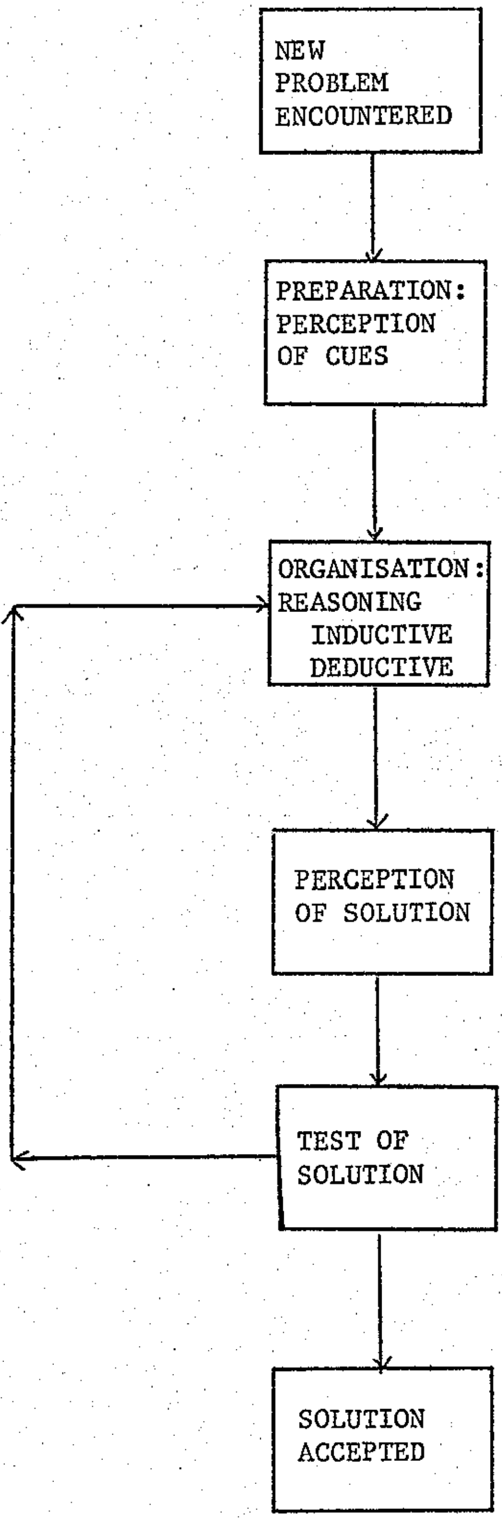


Figure 3: A Simplified Model of Human Problem Solving
(Based on Newell and Simon (1972)

The operation of the model is that:-

(i) the man receives instructions and directions and perceives cues to help in the organisation of a solution. A possible solution is arrived at after reasoning.

(ii) The possible solution is then tested either using an "internal" model (i.e. one held in the mind of the problem solver) or in the real "external" world.

(iii) If the tried solution fails the test, another possible solution is arrived at by going back to the organisation stage (stage (i)).

(iv) Solutions are tested until either an acceptable one is found or the problem solver gives up trying.

If a problem solver has a computer, then this may aid him in testing solutions. The problem in the real world may be modelled in the computer and it may be such that manipulating this is more convenient than manipulating the real world. However, the price to be paid is that the mechanism for manipulating this model has to be learned. Assuming that this price (which depends on many factors) is worthwhile, consider the human problem solving process with computer aid.

The process starts with the problem solver generating a set of solution steps. This solution is tested by means of inputting a string of messages in sequence to the computer. If the computer feedback is not pragmatic, then the input may be quick. If it is pragmatic, then the input is delayed until the result of the previous solution step has been processed.

The degree of delay will depend on, for example, whether there has been an unexpected result so that a minor adjustment to the solution plan is necessary, or whether the solution plan is totally inappropriate and must be completely re-created.

In some classes of problems, such as those requiring hierarchical forms of solution, Hayes (1966) has found that the fast rate of input increased as the solution goal was being attained. In these cases we would expect a positive skew on the distribution of measured rates of flow of information.

As in the previous model of human processes inputting to the computer, the model of problem solving with computer aid allows parallel processing; this time between man and computer. The degree to which it is possible for the man to, for example, review his strategy while waiting for a computer response is difficult to assess because this partially depends on the loading of his short-term memory. This, in turn, depends on the processes described in the model of human input to a computer as well as the need to remember the current state of his solution. Other factors are the expected computer response time for the current transaction, the extent to which a solution is "remembered" by other means (e.g. computer), and the motivation for solving the problem. None of these have been explored in the context of this thesis, although some have been discussed in relation to other areas, e.g. response time, Miller (1968).

4.3 A Model of Computer Processes

It is suggested that, where people interact with computers, short-term memory is used to store information ready to be input to the computer for processing and that the rate at which the store is filled varies according to the rates of processing at different information levels. Some of the factors which affect the rate at which short-term memory may be emptied are determined by computer processes. A particular factor is the information 'feedback' by the computer to the man during transactions.

Two types of feedback are identified: proximal (i.e. fast and direct) and distal (slow and indirect). Proximal feedback is generally non-specific whereas distal feedback provides information specific to the transactions being processed. This information may be at a syntactic, semantic or pragmatic level and may have particular dynamic characteristics. For example, syntactic input (e.g. letters) may be typed in, but only when the message is complete will the received input be displayed to the operator. In other systems, priority is given to providing syntactic feedback of every input and the user is then 'locked out' by the system until it can respond again at a syntactic level. The lock out time or system response times at the syntactic, seamantic and pragmatic levels are the dynamic characteristics of distal feedback, which also affects short term memory.

Figure 4 shows a general model of the computer processes involved in providing proximal and distal feedback at different levels. While proximal feedback is valuable it is of limited use in determining the rate at which the short-term memory is emptied, since it is non-specific. The main factors determining this rate are the distal feedback and the error proneness of the input device in its conditions of use.

In this model, the information processes have been emphasised in a general rather than a detailed way which attempts to, for example, define what is stored at what point by the computer. The simple model emphasises interface characteristics of a system, (i.e. hardware interface, software processing and peripherals). A hypothetical system has been examined in the following description of how the model operates. The system has particular characteristics related to when and how information may be received, processed and fed back to the human problem solver.

Human activity (usually speech or motor) causes the input device to operate and pass primitive information to the computer for syntactic processing. The input device may provide proximal feedback information about the fact that it has been activated by visual (e.g. lit-up keys), auditory (e.g. mechanical noise) or proprioceptive (e.g. alteration of impedance to the activator) means. The capacity of the input device for carrying information depends on the number of discriminable states it has (e.g. keys). Each state (e.g. key) may be allocated to a primitive permanently or dynamically by a computer process.

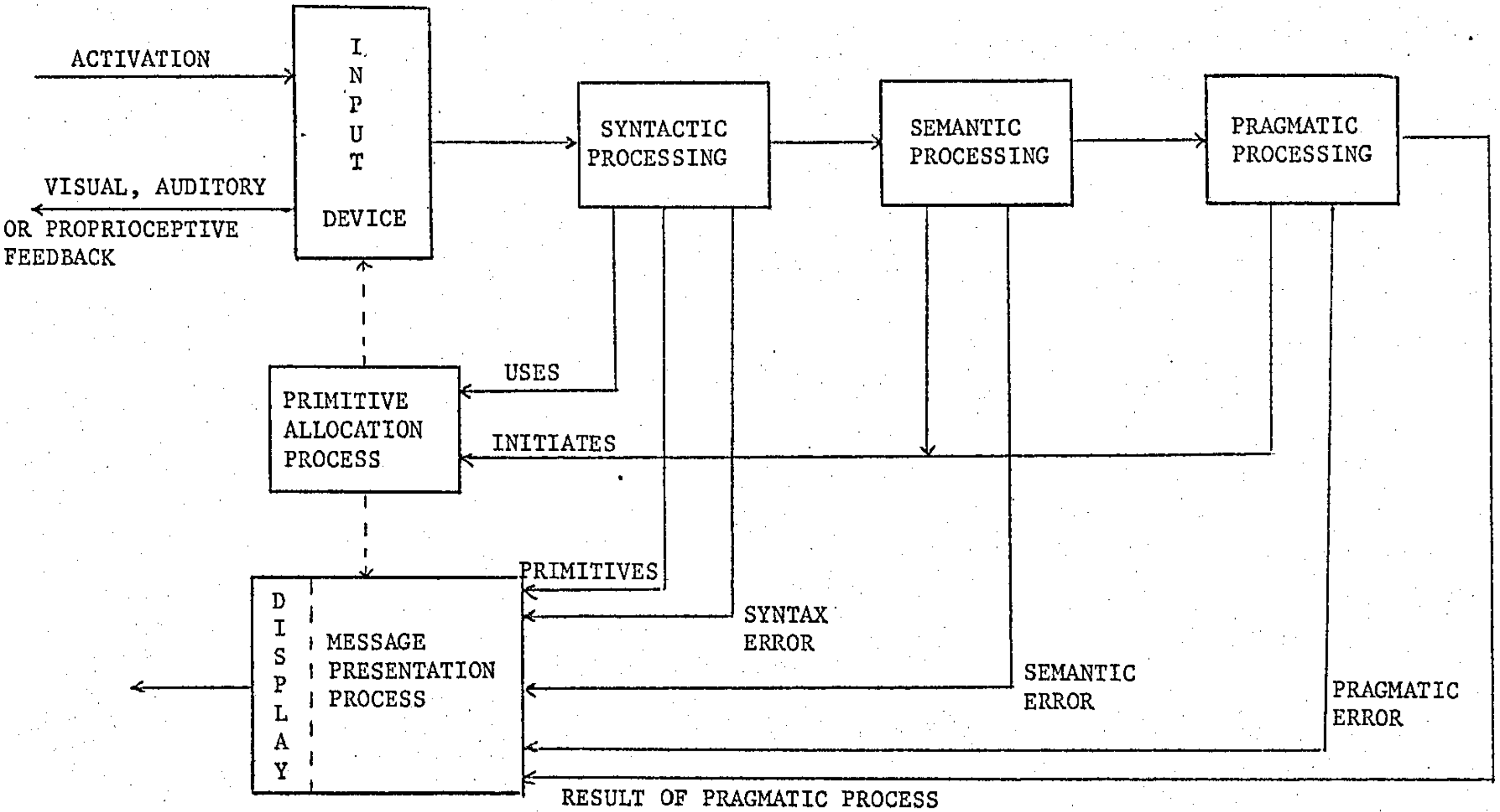


Figure 4: A Simplified Model of Computer Processes

On receipt of the primitives, the syntactic processor passes information to the message presentation process for display to the man in visual, auditory or other modes. The primitives are stored in a list corresponding to the order in which they are received. Particular primitives may have syntactic significance, e.g. a primitive may signify (i) delete last primitive in the list, (ii) delete total input received, or (iii) this is the end of the string, etc. On receipt of these, appropriate action is taken and messages displayed.

On receipt of particular primitives signifying the end of a message, the semantic process checks the legality of the combinations of primitives and words. If errors are found, these are displayed to the user. If not, then the verified input message is sent to the pragmatic processor where it is decoded and translated into computer instructions and operated on. If errors are found in this process, appropriate messages are displayed; otherwise normal actions are resumed until the initiative is again given to the man via a special message.

As an example of how this model works, consider the case where a man wishes the computer to set two numbers, A and B, equal to each other. The computer language determines the form of the message to be input is 'set A=B'. Assume that there are two alternative systems; one where a standard keyboard is used for input and the other where a light pen is used with a set of displayed options. Table 2 shows an analysis of the information received in the two hypothetical systems.

	Message Input in Human Primitives is 'SET∇A=B'												Information Type being Processed
	Letters Typed								Light Pen Picks				
Information Process	S	E	T	∇	A	=	B	↲	1 SET	2 A	3 B	4 ↲	Computer Primitives
Syntactic	d	d	d	d	d	d	d	d	d	d	d	d	Words
Semantic		d		p	d	d	d	p	d	d	d	p	Messages
Pragmatic		p			d		d		p	d	d		Verified Messages

KEY: p = prescriptive information for the computer

d = descriptive information for the computer

↲ = signifies end of input to the computer

∇ = signifies space typed

Table 2: Analysis of Hypothetical Input Characteristics

The analysis is explained as follows:-

(a) Typewriter Input

On typing each key the computer automatically checks that the received primitives (S, E, T etc.) are legal. This operation of the computer is equivalent to implicit prescriptive syntactic information with each primitive. The space symbol signifies the end of the word 'SET' and semantic processing is initiated (i.e. '∇' is prescriptive at the semantic level). The combination of letters 'SET' is checked as a legal word in the language. The rest of the characters are received until the symbol for end-of-the input '↲' is encountered. This prescribes more semantic processing which checks whether A and B exist, if they are suitable types for equating and whether the prescribed operation 'SET' is legal in terms of the equating operation. If all conditions are met, then pragmatic processing is carried out using the legal construct, A=B and the value of B is given to A inside the computer.

(b) Light Pen Input

In this case, all the input is made by pointing at one item of a 'menu' or list of items on a display. Each list consists of primitives which are, by definition, legal. That is, the syntax processing is implicit and does not need prescribing at that level. On picking the word 'SET', the next menu is displayed which contains a list of declared variables and an alphabet (for new declarations). On pointing at A, a new menu appears with the same information as before plus a numeric pad and the previously declared variable A. Pick 3 is of B from the new menu. Finally, pick 4 causes the three previous items to be checked at the semantic level ('=' is implicit) and, if acceptable, the legal construct A=B is processed at the pragmatic level by the computer as in (a).

These two examples are fictitious and have been chosen to show how the computer processes may affect particular characteristics of man-computer interaction.

The inferences of the operation of this model are that the rate and quantity of information emptied from short-term memory depends on:-

(i) the human mechanical process of operating the input device
(ii) the relationship between primitives held in memory and the primitives carried by the input device (coding)
(iii) the characteristics of proximal feedback for (i)
(iv) the characteristics of the distal feedback at the syntactic, semantic and pragmatic information levels.

The characteristics of the input device are particularly important in (i), (ii) and (iii). The degree of importance is discussed in the following section which describes the composite model of man-computer problem solving.

5. A Model of Man-Computer Problem Solving Emphasising Input Devices

Three models have been described; human input to a computer, human problem solving and computer processes. The common link between these processes is the transfer of information from the man to the computer. In forming a composite model of man-computer problem solving, however, it is also necessary to consider the characteristics of the information transfer between computer and man. This thesis emphasises the role of input devices to the computer. Thus a simple view of the processes involved in computer-man information transfer is being adopted. The view being taken is that information from the computer (other than direct proximal feedback) is sensed and perceived, then checked at the syntactic and semantic level and finally used at a pragmatic level.

Of the five levels of distal feedback considered by reference to the model of computer processes, three are concerned with the correction of errors reported by the computer, one with the solution of the problem, and one with the transfer of primitives. The correction of errors and the transfer of primitives has been accommodated within the structure of the model of human input to a computer by introducing simple error correction processes, e.g. 'modify connector store and input store.' The processing of distal pragmatic feedback is more complex. In order to develop the model to accommodate this aspect, reference was made to Carbonell et al. (1968) who provided the concept of utility or cost functions. These provide a means by which the interactive problem solver decides during the solution process whether or not to change different processes. Carbonell's (op cit) work shows that such functions are probably complex in that there are many variables involved and many rules of combination in describing these functions. In this thesis and the model no attempt is made to elaborate such a function except in relation to the use of input devices.

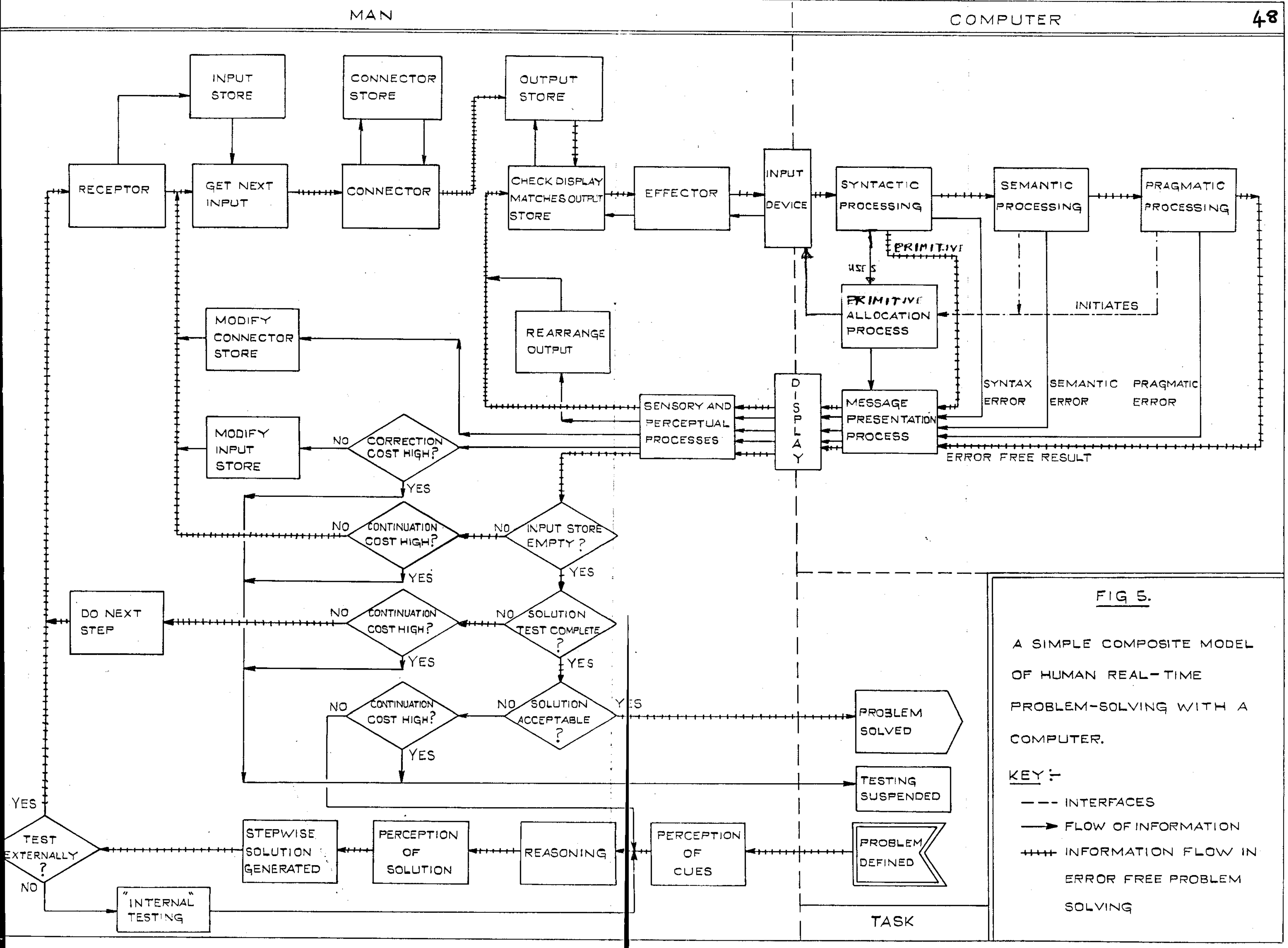


FIG 5.

A SIMPLE COMPOSITE MODEL OF HUMAN REAL-TIME PROBLEM-SOLVING WITH A COMPUTER.

KEY:-

--- INTERFACES

→ FLOW OF INFORMATION

++++ INFORMATION FLOW IN ERROR FREE PROBLEM SOLVING

Four evaluation decisions are proposed; one concerned with a global criterion of continuation cost and the others with local progress towards a solution. The continuation cost includes such factors as the problem solvers need to do other work, eat, sleep, and so on, in relation to his estimate of time and effort required to solve the problem at that time. It is the continuation cost function which is likely to be complex. The other evaluation decisions have been separated on the basis of a much simpler criterion: whether they are directly related to the current problem being solved with the current input device. The composite model may now be fully built.

The structure of the model in Figure 5 represents a collection of hypotheses about the effect of input devices on man-computer problem solving. In essence, the framework for thinking, talking and carrying on investigations is embodied in that model. It is therefore necessary that the operation of the model is reasonably well understood. The following section describes an example of the operation of the model and discusses the role of input devices within it.

5.1 An Example of On-Line Problem Solving in the Model

In this example it is assumed that the cost of solution testing using the computer is sufficiently competitive with alternatives for the computer to be used. Further, it is an example of error-free problem solving using a particular language.

Stage 1. New Problem Encountered:

The example problem is that a discrepancy has arisen between an expected experimental result (from theory) and an obtained experimental result (from observation). The problem goal is to explain the difference.

Stage 2. Perception of Cues:

The observed and expected results are compared to see where the difference arises in relation to how the results were obtained. Assume the difference arises when factors A and B are high.

Stage 3. Organisation:

Different hypotheses are generated to account for the differences which may involve strategic breakdown of the problem into parts, e.g. look for correlations, then check theory.

Stage 4. Perception of Solution:

The generated hypotheses are possible solutions and collectively may be perceived as 'a solution' when the hypotheses are interdependent. Hypotheses may be expressed in a problem meta-language; e.g. factors A and B are correlated.

Stage 5. Test of Solution:

(a) Some of the hypotheses may be tested without recourse to any computer aid; this example deals with a computer-aided test. Algorithms are available (e.g. Siegel (1970)) for the statistical testing of hypotheses. Some of the processes of these algorithms involve laborious and error-prone calculation when done manually. Hence, a freely available computer is a low cost (money, time and error judgement) alternative to manual processes, provided the problem solver knows how to use it. The test algorithm may be expressed in a solution meta-language. It is this which is encoded by the problem solver into the computer language, e.g. the correlation algorithm parts amenable to computer aid - (i) calculate the sums, sums of squares and cross products of the data, (ii) calculate the correlation coefficient given the formula.

(b) The distinct logical steps in the solution meta-language are encoded step-by-step by the receptor into an appropriate form for encoding into the computer language, e.g. 'calculate coefficient' is encoded into (i) put data into computer
(ii) calculate sums, etc.
(iii) print coefficient

The connector store then may encode (i) into a message to be input in terms of human primitives.

' FOR I = 1 to N; GET DATA(I) '

This is stored in short-term memory and input by the effector to the computer. Depending on the input device and its way of use, the command may be input a letter at a time, or more than one letter at a time (as in Table 2 example). For each input, direct feedback to the checker is received from the effector.

(c) On receipt of the whole message, the computer checks it to see if the syntax rules have been violated. For example, if a colon had been used instead of a semi-colon, an error would be reported and the whole message would have to be repeated. If no syntax errors are found, the message is checked against semantic rules. For example, if 'WHILE' replaced 'FOR' in the message, the message becomes meaningless.
Thus, for input of words (WHILE, FOR, etc.) a syntax error (mis-input in this case) can lead to a semantic error. Depending on the computer system programs, other semantic checks may be made such as whether N has been set to a value or DATA( ) has been declared. As with syntax errors, semantic errors cause an error message and the need for the user to re-input the message.

(d) If there are no syntax or semantic errors, the message is converted (compiled) into a suitable form for computer operations and the computer starts to carry out the command. In the course of the 'pragmatic' processing (where the goal of the computer is to store data for the user), various errors may occur. For example, N may be too large for the computer. However, if there are no pragmatic errors, then the computer requests each data point from the user.

(e) If the user is typing each data point in turn, then it is possible that more syntax errors (e.g. 12.A instead of 12.1) and semantic errors (121 instead of a number less than 100) may occur. If so, then the consequence of these errors may be such that the whole command may have to be re-typed.

(f) If no errors occur, then all the data are input and the next step in the solution ('calculate sums', etc.) is encoded into the computer language. The testing cycle repeats until the coefficient is printed. Then the hypothesis testing algorithm is re-entered and the significance of the calculated correlation coefficient obtained.

(g) Solution testing continues in this way; i.e. in a hierarchy of processes, until an acceptable solution is obtained.

In the foregoing example, the role of input devices and their way of use is clear; it is also clear that there are a number of human processes involved in on-line problem solving which may be in parallel and which rely on short-term memory. Thus, the consequences of an error-prone or slow input device may be as little as a change in way of use of the device (adaptation) or as large as a change in problem solving strategy so as to minimise the use of the computer. Nothing is known about the extent of these effects or how, for example, they relate to effects of long system response times. A difficulty in prescribing such effects is that, unlike the foregoing example, on-line problem solving performance is not deterministic. Thus, any quantitative description of the effects of different input devices must take into account differences between people.

Similarly, differences may arise because the problems collectively making up a task may vary. The following section describes the strategy designed to cope with some individual and task differences.

6. Experimental Strategy

Having derived the language and model of man-computer problem solving which emphasises the role of input devices, the development of a strategy continues by classifying the factors of interest and stating general forms of hypotheses. This leads onto the specific hypotheses tested in this thesis and the approach taken in testing them.

6.1 Classification of Parameters of Interest in this Thesis

6.1.1 Methods of Inputting Information to the Computer

There are two main components of interest in this thesis. These are the input device and its way of use (coding). The combination of these components are referred to in the thesis as the input method for any particular system.

(a) Type of Input Devices

A device may or may not provide proximal feedback (other than proprioceptive) on its operation. Devices have a number of separate 'states' which are available to be used for conveying primitive information. For example, input devices may produce a binary code on operation (e.g. keyboard). In this case, the maximum number of states corresponds to the maximum number of different codes produced by the device. Alternatively, input devices may produce voltages which are proportional to the position of the activator (e.g. joystick) and this is converted into a compatible form for the digital computer. The number of possible positions for the type of device is generally larger by orders of magnitude than for the discrete type of input device.

All the devices in the literature review may be encompassed by this simple point of view. But this is of little value unless the way in which the primitives are related to the states of an input device are taken into account. This is determined by coding.

(b) Classification of Ways of Use

The way in which an input device is used depends on how and when primitives are allocated to the states of the input device. Primitives may be allocated to states of the device either dynamically during interaction (e.g. changeable menus) or statically before interaction (e.g. QWERTY keyboard). In either case, primitives may be allocated by the computer (e.g. words in menu lists on a CRT) or by other means (e.g. an overlay of a keyboard put on by the user). Finally, the coding of computer primitives (as described in Chapter 1, section 2.3) may be at different levels (e.g. Table 2).

6.1.2 Characteristics of the Computer

Although there are many of these, the most pertinent to the study of the role of input devices are, (i) the provision of adequate distal feedback about information input to the computer, and (ii) the provision for input error correction and recovery, and (iii) computer language compatability with solution.

6.1.3 Characteristics of People

The model implies that the characteristics of people which would most affect how input devices are used are short-term memory, motor skill, language knowledge, problem solving ability and the type of 'cost' functions used to evaluate performance.

6.1.4 Characteristics of Tasks (Problems)

Task difficulty would be expected to affect the rate of information transfer across the man-computer interface and hence must be considered in relation to the role of input devices. The model infers that solution structure, problem representation and information content are relevant in determining task difficulty.

6.2 Hypotheses

The model of human interactive problem solving represents a collection of general hypotheses.

6.2.1 General Hypotheses

The research hypotheses which are concentrated on in this thesis are:

1. H1 : Problem solving processes are affected by the input method required by the computer.
2. H2 : The input method determines the transfer of information between man and computer depending on the interactions between the characteristics of computers, people and problems.
3. H3 : The acceptability of different input methods is based on an individual's judgement of a combination of factors affecting information transfer. These factors are the characteristics of input devices, computers and problems.

6.2.2 Experimental Hypotheses

The derivation of more specific experimentally testable research hypotheses was based on both the need to be selective and practicalities such as availability of research facilities. Within these constraints, laboratory rather than field investigations were carried out, using particular facilities. These are described in the appropriate chapters.

In planning the experiments, a degree of freedom existed in the choice of experimental hypotheses. In the work of this thesis, the input method and the type of task used were the main variables. The choice of input method was restricted by the availability of particular input devices and associated software. The type of task was chosen to be either problem solving or non-problem solving. Because of these practical limitations and the chronology of development, the experimental hypotheses were not systematically related to the general hypotheses by using the framework of the descriptive model.

Table 3 summarises the experimental hypotheses which have been tested in this thesis and provides subsidiary information on the structure of the thesis.

6.3 Hypothesis Testing

The approach taken in this thesis is to develop each experimental hypothesis in the chapter which describes how it was tested. The chapters are self contained with a summary at the beginning of each which allows the reader to glean the relevant information. Experimental details and data are contained within each chapter

Relates to General Hypotheses No.	Experimental Hypotheses (EH)	Experiment No.	Independent Variables	Dependent Variables	Chapter
2, 3	EH1 The rate at which information is transferred by an input device does not depend on the input device	1	Input device	Information flow rates	3
2, 3	EH2 The rate at which information is transferred by an input device does not depend on its way of use	2	Way of Use	Information flow rates Acceptability	4
1, 2, 3	EH3 (i) Problem solving does not depend on the way of use of an input device. (ii) Problem solving does not depend on the input method.	3	Input Device Way of Use	Information flow rates Acceptability Problem solving performance	5
1, 2, 3	EH4 Problem solving does not depend on the combinations of input method, and the characteristics of people and systems.	4	Input Method Distal feedback User Personality	Information flow rates Acceptability Problem solving performance	6
1, 2, 3	EH5 Problem solving does not depend on the combinations of input method and system characteristics	5	Input Method Task Difficulty Distal feedback	Information flow rates Acceptability Problem solving performance	7

TABLE 3: Experimental Hypotheses under Test in this Thesis

except where it is appropriate to put data into Appendices. In general, the approach in problem solving experiments is to keep pragmatic and semantic information similar and examine alternative methods for inputting syntactic information.

7. Summary

This chapter has described the experimental strategy which is proposed for investigating the role of input devices in man-computer problem solving in real time systems.

The approach taken was to use the language of semiotics in the derivation of a descriptive model of man-computer interaction. The model emphasised the role of input devices and used models derived by other workers.

The model was used as a basis for classifying factors of interest which were in turn used for deriving experimental hypotheses from general hypotheses. The particular hypotheses under test in this thesis are a consequence of circumstances which are described rather than being a systematic balanced approach to testing the general hypothesis. The thesis continues by detailed descriptions of the testing of the hypotheses and finishes by summarising them and recommending future work.

CHAPTER 3

EXPERIMENT 1

SUMMARY

The investigation described in this Chapter concerns one main variable - the input device used to select words from a displayed list, (Experimental Hypothesis 1). Two devices were used; a light pen and a keyboard (QWERTY). The experimental design included four variables which were the number, brightness and separation of words in the list and the method of using the devices. The number of words was arbitrarily fixed at nine while the other screen variables were controlled within two levels of brightness and three levels of word separation. The method of choice of a word for the keyboard was to type the character (A-H) which identified the required word in the list. The method of using the light pen was merely to point at that word. The light pen had no switch attached.

Two measures were taken. These were the mean time taken for the subject to select a given word and the number of selection errors made for a number of selections.

The experiment was run using ten subjects each being required to select ten words in succession for each of the combinations of the variables. Practice with each device was allowed, and all conditions were presented in a balanced way. Emphasis was placed on speed rather than accuracy.

The results showed that the light pen was faster in use but more error-prone than the keyboard. However, with the light pen, subjects traded off speed for accuracy in accordance with the menu characteristics of item separation and brightness. With higher brightness levels, the selection time decreased

as separation increased. With lower brightness levels, the opposite was true. Increases in brightness and decreases in separation both increased the light pen selection error rate. The effects of individual differences were important in these selection time results.

The conclusion is that the rate at which syntactic descriptive information is transferred by an input device depends on the input method, i.e. Experimental Hypothesis 1 is rejected. The relevance of this is not as great as the fact that in comparing input subsystems even in simple tasks, some input methods may be optimised independently. Thus, the approach whereby two alternatives are compared must be taken with caution.

CONTENTS

TABLES

GRAPHS

PLATES

1. Introduction

This Chapter deals with testing the experimental hypothesis that the rate at which information is transferred by an input device does not depend on the input device. The level of information chosen for the test was descriptive syntactic. The hypothesis was tested at the non-problem solving level using two available alternative input devices; light pen and keyboard. The use of the keyboard was such that 3 letter words (source primitives) were coded into single alphabetic characters (receiver primitives). The use of the light pen was to point at the word to be transferred.

2. Objectives

The objective was to compare the use of a light pen to a keyboard when used for selection of a word from a displayed menu list. A secondary objective was to provide information on the factors affecting this comparison.

3. Constraints

The hardware and software constraints are described in detail in this experiment, since the constraints of experiments 2, 4 and 5 are similar.

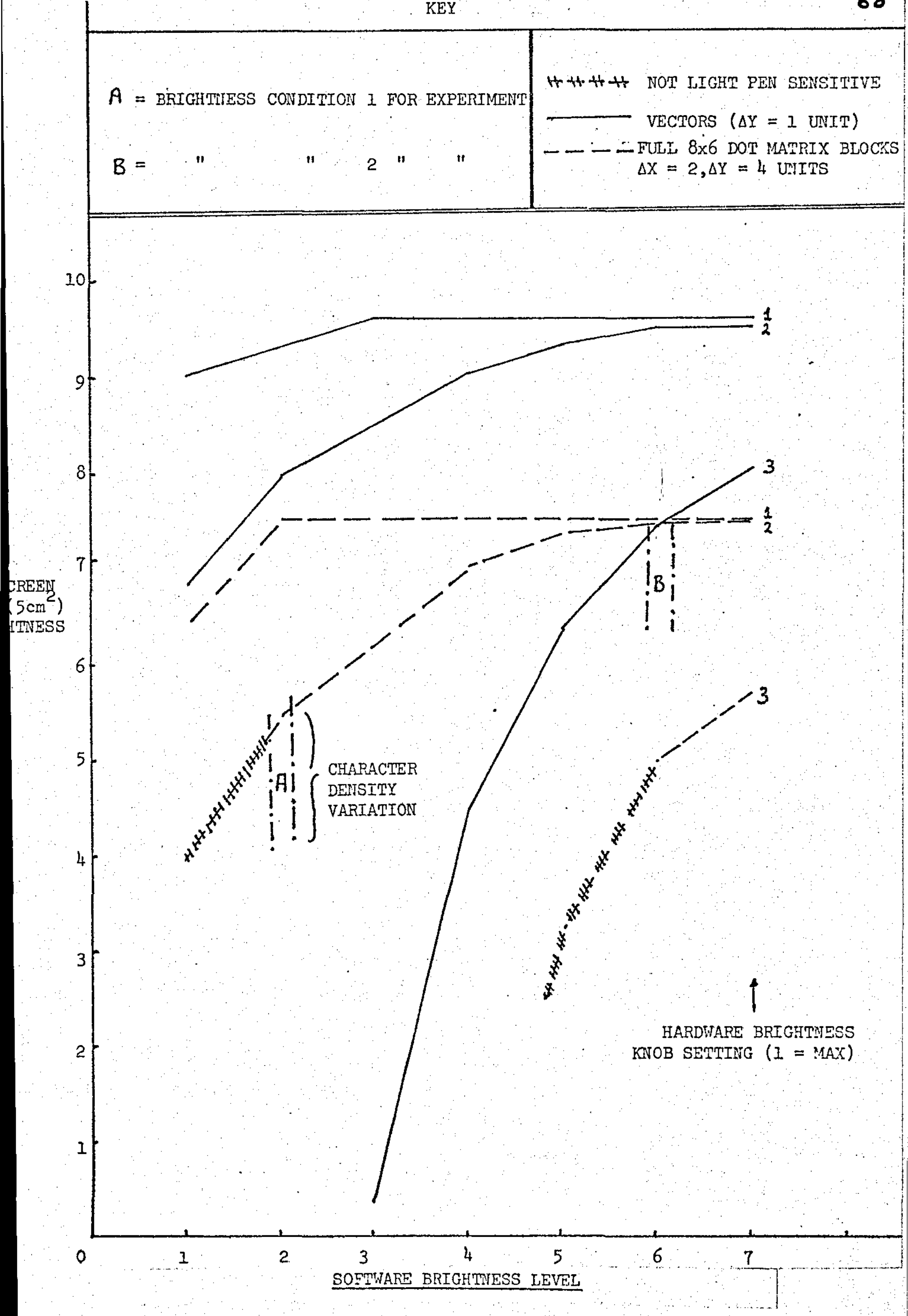
3.1 Hardware

The hardware used was a DEC PDP 11 GT42 intelligent graphics terminal which had 16K of 16 bit words, paper tape input and output facilities, but no other storage medium. There were three input devices; a 100 character/sec. paper tape reader;

an ASR 33 teletype and an unswitched light pen (i.e. always sensitive to screen items). The output devices were an ASR 33 teletype and the standard VR17 display of the GT42. The VR17 display had its own processor and character generation hardware. Data and vectors could be drawn in the addressable area but there was no hardware circle drawing facilities. Text could be either graphic (i.e. based on a graphic dot matrix) or non-graphic (restricted to fixed areas of the screen).

Table 1 gives some brief details about the display; in particular that there are eight software controlled brightness levels, and that the characters are 6 x 8 dot matrices. There is also a hardware control on the VR17 display itself which is analogue adjustable through all 8 levels. The problem of light pen sensitivity was that the combination of hardware setting and software level affected the response of the light pen.

Graph 1 was drawn as an attempt to calibrate the brightness levels. The light units were read off a standard Weston light meter held against the screen which contained a centralised 5 cm^2 of either vectors or full matrix characters. Three arbitrary marked hardware settings were used (1 (full on), 2 (2/3 on), 3 (1/3 on)). At position 1, only three of the 7 non-zero software brightness levels were effective with the vectors and only 2 for characters. At the other extreme, (position 3), only 5 levels were effective for vectors and 3 for characters. Position 2 represents the hardware setting


KEY
A = BRIGHTNESS CONDITION 1 FOR EXPERIMENT
B = " " 2 " "
NOT LIGHT PEN SENSITIVE
VECTORS (ΔY = 1 UNIT)
FULL 8x6 DOT MATRIX BLOCKS ΔX = 2, ΔY = 4 UNITS
CREEN
(5cm^2)
ITNESS
10
9
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
1
2
3
1
2
3
A
B
CHARACTER DENSITY VARIATION
HARDWARE BRIGHTNESS KNOB SETTING (1 = MAX)
SOFTWARE BRIGHTNESS LEVEL

GRAPH 1

where 6 levels are effective, and at this setting, software level 1 is not light pen sensitive for all the character set. At a slightly lower knob setting all levels would be effective but only levels above 4 will be light pen sensitive.

3.2 Software

The GT42 was used by programming in FOCAL-GT which is a real time command language developed by DEC. FOCAL-GT normally operates in 8K and the GT42 had 16K, therefore some minor changes were patched into core which allowed FOCAL to use the top 8K for display files. This allowed a larger program to be developed for the purpose of this investigation.

All terminal hardware characteristics were controllable through FOCAL although a constraint was that the input buffer for the teletype was limited to two characters. If this was exceeded, FOCAL reported an error and interrupted the main program.

This also occurred for such events as arithmetic overflow, stack overflow, etc. Since FOCAL had no facility to resume at break points, care had to be taken to minimise the chance of stack overflow by appropriate experimental design.

A further constraint was that the text displayed through FOCAL as graphic text (i.e. in any screen position) was italic and upper case only. Non-graphic text was upper case non-italic and subject to the constraints of Table 1.

DISPLAY PROCESSOR:

Drawing Times:

Character	∿26μs
Vectors	∿18μs for 0.5"
Character font	6 x 8 matrix
Intensity level	8 (including 0)

C.R.T.

Light pen	Non-switched solid state
Tube size	17 in. diagonal
Viewing area	8.25 in. x 11 in.
Phosphor	P39 doped with IR
Characters/Line	72
Lines/Frame	31

Table 1

Some details of the Computer Display

3.3 Subjects

At the time of the experiment, the number of available subjects was limited to ten.

4. Experimental Details

4.1 Design

The objective concerns the comparison of two input devices (light pen or keyboard). The constraints show that this must be examined in relation to the brightness, separation and number of the words in the list on the display. It was decided to use two brightness levels; the lowest level representing the position where small changes in software level was not significant in determining the sensitivity of the light pen; the highest level where small changes were significant to the sensitivity of the pen.

Three word separation distances were chosen; the smallest allowed each word to just be perceived separately at normal viewing distance; the middle value was the 'normal' FOCAL separation and the largest value was about twice the 'normal'. The variable of the number of words in the list was not included in the design, but was fixed at nine.

The design chosen was to present each subject with all conditions of brightness and separation conditions for the light pen but with only a particular subset of conditions for the keyboard. This was arrange because the keyboard did not have the same sensitivity to the brightness and separation as the light pen and the overall time of the experiment could be reduced for each subject by eliminating redundant conditions. This was of importance because of the limits on the available time of subjects.

Condition	Meaning	
S1	Menu Item Separation level 1	
S2	Menu Item Separation level 2	
S3	Menu Item Separation level 3	
B1	Item brightness level 1	
B2	Item brightness level 2	
PL	Practice condition with light pen	S2,(B1+B2)/2
PK	Practice condition with KEYBOARD	S2,(B/+B2)/2
A	Experimental condition S1, B1 with light pen	
B	Experimental condition S1, B2 with light pen	
C	Experimental condition S2, B1 with light pen	
D	Experimental condition S2, B2 with light pen	
E	Experimental condition S3, B1 with light pen	
F	Experimental condition S3, B2 with light pen	
G	Experimental condition S2,(B1+B2)/2 with keyboard	
H	Experimental condition S3,(B1+B2)/2 with keyboard	

Table 2

Experimental Conditions

Table 2 shows the experimental conditions and Table 3 shows the order of presentation of these conditions for ten subjects. Table 3 was constructed by reversing the order of presentation of each condition within each input device. For example, subject 1 carried out the keyboard conditions before the light pen conditions in order GH. But subject 2 had the light pen conditions before the keyboard conditions and the keyboard conditions are reversed to H,G. This procedure was carried out in an effort to reduce any effects of the order of presentation of conditions.

4.2 Sessions

Each subject was required to carry out all the conditions in one session. Table 4 shows the procedure for each subject. The experimenter stayed in the same room all through the session but subjects were not allowed to talk while interacting with the computer.

Each condition consisted of ten trials with the appropriate variables set. One practice trial was used for training the subject. A trial was the selection of a list item after the target had been presented.

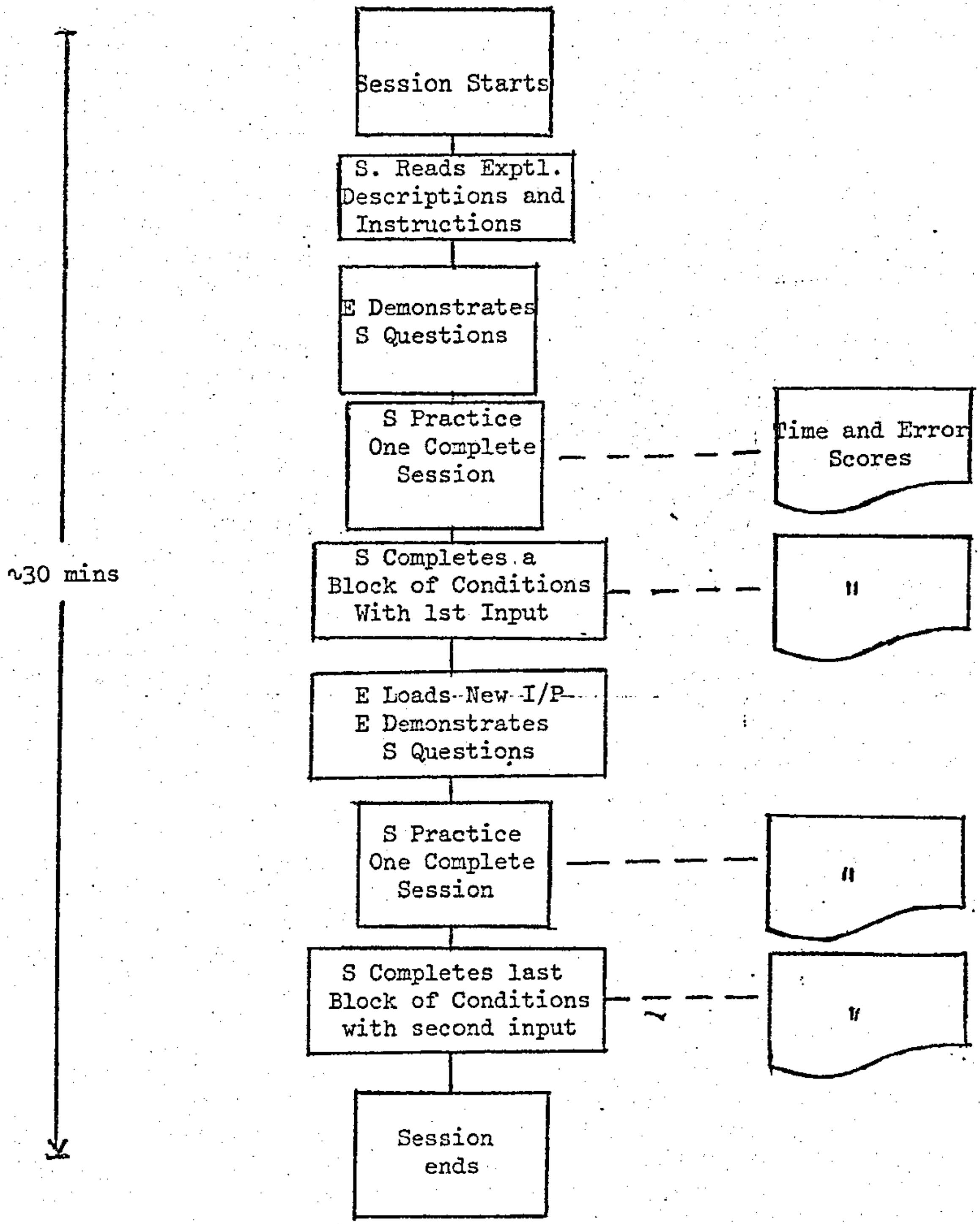
All subjects were allowed to see their trial and condition results as a way of improving motivation during a session. No-one was allowed to see other subjects' results which were confidential. Where appropriate, subjects' comments and experimenters' observations were noted during a session.

Subject number	Order									
	1	2	3	4	5	6	7	8	9	10
1	PK	G	H	PL	A	B	C	D	E	F
2	PL	F	E	D	C	B	A	PK	H	G
3	PK	H	G	PL	A	B	C	F	E	D
4	PL	C	B	A	D	E	F	PK	G	H
5	PK	G	H	PL	F	E	D	A	B	C
6	PL	D	E	F	C	B	A	PK	H	G
7	PK	H	G	PL	D	E	F	A	B	C
8	PL	C	B	A	F	E	D	PK	G	H
9	PK	H	G	PL	A	C	E	B	D	F
10	PL	F	D	B	E	C	A	PK	G	H

Table 3

Order of Presentation of Experimental Conditions

TABLE 4 FLOWCHART OF PROCEDURE FOR EACH SUBJECT



Key: S means Subject
E " Experimenter
I/P " Program input to computer

4.3 Task

The basic task was concerned with simple selection of an item from a list of graphically displayed items (menu) as quickly as possible.

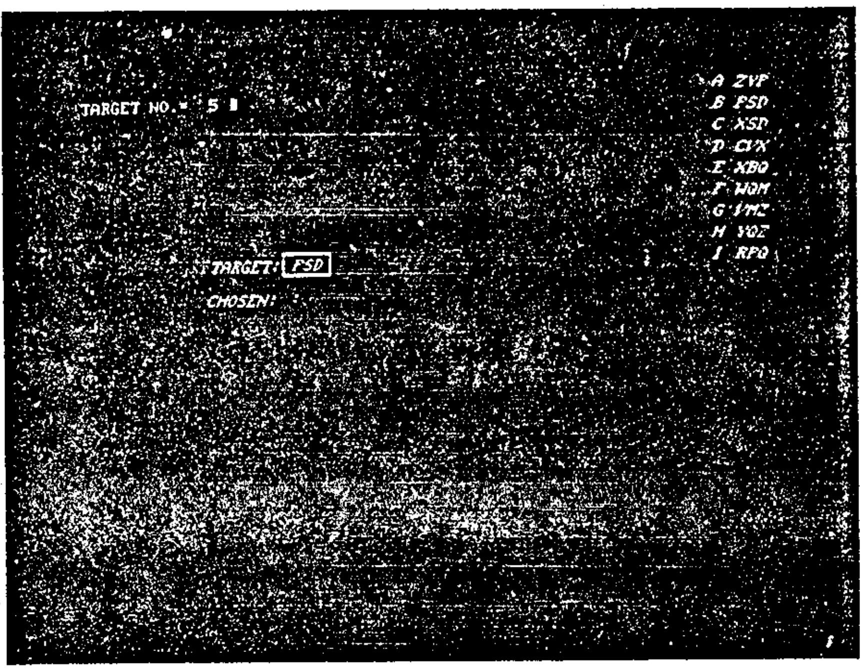
The menu list shown in Table 5 was deliberately made up of nonsense letter groups to discourage memorisation by phonetic rehearsal. A consequence of this was that the groups were not perceived as words and had to be read letter by letter. This may be equivalent to reading three full words. Each menu item was three letters; again, this was deliberate, since length information is helpful to fast location of an item. The list consisted of items positioned down the right-hand side of the screen (see Plate 1).

Two further task variables were considered; the presentation of the target and knowledge of results. In such a simple task as this the latter was necessary to increase motivation. It was also necessary that the subject could anticipate the appearance of a target. Two methods were used to enhance this anticipation; first when a new target was being prepared by the computer, a 'clock' would show the stage reached (see Plate 2). Second, when the new target was displayed, the teletype would print a space, thus providing an auditory cue.

LIST OF MENU ITEMS
CVX
XBQ
YQZ
ZYP
RPQ
WQM
VMZ
XSD

Table 5

Menu Item List



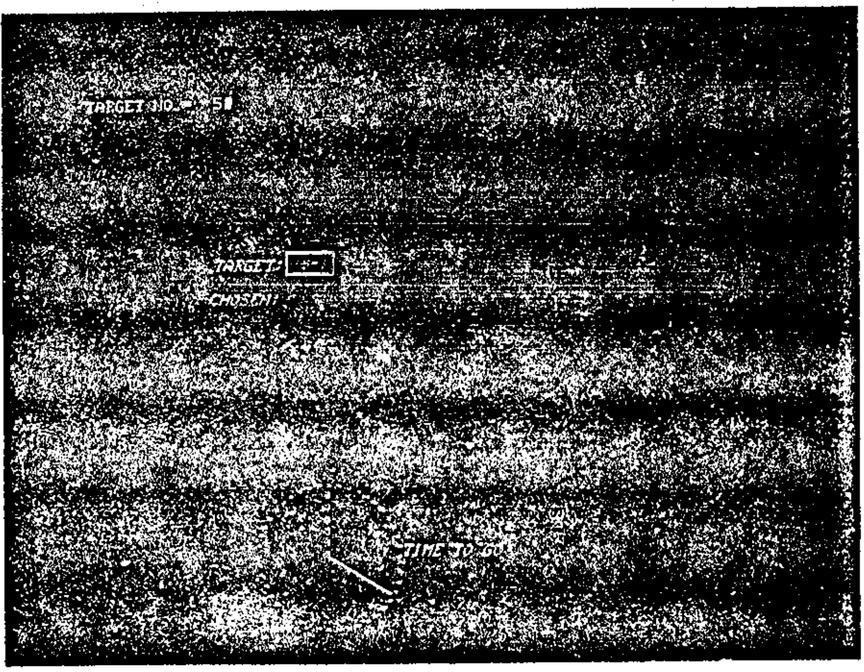
TARGET NO.= 5

A ZVF
B FSD
C XSD
D CVX
E XBQ
F WQM
G JMZ
H YQZ
I RPQ

TARGET: FSD

CHOSEN:

PLATE 1



TARGET NO.= 5

TARGET: [illegible]

CHOSEN:

TIME TO GO

PLATE 2

Knowledge of results (distal feedback) was given to the subject in audio-visual modes. On selection of a menu item, first the menu disappeared, the clock started, a new target number appeared, and the chosen item appeared under the target; secondly, if the selection was correct, the teletype bell rang, otherwise a space was printed. Thus, the inherent characteristics of the task reinforced accuracy while the experimenter emphasised speed of response.

Finally, for every new target of the ten trials, the menu item order was randomised, so that the subject could not benefit by learning the list order. Targets were also randomised within the constraint that it may appear only once unless it was target 1, which could appear twice.

4.4 Subject Information

Apart from simple training instructions given verbally during demonstrations, Tables 6 and 7 show the instructions for the light pen trials. Similar instructions were given for the keyboard trials with appropriate wording altered.

4.5 Subject Population and Context

In the simple task used in the experiment, few specific details were taken. Each subject was asked whether he/she had used a similar light pen or keyboard before for a similar task. If so, they were excluded. Left handed subjects were also excluded.

SUBJECT INFORMATION

The experiment you are about to take part in is concerned with evaluating the use of the computer graphics light pen. You are asked to read the attached sheet of instructions carefully before proceeding. Any queries you have should be raised at this time.

Please note that your identity and the information collected are confidential to the experimenter.

Thank You.

Table 6

SUBJECT INSTRUCTIONS

1. You will be asked to place the light pen in your right hand and look at the graphics screen.
2. When the "clock" (being displayed) reaches zero, a target word of 3 letters will appear in the rectangular box on the screen.
3. At the same time a list of similar 3 letter words will appear at the top right hand side of the screen, and the printer will operate briefly to remind you of this.
4. Your task is to point at the target word in the list on the screen as quickly as possible with the light pen.
5. If you choose the right word, the teletype bell will ring, otherwise it will not. Once a selection has been made, you cannot change it.
6. The word you choose will be displayed to you on the screen.
7. There are ten targets per session which will only take you about five minutes.
8. The first session is for you to practice and ask questions and start as soon as you indicate that you have read and understand these instructions.

Table 7

5. Independent Variables and Measures

5.1 Input Method

(a) Input Device

Two levels of this variable were used: (i) an unswitched light pen and (ii) the standard QWERTY keyboard of an ASR33 teletype. Because of the nature of the light pen, levels of display variables were considered as part of the independent variables (Section 5.2).

(b) Way of Use

The way of selecting the word was at two levels for each level of input device. These were using the light pen to point at the word directly, and using the teletype QWERTY keyboard to type a word-associated alphabetic (A→H) character. In the latter case, the menu was displayed together with this list as shown in Plate 1. With the light pen, the letters A→H were not displayed. Subjects were not allowed to correct errors. The first letter typed or the first item picked was final.

(c) Formal Description

Receiver Information Level	Source Primitives "Select a Menu Item"	
	Light Pen	Keyboard
Syntactic	Pick 1 (p)d	t (p)d

t means typing any alphabetic character.

d is descriptive information.

(p) means implicit prescriptive information because input of it causes decoding of the input and checking of the result against syntactic rules.

5.2 Display Variables

These were the brightness and separation of menu list items. Table 8 shows the levels of separation and brightness chosen to meet the experimental design conditions given the constraints imposed by the hardware.

6. Dependent Variables and Measures

6.1 Information Rates

The main measure is the speed and accuracy of input of descriptive syntactic information. The speed of selection of an item and the number of items correctly selected were the two main dependent variables.

The speed of selection (selection time) was measured for each selection by the elapsed time between the presentation of a new target and the first light pen pick or keyboard key struck thereafter. An error was tallied whenever the selected word did not correspond to the target. Thus, for each subject and each condition, there were ten pairs of selection time and error counts. These were reduced to mean selection time for correct selections and error frequency for each condition and subjects. The standard deviation or other measure of dispersion was not taken for each condition within subjects.

Label	Value	Software condition
S1	0.125 x character height (0.16 in.)	20 Units *
S2	0.405 x character height (0.16 in.)	25 Units
S3	0.68 x character height (0.16 in.)	30 Units
B1	3.5 - 5.5 arbitrary light units	Level 2 **
B2	5.5 - 7.5 arbitrary light units	Level 6

* Vertical screen size is 1024 programmable points one unit apart

** Software levels from one to seven. One set at just light pen detectable brightness using hardware control.

S means separation
B " brightness

Table 8

Independant Task Variables and Values

7. Analysis and Discussion of Results

Tables 9 and 10 show the data collected on information rates in the experiment.

For the light pen data, a 2 x 3 ANOVA was carried out on the data for correct selections. The model used was a mixed effects related sample model. The results in Table 11 show that there was an effect of brightness level and subject variation on the mean select time for correct responses. Non-parametric tests were used for further within-subjects analysis of effects between conditions.

Table 12 shows the results of paired t-tests on the data which is presented in Graph 3.

7.1 Order Effects

Graph 2 shows the selection time and error rates averaged over subjects for each order of presentation. The design was such that the two keyboard conditions were either after or before the light pen conditions. The lower errors and larger selection times for the first 3 and last 3 conditions presented compared to the overall means are not statistically significant (t-test between means). The statistical fluctuations are large, partly because of the small sample size, but the conclusion is that order effects are not statistically significant.

	Practice	Light Pen						Keyboard		Practice
SUJECT	Mean Brightness Level	Brightness Level 1			Brightness Level 1			Mean-Brightness Level		Mean Brightness Level
	Separation 2	Separation 1	Separation 2	Separation 3	Separation 1	Separation 2	Separation 3	Separation 2	Separation 3	Separation 4
1	1.87	1.94	2.51	2.96	2.46	3.01	1.97	3.89	3.16	3.45
2*	1.71	2.11	2.30	2.51	2.10	2.04	1.93	3.46	3.36	2.89
3	2.89	3.79	4.67	5.16	3.20	2.71	3.16	3.73	3.36	3.53
4	1.95	2.33	2.00	1.56	1.74	2.36	2.03	3.44	2.91	2.49
5*	2.28	2.35	3.30	3.14	2.28	3.58	2.37	5.14	5.08	4.05
6	1.66	1.66	1.69	1.87	1.76	1.41	1.60	2.32	2.33	2.36
7*	3.17	2.29	3.27	3.25	1.92	2.68	2.34	4.20	4.28	4.57
8*	1.79	1.72	1.79	1.70	1.61	1.51	1.51	2.53	2.66	3.08
9*	2.93	2.69	2.72	3.70	2.12	2.75	1.89	3.49	3.71	3.83
10	1.55	2.8	2.05	2.06	1.96	1.62	1.89	3.11	2.94	2.91

* Means used pen stroke technique.

TABLE 9

MEAN SELECTION TIME FOR CORRECT SELECTIONS (SECONDS)

SUBJECT	Practice	Light Pen						Keyboard		Practice
	Mean Brightness Level	Brightness Level 1			Brightness Level 2			Mean-Brightness Level		Mean-Brightness Level
	Separation 2	Separation 1	Separation 2	Separation 3	Separation 1	Separation 2	Separation 3	Separation 2	Separation 3	Separation 4
1	4	7	9	9	10	4	7	9	10	10
2*	7	6	6	6	4	7	9	10	10	10
3	4	8	2	7	7	8	5	10	10	10
4	6	6	8	10	7	7	6	10	10	10
5*	9	9	7	10	4	10	10	10	10	10
6	6	7	8	6	5	5	7	10	10	10
7*	10	6	10	10	7	8	10	10	10	10
8*	8	5	10	8	5	7	5	10	10	10
9*	4	5	6	8	5	7	8	10	10	10
10	6	9	8	9	5	5	4	10	9	10

* Subject 'Stroked' Pen.

TABLE 10

FREQUENCY OF CORRECT RESPONSES OUT OF 10

		CORRECT RESPONSE RATES			MEAN INPUT TIMES		
EFFECT	DF	SSQ	MSS	F	SSQ	MSS	F
Brightness - B	1	2.554	2.554	4.63*	12.15	12.15	3.49
Separation - S	2	0.754	0.377	-	18.3	9.15	2.62
B x S	2	0.676	0.338	-	0.90	0.45	-
Residual	55	30.41	0.552		191.5	3.48	
Total	59	34.39	0.58		222.85	3.77	

ANOVA - Random Effects Model

MEASURE		BRIGHTNESS B(1)			BRIGHTNESS B(2)			KEYBOARD	
		S1	S2	S3	S1	S2	S3	S2	S3
Times (Seconds)	Mean	2.37	2.63	2.79	2.11	2.37	2.07	3.53	3.37
	SD	0.59	0.86	1.03	0.43	0.67	0.44	0.76	0.76
Rates	Mean	6.8	7.4	8.3	5.9	6.8	7.1	9.9	9.9
	SD	1.4	2.24	1.48	1.75	1.66	2.02	0.3	0.3

			CORRECT RESPONSE RATES			MEAN INPUT TIMES		
	EFFECT	DF	SSQ	MSS	F	SSQ	MSS	F
Random	Subjects S'	9	39.68	4.41	1.07	21.94	2.44	13.8*
Fixed	Brightness B	1	12.15	12.15	3.77	2.554	2.554	8.8*
Fixed	Separation S	2	18.2	9.1	3.35	0.754	0.338	2.13
	S' x B	9	29.02	3.22	-	2.65	0.29	
	S' x S	18	48.77	2.71	-	2.64	0.15	
	B x S	2	0.9	0.49	-	0.676	0.338	
	Error	18	74.13	4.12		3.179	0.177	
	Total	59	222.85	3.77		34.39	0.58	

ANOVA - Allowing for Subjects Variation (Mixed Model)

KEY: * means significant at 5% level using conservative F-test.

NOTE: Assumptions :-

(a) Independance - no significant treatment x subject effects.

(b) Normality - skewed distribution of response times. Cut-off of rates at 10

(c) Populations have similar variance (F-test between variances).

(d) All measures in interval scale.

(e) It is assumed that a linear model applies :-

i.e. Y = A(I) + B(J) + C(K) + AB(IJ) + AC(JK) + BC(JK) + ABC(IJK) + E

TABLE 11
Light Pen ANOVA Results

					Light Pen						Keyboard	
					B1			B2			(B1+B2)/2	
					S1	S2	S3	S1	S2	S3	S2	S3
% errors	LP	B1		S1		-0.61	-2.62	1.17	0.00	-0.67	-6.43*	-6.42*
				S2	-		-1.30	1.69	0.57	0.11	-3.21*	-3.2*
				S3	-	-		3.77*	2.36	1.46	-3.07*	-3.06*
		B2		S1	-	-	-		-0.94	-1.50	-6.00*	-6.00*
				S2	-	-	-	-		-0.81	-6.15*	-6.14*
				S3	-	-	-	-	-		-4.33*	-4.21*
	KYB	Mean B		S2	-	-	-	-	-	-		1.00
				S3	-	-	-	-	-	-	-	

TABLE 12

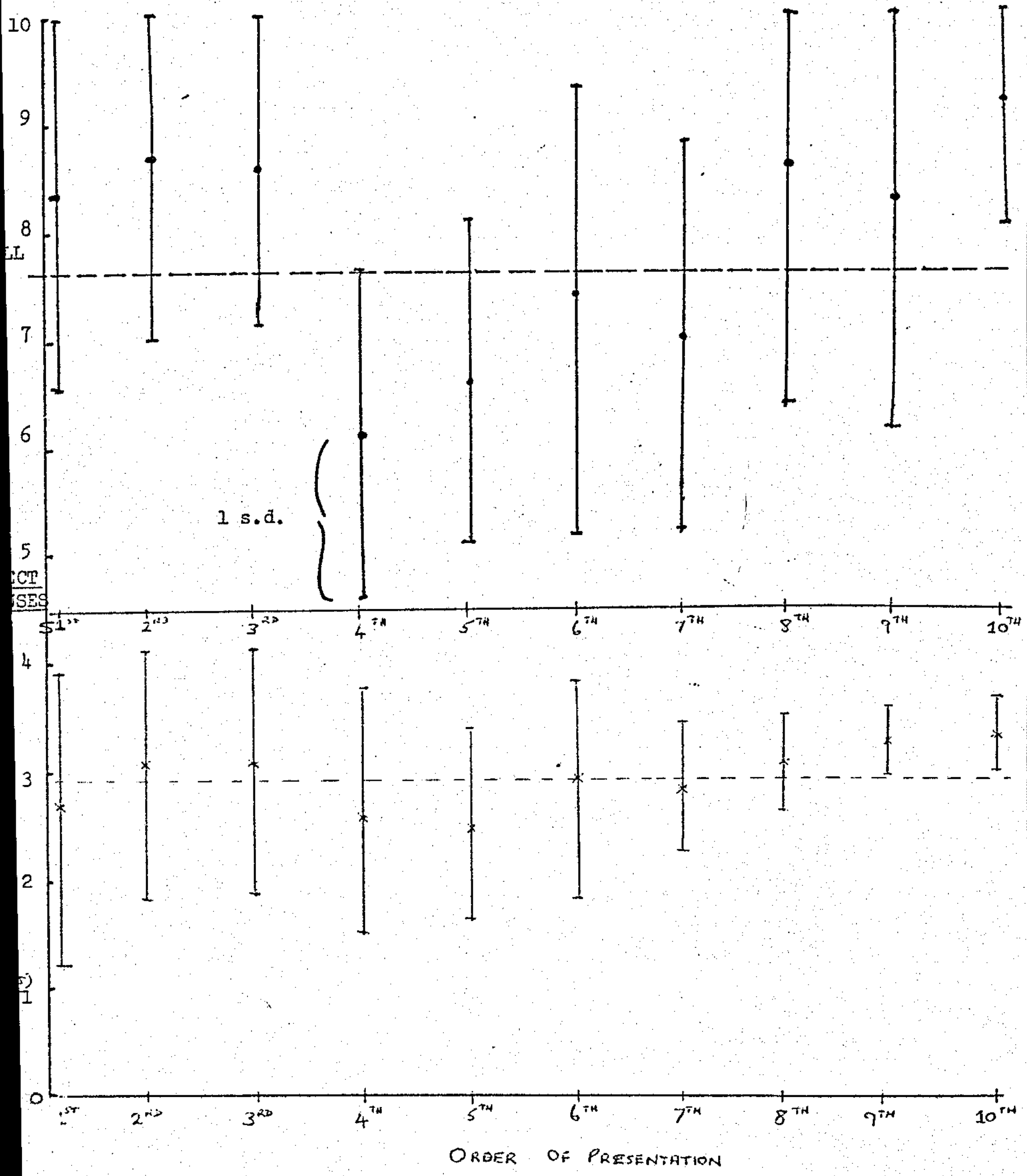
Paired t values

* Significant @ 5% level (2 tailed).

				Light Pen						Keyboard	
				Brightness 1			Brightness 2			$\frac{B1+B2}{2}$	
				S1	S2	S3	S1	S2	S3	S2	S3
CORRECT TIMES	LP	B1	S1	–	-1.61	-1.78	1.95	0.00	2.65	-4.30	-3.59*
			S2	–		-1.27	2.88*	1.19	3.68*	-3.82*	-3.00*
			S3	–	–	–	2.94*	1.53	3.02*	-2.37*	-1.85
		B2	S1	–	–	–	–	0.00	2.65	-1.97	-2.70*
			S2	–	–	–	–	–	1.53	-12.13*	-6.98*
			S3	–	–	–	–	–	–	-7.18*	-5.86*
	KY	$\frac{B1+B2}{2}$	S2	–	–	–	–	–	–	–	1.58
			S3	–	–	–	–	–	–	–	–

Table 12 (continued)

Paired t values

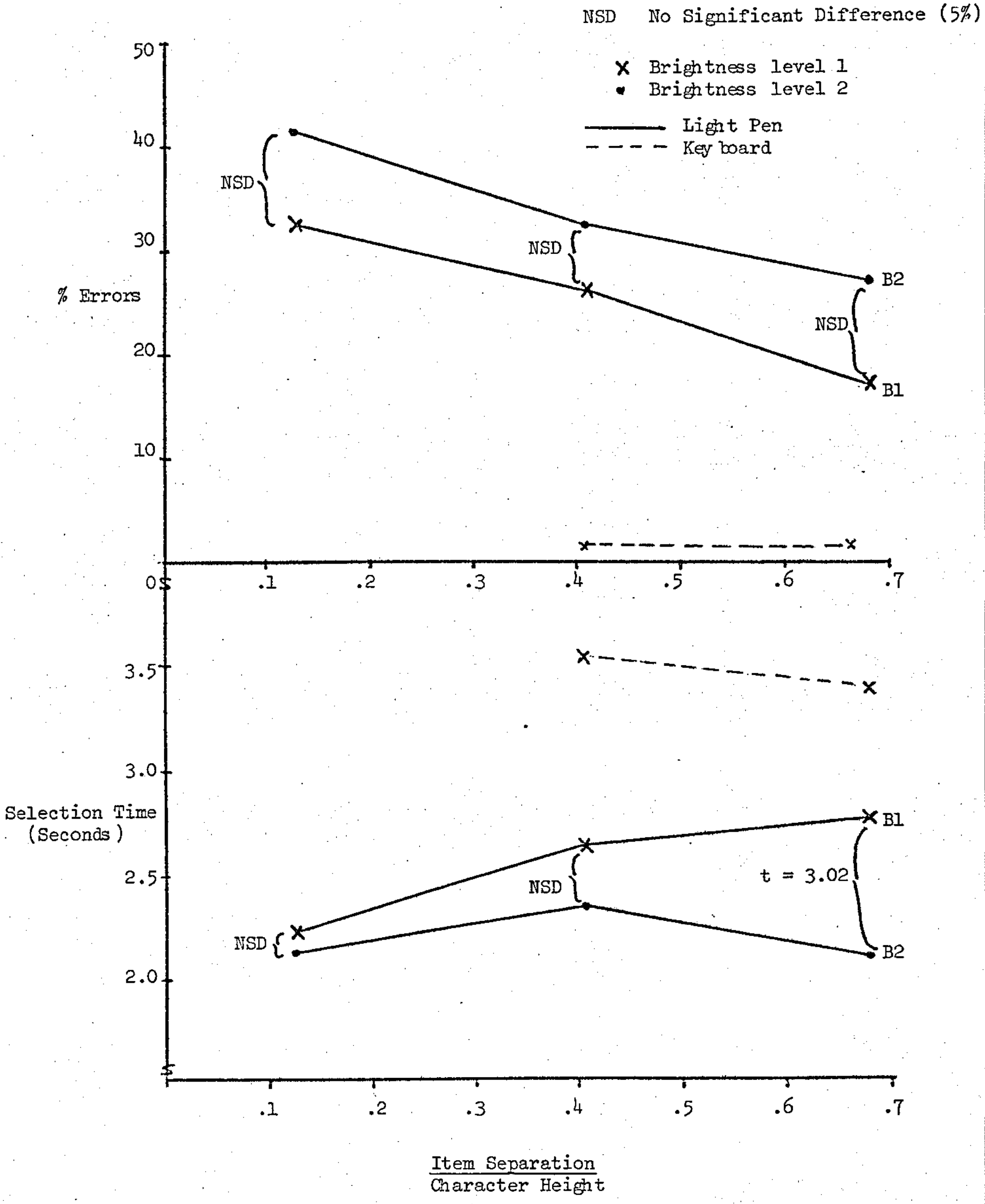


ORDER EFFECTS IN SELECTION TIME AND ERROR RATES

GRAPH 2

7.2 Errors and Input Method

The top most half of Graph 3 indicates that there were significant differences (as indicated by accuracy of item selection) between the use of a light pen and a keyboard. More care was taken to choose the right key than to point the light pen accurately. A possible explanation is that the method of hunting and locating the right key had only time penalty with a very small risk of error, whereas for the light pen, the device had a high risk of error, but a small time penalty. The emphasis was on speed from the experimenter (E) but on accuracy by the task. For the light pen, the conditions were such that the subject could adapt his responses. One such condition was that the light pen did not force a subject to switch his attention from the display so he assigned a different penalty for error than with the key input. With the keyboard such an adaptation was not apparent.



GRAPH 3

Speed vs Accuracy

7.3 Subject Selection Time and Input Methods

Referring to the lower half of Graph 3, keyboard input times are around 3.5 seconds whereas light pen select times are around 2.5 seconds. These times are statistically different using the non-parametric sign test. Thus, the time of using a keyboard was about 1 second more than using a light pen, i.e. an extra 40% time was used for an approximate 25% gain in accuracy. This time was needed for the subject to change attention to the keyboard from the display, hunt for the required key and strike it. With the light pen, no such division of attention was necessary.

Although not statistically significant, the trend of increasing time-to-select with item separation indicates that search time increases with item separation. This was offset for the brightest light pen condition by the easier and quicker light pen hint and by using the light pen as an eye position confirmation device. At lower brightness levels, this was not the case.

7.4 Selection Times and Errors

The selection times for incorrect responses are not shown but in general were smaller than those for correct responses and appeared to be from a different population. Hence, it was inappropriate to use times for incorrect responses in this analysis because of the diverse reasons for their differences.

8. Conclusions

8.1 Light Pen vs. Keyboard for Simple Menu Selection

(a) Speed

The unswitched light pen allowed faster selection than the keyboard methods because no division of attention was necessary between the display and the keyboard.

(b) Accuracy

The unswitched light pen resulted in more errors of selection than the keyboard because of the tendency to use the light pen as a pointer to scan the list.

(c) Trade-off Between Speed and Accuracy

With the light pen, subjects traded off speed against accuracy in such a way as to suit the emphasis given by the context. This was not done with the keyboard.

8.2 Light Pens for Simple Menu Selection

(a) Speed

The selection time depended on the brightness level of the display and usually decreased as brightness increased. Selection time also depended on the item separation. With higher brightness levels, selection times decreased as item separation increased. With lower brightness levels, the converse was true.

(b) Accuracy

Increased display brightness and decreased separation between lines increased the error rate of item selection.

8.3 Keyboards for Simple Menu Selection

(a) Speed

The keyboard select time was independent of the brightness and separation of menu items.

(b) Accuracy

The error rate was also independent of the brightness and separation of the menu items.

8.4 Experimental Hypothesis

The experimental hypothesis that "the rate of transfer of descriptive syntactic information does not depend on the input characteristics" is rejected.

8.5 Implications for the Model

(i) The input characteristics of mixed input/output devices affect man-computer interaction by necessitating switches of attention before effector action. The need for this depends on the user's knowledge of the input device and how it is used.

(ii) Some input devices allow users to adapt to input methods so as to arrive at a suitable balance between speed and accuracy of use.

CHAPTER 4

EXPERIMENT 2

SUMMARY

This Chapter deals with the test of experimental hypothesis (2) that the rate at which information is transferred by an input device does not depend on its way of use. The hypothesis is tested using one input device, a standard keyboard, and two ways of use for selecting an item from a menu list. The information being measured was descriptive and prescriptive at the syntactic level.

The first way of use was to step a cursor next to the required word to provide descriptive syntactic information then confirm entry by providing prescriptive syntactic information; the second way of use was to type in the number of the word in the list, then confirm entry.

The experimental design was a 3-factor 2-level ANOVA. The three factors were way of use of an input device, subject experience and order of presentation. Twelve subjects completed the experiment and the measure of time to input prescriptive syntactic information showed both main and interactive effects due to all three factors.

The conclusion is made that the rate of input of syntactic information depends on the way of use of an input device. The dependency was such that the way of use interacted with task variables (the position of them in the list). The subject's general computer experience and task specific

experience affected the rate of information flow and the results supported the idea that the time needed for error-free input is an important component in the model of man-computer interaction.

CONTENTS

TABLES

PLATES

GRAPHS

1. Introduction

This chapter deals with the experimental test of the hypothesis that the speed and accuracy of information transfer does not depend on the way of use of the input device. The information type used in the test is descriptive syntactic information; namely, items in a list on a display (as in Experiment 1).

The input device used was a standard keyboard. The independent variable is the way of using the device. A subsidiary hypothesis is tested that the speed and accuracy of input of information does not depend on the subject's experience. This hypothesis was suggested by the results of the first experiment.

2. Objectives

These were:-

(a) To investigate the effects of two different methods of using the keyboard to select an item from a displayed menu list on the speed and accuracy of inputting descriptive syntactic information.

(b) To examine the effects of a subject's general and specific ability on information flow in task (a).

3. Constraints

These are the same as for the previous investigation (Chapter 3, Section 2) except that more subjects were available for longer periods.

4. Experimental Details

4.1. Experimental Design

The objectives of the experiment concern two main variables;

first, the way of using the input device and second, the subjects' experience. The first variable was the selection technique and the second was divided into two variables; general experience and specific experience. In the context of the investigation, specific experience means that acquired during the course of the experiment. Such experience would be shown as an order effect if there was a difference between the rates of learning due to the different ways of using the input device.

The design used in this experiment was a complete balanced 2x2 factorial with each subject carrying out all the trials in a balanced order (O) given by Table 1. The factors were selection techniques (S), subject experience (E) and order of presentation (O). The latter was included because of a possible learning effect between and within trials. Position of the target in the list was not included in the design as a main variable.

4.2 Sessions

The purpose of the experiment was explained to each subject as 'investigating techniques of menu selection'. Following a small demonstration of their first experimental condition, the subject was allowed a practice (and questions) before being asked to undergo the first series of 20 trials. A trial consisted of the presentation of a new target, its selection from a list of nine items, followed by an input signifying confirmation of choice. Each target was chosen at random from the list with the proviso that, except for item 1, it could only appear twice. Item 1 could appear 4 times. A similar procedure was used for

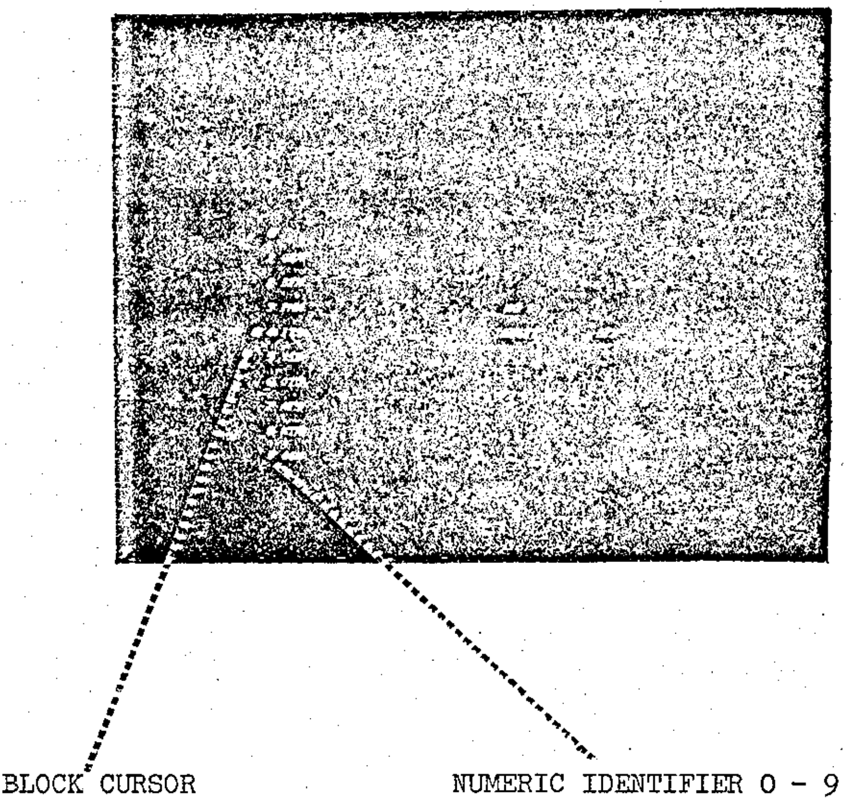


PLATE 1

DISPLAY FORMAT FOR EXPERIMENTAL TASK

the second condition. Before the trials, verbal instructions were given. No incentives or rewards were given other than the investigator urging subjects to perform well in the trials.

The sessions were completed by subjects giving comments and stating preferences about techniques.

4.3 Task

The general form of the task was menu selection as described in the investigation of the previous chapter. However, in this task the subject was able to re-select if a mistake had been made. On deciding that a particular word was the desired one, the subject confirmed the selection.

The subject was presented with a target word which had to be selected from a menu list and confirmed as quickly as possible. The target was one of the menu items. The size of the list was limited to less than 10 items, since selecting the tenth (or more) item required typing two digits and it was desired to make results comparable with those of the previous investigation. Nine menu items were used in the list. Each consisted of a three-letter nonsense syllable and were displayed as in Plate 1. The cursor always started at a position one step above the first item (position 0) with every new target.

After a confirmation, the display was cleared and the message "WAIT FOR BELL" displayed for 5 seconds. After that time, the bell would ring and 0.5 seconds later, a new target appeared.

4.4 Subject Information

Verbal instructions were given emphasising first accuracy and secondly speed of selection and confirmation. The procedure was explained and demonstrated to the subject before each trial and a short practice was allowed (3 targets).

4.5 Subject Population and Context

A sample of 12 people was chosen whose experience with computers and familiarity with keyboards were known. This is summarised in Table 1.

The 12 subjects were divided into two groups of 6 according to two criteria; their typing ability and their experience with computers. Those with more experience of computers were labelled the 'experimental group' (Table 1). It was assumed for the purpose of this Table, that since the difference in techniques is concerned with switching attention from display to keyboard, the confidence of the subject was a more important factor than his familiarity with the keyboard and that this was a function of experience with computers. It may be noted that subjects 5 and 10 differ in their group allocation according to the two criteria.

5. Independent Variable

5.1 Input Method

(a) Input Device (constant)

Teletype (ASR33), keyboard (QWERTY).

Subject Number

	1	2	3	4	5	6	7	8	9	10	11	12
Typing Rate* (SECONDS/KEY)	0.563	0.713	1.051	0.867	0.208	0.937	0.166	0.305	0.304	0.864	0.562	0.296
Typing Group (1 above median) (2 below median)	1	1	1	1	2	1	2	2	2	1	2	2
"Experience" (E) + Category	2	2	0	1	2	0	3	3	4	3	3	4
Experimental Group (Experience E2 ≤ 3)	1	1	1	1	1	1	2	2	2	2	2	2
Order (0)	1	2	1	2	1	2	2	1	2	1	2	1

Key

(* Mode of inter-character time distribution obtained in a simple typing test.
(
(+ 0 = No computer experience; 1 = Data input/output only:
(2 = programs for self; 3 = programs for others;
(4 = hands on and systems programmer.
(
(0 Order of presentation of techniques; 1 = step technique first.

TABLE 1 : Subject Population Details

(b) Way of Use

The alternatives of selecting an item were grouped according to whether, or not, the sequence number of the item was used in the procedure for selection. Two techniques were considered; the first being representative of group one (digit select), the second of group two (step select).

The subject's procedure for the digit select technique was to type in the number of the required item (select) then confirm by typing the carriage return (or CR) key. At the selection stage, a cursor was displayed next to the selected item; if a selection error had been made, then another digit was typed and the cursor moved accordingly, (syntactic feedback).

The subject procedure for the step select technique consisted of typing any key (except the carriage return (CR) 'confirm' key) to move the cursor down the menu list until it was next to the desired item. It was not possible to move the cursor up the list back to the previous item. Confirmation of selection was by typing the CR key. If a selection error was made, the cursor was stepped down the list until the desired item was found. On stepping off the bottom of the list, the cursor returned to its start position at the top of the list.

(c) Formal Description

Information Process	Source Primitive	
	'Choose a menu option'	
	Way of Use 1	Way of Use 2
	Digit Select / Confirm	Step Cursor / Confirm
Syntactic	$\begin{array}{c}D\\ \vert\\ d\end{array}$ $\begin{array}{c}\hookleftarrow\\ \vert\\ p\end{array}$	$\begin{array}{c}\alpha_r\\ \vert\\ d\end{array}$ $\begin{array}{c}\hookleftarrow\\ \vert\\ p\end{array}$

D	means typing any numeric key
α_r	means multiple (r) typing of any alphanumeric key
$\hookleftarrow$	means typing the carriage return (CR) key
d	means descriptive information
p	means prescriptive information

5.2 Subject Variables

Subjects were allocated to the general computing experience group if they were above the group median of subject experience as given in Table 1.

6. Dependent Variables and Measures

6.1 Information Flow

This was measured by the speed and accuracy of confirmation of input and the response time to the first input.

6.2 Acceptability

The subject's preference for technique was used as a simple measure of relative 'acceptability' of an input technique.

7. Analysis of Results

7.1 Raw Data

The data consists of two parts; the mean response times for each subject and preferences for technique. Table 2 shows the mean response times for striking the first key (select time) and for confirming entry (confirmation time) for each subject (1-12), technique (2) and order (2). Table 3 shows the subjects' preferences for technique.

7.2 Treatment of Data

Within each technique, the samples are independent and were subjected to a 2x2 ANOVA using a mixed model (technique, order fixed effects, experience random). Between techniques, selection of the samples carrying out the conditions for the first time allows a 2x2 ANOVA using a mixed model, with independent samples. Table 4a shows the results of the within-technique ANOVAS for select and confirm time. Table 4b shows the results of the between-technique ANOVAS for select and confirm time. Table 5 shows the mean selection times for those effects which were found to be significant.

8. Discussion of Results

The small number of subjects and large variation in results contribute to the lack of effects found in Tables 4a and 4b. There was no significant difference between the confirmation times because of subjects' experience, technique, order of presentation or interactions. This may arise because of the

TECHNIQUE	STEP				DIGIT			
EXPERIENCE	E1 (high)		E2		E1		E2	
ORDER	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Mean Confirmation Time (Seconds)	4.93 5.67 4.89	3.00 6.52 5.33	6.39 7.12 5.30	4.28 6.43 4.05	2.72 6.61 5.98	3.91 4.20 2.71	4.55 4.89 3.79	4.23 5.19 3.23
Mean Select Time (Secs.)	1.80 1.95 1.85	1.20 2.40 2.45	3.93 4.64 2.51	2.40 3.36 1.84	2.32 5.43 4.62	3.17 2.82 2.30	3.81 4.14 3.00	3.34 4.15 2.76
Analysis Groups	A A'	B	C B'	D	B C'	A	D D'	C
Subject Nos. in Group	8,10,12	7,9,11	1,3,5	2,4,6	7,9,11	8,10,12	2,4,6	1,3,5

TABLE 2

Mean Response Times for Correct Responses

	SUBJECT NO.											
	1	2	3	4	5	6	7	8	9	10	11	12
Preference	S	S	D	S	D	D	D	S	D	S	S	D

TABLE 3

Subjects' Preference for Techniques

(D = digit : S = step)

		STEP						KEYBOARD					
		CONFIRM TIME			SELECT TIME			CONFIRM TIME			SELECT TIME		
EFFECT	DF	SSQ	MSS	F	SSQ	MSS	F	SSQ	MSS	F	SSQ	MSS	F
Order O	1	0.87	0.87	-	4.14	4.14	6.54*	0.04	0.04	NS	1.9	1.9	2.34
Experience E	1	1.83	1.83	1.91	0.77	0.77	NS	2.18	2.18	NS	0.02	0.02	NS
Ox E	1	0.96	0.96	NS	1.30	1.30	NS	1.23	1.23	NS	0.06	0.06	NS
Error	8	12.03	1.50		4.40	0.55		12.54	1.57		7.25	0.91	
Total	11	15.69			10.61			15.99			10.13		

TABLE 4a

2x2 ANOVA Results Within Techniques

		SELECT TIME			CONFIRMATION TIME		
EFFECT	DF	SSQ	MSS	F	SSQ	MSS	F
Technique T	1	3.7	3.7	5.96*	2.75	2.75	NS
Experience E	1	1.4	1.4	2.25	0.135	0.135	NS
T x E	1	3.94	3.94	6.35*	2.58	2.58	NS
Error	8	4.97	0.62		16.76	2.09	
Total	11	14.02			22.43	2.04	

TABLE 4b

2x2 ANOVA Results Between Techniques for First Trials Only

(fixed effects for T,O : random for E)

	E1 (high)	E2	Overall Mean
S	1.86	3.69	2.77
K	4.1	3.65	3.88

t = -1·64, DF=10, NS.

TABLE 5

Interactive and Overall Effects in Table 4b - Mean Selection Times variations that occur because of differences in target position discussed in this next section. There were interactions in the selection time results between technique and experience.

8.1 Select Time Data

Table 5 shows that the average select time for the step technique was faster than that for the digit technique, and that the difference was amplified for subjects with general experience. Observation and subjects' comments indicated that the differences between techniques arises because of different needs to change attention from the display to the keyboard. With the digit technique, a particular key had to be located before striking: this required a change of attention from the display to the keyboard. With the step technique, any key could be pressed and it was desirable for the subject to keep attention on the display. Hence, no search and locate time was necessary for the step technique and one hand could be placed permanently over a key on the keyboard. This was done particularly by the experienced subjects: hence the interactive effect between technique and experience.

8.2 Confirmation Time Data

Measurements of the confirmation time included the selection time of each subject and within each technique. However, the effects shown in selection time were not shown in the confirmation time. For the digit select technique, this may be expected if it is assumed that the search and strike time of the confirmation key is constant, within the tolerance of individual differences. For the step technique, the first key strike is one of a number depending on the position of the target in the list. Hence, variations in successive striking times may be greater than variations in times to strike the first key. Thus, confirmation times would not be expected to show differences with the small number of subjects used in this experiment.

8.3 Subject Preferences for Techniques

Table 3 shows that half the experienced group preferred the step technique and the corresponding figure for the inexperienced group was also 50% - i.e. no difference.

Table 6 shows the number of errors in relation to technique, subject and preference. An interpretation of this is that the technique perceived as requiring least division of attention was preferred, although if an error was made with both, the technique with least penalty for error for that subject was preferred. If an error was made with one technique, the other technique was preferred. The data was insufficient to perform statistical tests on this hypothesis. There was no relationship between preference and order.

8.4 The Effects of Target Position

Graph 1 shows the results broken down according to the position of the target in the menu item list. The select time data shows a distinctive pattern reflecting the search and locate strategy of the subjects. The pattern was common to both techniques and indicates that the middle of the menu item list was scanned first (items 4,5,6) then the items at the top of the list (1,2) then near the bottom of the list (7,8) then,

		Preference Group					
		Step Technique			Digit Technique		
		N	Mean	SD	N	Mean	SD
Mean Confirmation Time	Step Technique	119	5.074	2.117	118	5.624	2.137
		t = 2.02*					
	Digit Technique	118	4.061	1.666	120	4.489	1.614
		t = 1.98*					

N = Number of responses

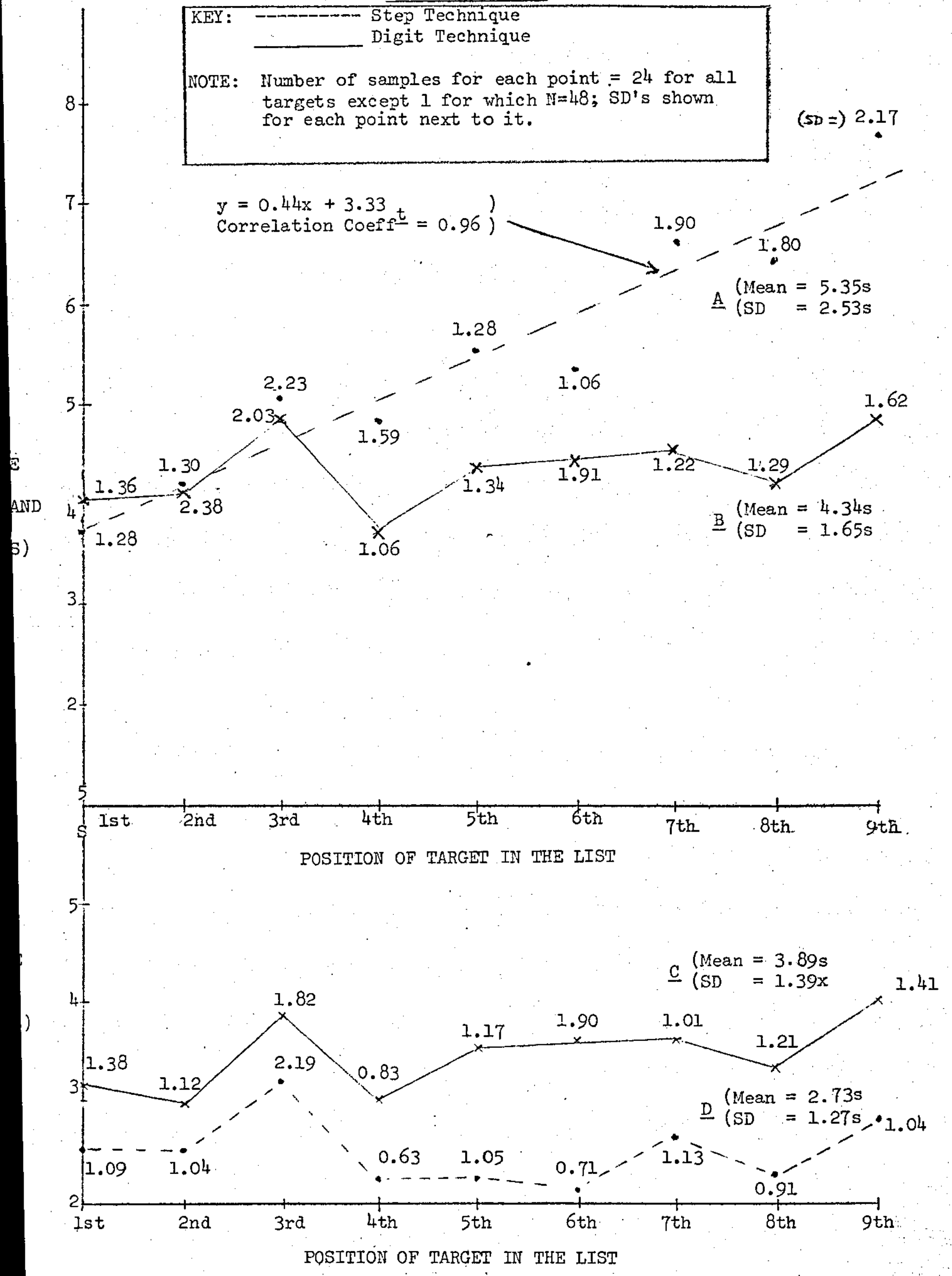
* = Significant at 5% level using t-test

TABLE 6

Mean Confirmation Time as a function of Technique and Preference

GRAPH 1

SUBJECT RESPONSE TIMES AS A FUNCTION OF TARGET POSITION



finally, position 3 and the last position in the list. The pattern is also present in the overall select and confirm time although it is more obvious in the digit technique than the step technique. The effect of target position on confirm time was strongest in the step technique data and a straight line has been fitted (A) to the mean values represented by the points. The proportion of total variation about the mean $(\bar{y})$ explained by the regression is .91, which indicates a good fit. There is no relationship between target position and confirm time for the digit technique. The techniques are equivalent in terms of confirm time if the target appears in position 2.33. More reasonably, this means that, when the target appears in positions 1 or 2, the step technique was faster overall; otherwise the digit technique was faster.

The average step select time (2.77 seconds) was less than the 3.33 seconds found from the straight line intercept of A. This indicates that the time to enter the confirmation key was about 0.80 seconds. The corresponding figure for the digit technique is (4.34 - 3.49) = 0.85 seconds.

These figures are comparable but are much greater than the 0.44 seconds per step in the step technique wherein it is assumed no division of attention occurs. The time to shift attention from keyboard to display and to locate a key is estimated by the difference in select times between techniques to be about 0.96 seconds. Assuming that the location of a key takes about 0.8 seconds (c.f. CR key), the time to shift attention was about 0.2 seconds and the time to locate a target was about (2.77 - 0.2) = 2.57 seconds.

9. Conclusion

The experimental comparison between the step technique and the digit technique showed that:-

(i) The preferred technique depended on the subjects' general experience as well as the specific experience on the task. If no error had been made using either technique, then the technique requiring least division of attention for that subject was preferred. If an error was made with one technique but not the other, then the error-free technique was preferred. The subjects' experience affected their judgement of required division of attention and penalty for error.

(ii) In general, the technique allowing faster select and confirm time depended on the position of the target in the list, but overall was the digit select technique. When the target was in position 1 or 2, the step technique was best; otherwise, the digit technique was better.

(iii) The time to make the first input in the selection procedure (select time) was lower for the step technique than for the digit technique.

(iv) Select time was a function of both the general computer experience of the subject and specific experience on the task. Subjects with more general computer experience responded faster than inexperienced subjects and this was particularly true for the step technique. The step technique required less division of attention than the digit technique and so, by doing it first, the subject acquired greater confidence than if the other technique had been used first. This, however, could be an experimental design artifact since such effects had not been observed in the previous experiment.

9.1 Experimental Hypothesis (2)

The hypothesis that the rate at which syntactic information is transferred does not depend on the way of use of an input device is rejected. The rate depends interactively on the way of use, the task and the subjects' experience.

9.2 Implications for the Model

The experimental conclusions support the view that people behave as if they evaluated and adapted to alternative input methods using a simple estimate of the time needed to input correct data. The mechanisms of adaptation operate on sub-processes such as attention switching and motor processes. The degree to which such adaptation is possible depends on the input device and how it is used. The degree to which it is necessary depends on the person's experience of using that input device.

CHAPTER 5

EXPERIMENT 3

SUMMARY

This Chapter deals with testing the hypothesis that problem solving does not depend on input method and coding of an input device. It was tested using a 3-factor experiment; the factors were input method, problem difficulty and order of presentation. Four different input methods were used; three used a coded keyboard, the other a joystick. Sixteen subjects completed the problem. All had similar levels of computer experience and were equally naive of the problem and the experiment.

Problem solving performance was measured in a number of ways; two classes of measures are identified. These were overall measures of time and frequency, and information flow through the interface. Both sets of measures depended on the input method, the problem difficulty and their interactions. The dependency was that the mean number of steps in an attempted solution decreased in proportion to the time needed to input information. The effect was amplified by increasing problem difficulty. These results imply an adaptive element in the model of interactive problem solving which relates the accurate input of information to the problem solving strategy used.

A verbal protocol and frequency analysis of the input rate of information supported the model of interactive problem solving in that three transfer rates were observed; fast (pre-programmed step); medium (called 'evaluating and recalling') and slow (called 'Planning and organising' a solution).

The hypotheses that problem solving does not depend on (i) the input method and (ii) coding of an input device are rejected. A subsidiary hypothesis was tested that problem solving does not depend on the interaction between user personality and input method. This hypothesis was not rejected at the 5% level of significance but the data showed that the hypothesis may be rejected if it was tested in a more sensitive experiment. The inference was that neuroticism may interact with the input method to affect problem solving performance.

Preference for the different input methods appeared to be based on the subjects' assessment of them using two criteria; first the rate at which they could input information at an acceptable error rate and second, the error proneness of the input method.

CONTENTS

GRAPHS

PLATES

TABLES

FIGURES

1. Introduction

The experimental work of this thesis has shown that with simple tasks, the general and specific computer/task experience and the input method are important in determining the rates of information transfer across the interfaces. However, the information transferred in each transaction in these experiments was independent, and not related to any overall goal of the subjects. This chapter concerns an experiment aimed at testing the experimental hypotheses that problem solving does not depend on the input method and the coding of an input device.

A consequence of testing these hypotheses is that a decision must be taken about the form of dialogue, i.e. the provision of feedback and the computer language structure. In general, real time problem solving involves the presentation of messages at the interface which define the operations to be carried out and the data to be operated on. For example, RUN MYPROGRAM; SET Y=5, etc. In different dialogues, either the operation and/or data may be implicit. For example, input of Y=5 would have the same result as SET Y=5 and RUN may cause the current program to be executed. The exploratory experiment described in this chapter uses a command language in which each command defines both the operation and the operand.

2. Objectives

The specific objectives were:-

(i) To examine the relative performance characteristics (work production) and acceptability of a joystick/special function keyboard and a QWERTY keyboard for solving a computer based problem.

(ii) To examine the relative work production and acceptability between different levels of use of the keyboard in the same task as (i).

(iii) To examine the relationships between work production and acceptability as a function of traits of subject behaviour between and within different input methods.

3. Constraints

The investigation described in this chapter was carried out in a different environment from those described in the rest of the thesis. This is described as follows.

3.1 Hardware

A 16K 12 bit-word PDP-12C computer was used with 2 magnetic tape devices, a general purpose analogue-digital interface, a VR14 point plot display (1024 X1024 addressable points), and a programmable real time clock. Display characters were 2x4 dot matrix and no hardware vector plot was available. Character brightness was not software controllable. The clock was designed for accurate timing (better than ± 1 ms) and was interruptable on external events such as those occurring on an analogue-digital (A/D) channel.

A Honeywell solid state keyboard was interfaced through the A/D channel which also provided a clock interrupt when any key was struck. The decoding of the input was done by software and the keyboard was also flexible in that the positions of the keys could be altered at will.

The joystick and its special function buttons were also interfaced through the A/D channels but while the function buttons provided a digital interrupt for the clock, the analogue joystick output did not. For the joystick, the clock interrupt was generated by software embedded in the task. This detected when certain thresholds had been exceeded. The thresholds were when the 'current position' was taken from one area of the screen to another through a software boundary. This is made clear in section 4.

3.2 Software

The PDP-12 has few constraints due to software since all the devices are available through functions embedded in the high level language it supports, e.g. FOCAL, FORTRAN, ALGOL. The task software was written in the low level assembler language to optimise the response and reliability. The analysis programs were written in FOCAL and accessed the data through the magnetic tapes.

3.3 People

The experiment was carried out within a research group of approximately 20 people, all of whom had similar knowledge and experience of computers. For the majority of these people the time and availability for taking part in the experiment was not a difficulty.

4. Experimental Details

4.1 Design

The general and basic form of the hypothesis to be tested was that the input method does not affect man-computer problem solving. The design used was a 2-way (subjects and treatments) ANOVA. Four levels of treatment are used.

Four different input methods were used and the design was such that each subject was required to solve problems using all four. Each problem had to be unique (because of remembering the solution) and training was required at each level.

Thus, there were 4 problems, each of which was presented in a particular order to each subject.

The order of presentation of the conditions to each subject is shown in Table 1. The table was constructed by rotating the two variables (input method and problem difficulty) in a balanced way so that each combination occurred once only in each block and in a different order in each block.

Order and Combination of Experimental Conditions for each Subject (S)					
S1	A1,	B2,	C3,	D4	Block 1
S2	A2,	B3,	C4,	D1	
S3	A3,	B4,	C1,	D2	
S4	A4,	B1,	C2,	D3	
S5	D1,	C2,	B3,	A4	Block 2
S6	D2,	C3,	B4,	A1	
S7	D3,	C4,	B1,	A2	
S8	D4,	C1,	B2,	A3	
S9	C2,	D1,	A4,	B3	Block 3
S10	C3,	D2,	A1,	B4	
S11	C4,	D3,	A2,	B1	
S12	C1,	D4,	A3,	B2	
S13	B2,	A1,	D4,	C3	Block 4
S14	B3,	A2,	D1,	C4	
S15	B4,	A3,	D2,	C1	
S16	B1,	A4,	D3,	C2	

Table 1.

Key
A = Input Characteristics
1-4 = Problem number at a task difficulty level

4.2 Sessions

A pilot experiment had indicated that the presence of the experimenter (E) during the course of an experiment had a significant effect on the way a subject (S) performed the task. Therefore, the experimental procedure allowed for minimal S-E contact mainly during the training phase. The procedure for each S was similar and followed the following pattern:-

(1) E read aloud information to the S which generally described the aims of and his part in the experiment.

(2) S read the Task Description.

(3) E demonstrated the task with the first trial conditions using a demonstration problem. S was allowed to question E.

(4) S was shown the pre-trial training problem and its function was explained.

(5) S was presented with an on-line EPI* questionnaire and was left alone to complete it.

(6) S was given the four conditions in the predetermined design order of the problem and input method with the pre-trial problem preceding every trial. E was absent during the problem solving and S was not allowed to use anything but the computer. A trial was completed only when S had solved the problem. No time limit was given for any trial.

(7) Finally, S was asked to:

    (a) Rank and comment on the input method.

    (b) Comment on his problem solving strategies.

---

* Eysenck Personality Inventory.

4.3 Task

The development of a suitable problem solving task for this investigation was based on the following assumptions:

(i) The display variables should not interfere with the variables under investigation.
(ii) The number of steps to solution should be controllable.
(iii) All steps should be of equal difficulty (homegeneity).

A special problem solving task was developed which was based on the idea of a three-dimensional maze, (2 dimensional mazes have commonly been used in problem solving research). The given situation is a position in a room of a 'building' with 9 rooms on each of 4 floors, each of which is interconnected by walls. The desired condition is to leave the exit having reached it without having attempted to go through solid floors, walls or ceilings in the process.

The following section describes the task in detail.

(a) Task Description

The general form of the task was a maze in three-dimensions with a single entrance or starting point and a single exit or finishing point. The basic structure of the maze can be likened to a building having four floors with nine rooms on each floor.

A pictorial representation is shown below. (Figure 1).

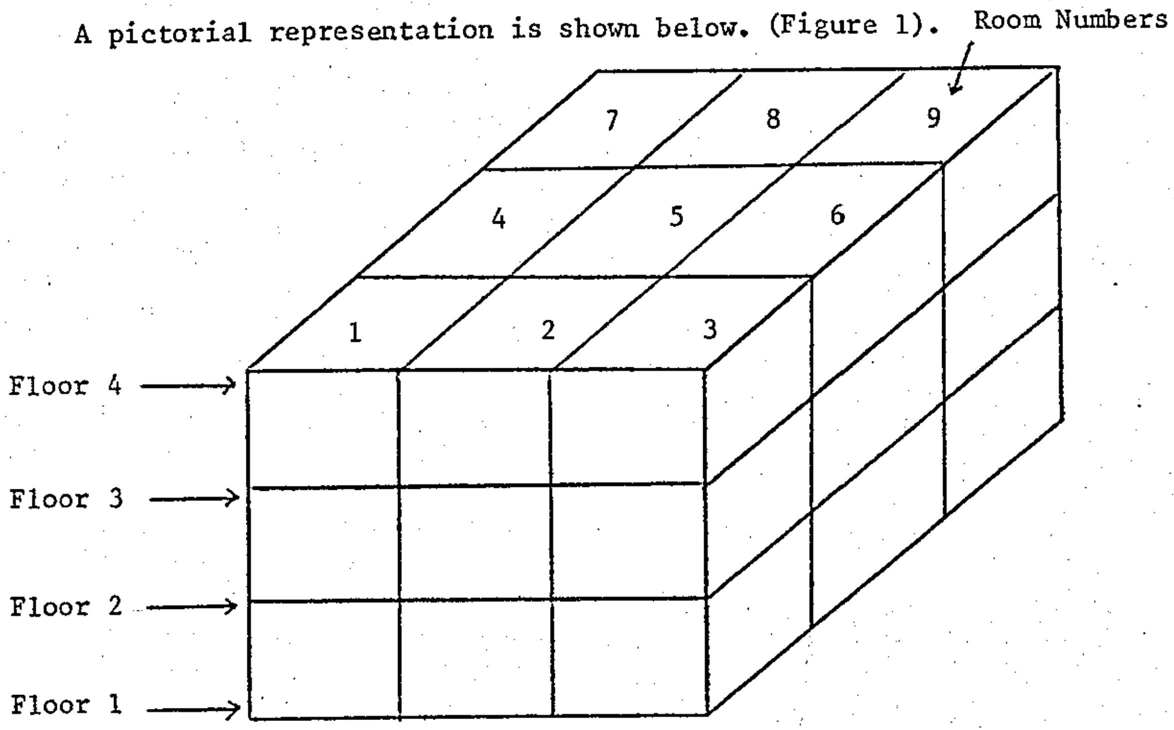


Figure 1

The starting point was always in Room 5 on Floor 4, although the exit could be on any floor in any outward facing position. As in a real building, some rooms had interconnecting doors and others had none. Similarly, some floors had staircases in rooms and others had none. The task, as in more conventional 2-D mazes, was to find the way out of the 'building' as quickly as possible and to do so without making errors of trying to go through brick walls or through solid floors. If any errors of this kind were made, the maze had to be attempted again until an error-free exit was made.

The maze does not exist in three-dimensions either as a real object or as a projection; it was presented as four floor plans (each being in two dimensions) on the PDP-12 visual display unit (see Plate 4). The idea of using the 'building' analogy was to simplify the explanation of the interconnections between the plan views.

Consider floors 4 and 3 as shown below. (Figure 2).

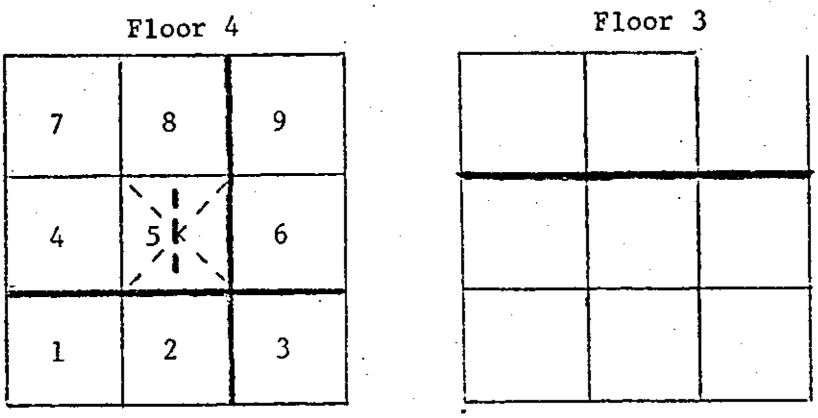


Figure 2

(The room numbers are shown for convenience of description; they were not displayed in the real task.)

The conventions of the display were as follows:

1. The current position in the maze was always represented by a vertical line as shown in the normal starting position of Floor 4 room 5, above.

2. A solid line represented a brick wall between two rooms i.e. one should not try to pass through it! Conversely, one may freely move where there are no solid lines.

3. A diagonal in a room (e.g. Floor 4, room 5) indicates that there was a solid floor in that room, and one should not attempt to go through it. Conversely, where there was no diagonal in the room, one may go down to the same room of the next floor.

4. The exit was immediately obvious as the only gap in an external wall, e.g. Floor 3, room 9.

These very simple rules will become more readily understood by using the example in Figure 2 to show how one gets out of the exit. There is the assumption that only 2 floors exist. Using an optimum goal-oriented strategy rather than the obviously less efficient trial and error method a solution argument goes as follows:

Step (a) The exit is in floor 3, room 9. Since there are solid walls around this room, we can only get into it from another floor. The only other floor is floor 4.

Step (b) Our problem is to get to room 9 from room 5 on floor 4 given that there are solid walls to be circumvented. This is essentially the same as getting from room 5 to room 6.

Step (c) The essential part of the solution is the recognition, that while there is no freedom of movement between rooms 5 and 6 on floor 4, there is on floor 3. Thus, we now need to get to floor 3, room 5.

Step (d) The diagonals mean we cannot go down to floor 3 from our present position, but move to a position when we can, i.e. room 4 (or 8).

Step (e) Our solution is therefore as follows:-

Start: Floor 4 Room 5
Move 1: Floor 4 Room 4
Move 2: Floor 3 Room 4
Move 3: Floor 3 Room 5
Move 4: Floor 3 Room 6
Move 5: Floor 4 Room 6
Move 6: Floor 4 Room 9
Move 7: Floor 3 Room 9
Move 8: <u>Out</u>

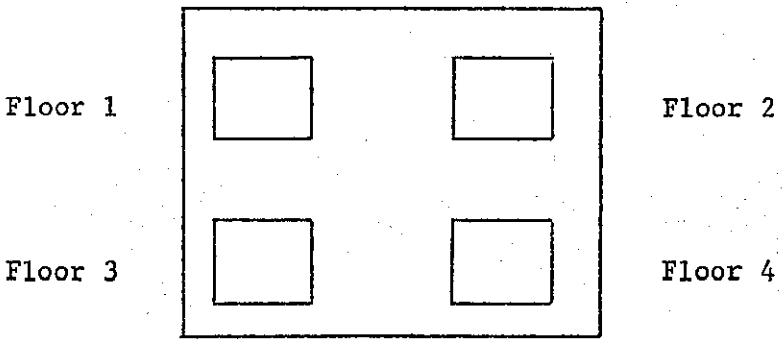
There is one very important and at first frequently confusing aspect to this convention; this can best be illustrated by considering the following situation. Suppose that your position is floor 3 in room 5. Can you move up to floor 4? The answer is 'no' because there is no 'hole' or 'staircase' to move into in room 5 on floor 4. Errors can be made because there was no indication in room 5 on floor 3 that this was so. This is a very important point to remember. If there was not a 'hole' or 'staircase' (indicated by a diagonal on the floor above) one could not move up to the next floor. If this was attempted, an error had been made and this meant repeating the task from the beginning. Logical errors of this kind were fed back to the subject by not moving the displayed position as requested.

There were three basic facilities to help in the solution of this maze:-

(1) One could go back to the entrance start position (floor 4, room 5) at any time. All previous errors were cancelled on this action (Restart).

(2) One could 'mark' the current position in any room or floor to remind oneself of a previous action. Using the facility forced a restart through the maze (as in (1)).

(3) Since only one floor was displayed at any time, there was a facility to display all four floors simultaneously ('aid'). Using this also forced a restart (as in (1)). The display of the floors was as shown below.



(b) Problem Difficulty

The problem difficulty was controllable by altering both the number of possible routes and the number of dead-ends between the given position and the desired position. This was achieved by altering the boundaries between each region and floor. The experimental design required 4 problems at the same level of task difficulty, a training problem at a low level of difficulty and a practice problem.

A 2-dimensional representation of the 3-dimensional task was developed and used for designing mazes of known difficulty. A small pilot study was carried out using 3 subjects to examine the range of difficulty and its sensitivity to the variables mentioned. Six mazes were developed and are shown in Appendix 1 (Figures 1-12). Their characteristics were as shown in Table 2 below.

Maze No.	Minimum Number of Solution Steps	Number of Dead ends	No. of loops		Total Count
			≤4 moves	>4	
1	18	6	3	2	27
2	19	6	0	2	27
3	20	4	3	1	28
4	24	3	4	0	31
5	27	0	0	0	27
6	22	5	-	-	Not Used

Table 2: Problem Difficulty

Although the figures in Table 2 give some indication of the problem difficulty, they may be misleading in that, for example, the problem with the highest total number of steps (Maze 5) had the lowest difficulty since the choice of moves was limited to one direction through the maze. The small pilot study (using 3 people and maze 6) showed that the task was not completely homogeneous since steps from floors to the higher floors were more difficult (because of the need to remember if it was possible) than steps from floors to lower floors. It was an assumption that the use of a mark would avoid this problem if it was used to remind the subject that he could or could not go step up to the floor above. The pilot study showed that the mark was not used by all subjects in this way because of possible ambiguity of interpretation.

The task was capable of many different levels of problem difficulty from trivial (maze 5) to impossible (no route to exit). The 4 chosen were comparable with each other in homogeneity and other characteristics within the range and therefore suited the experimental design.

5. Independent Variables

5.1 Input Method

(a) Input Devices

Two types of input device were used. These were:-

(i) a joystick with 6 special function buttons, and

(ii) a Honeywell solid state keyboard (silent in operation).

Plates 1 and 2 show the physical characteristics of each device.

The joystick control was centralised by means of elastic bands and had little stiction. The full range of physical movement gave a signal which matched the inputs of the computer (± 5 volts). The depression pressure of the special function buttons was slightly greater than a standard ASR 33 teletype keyboard. The pressure needed to operate the keyboard was less than that of an ASR 33 and the layout of keys could be easily altered.

(b) Ways of Use (Coding)

The joystick signal was continuously sampled and converted into x-y co-ordinates such that an 0.8" movement of the top of the stick produced a displacement of the displayed mark by about 1". The control was direct and not aided in any way. The x-y co-ordinates were used to decide which areas of the screen the current position was in terms of a region of a floor. This was handed over to the computer as a logical signal for a new desired position. The special function buttons were directly connected to interrupt lines and decoded into appropriate messages for the computer.

The keyboard was used with two coding levels; a single character message (no confirmation needed) and multiple character (two word) command messages (confirmation of entry needed). Two codes for single characters were used; mnemonics and cursors.

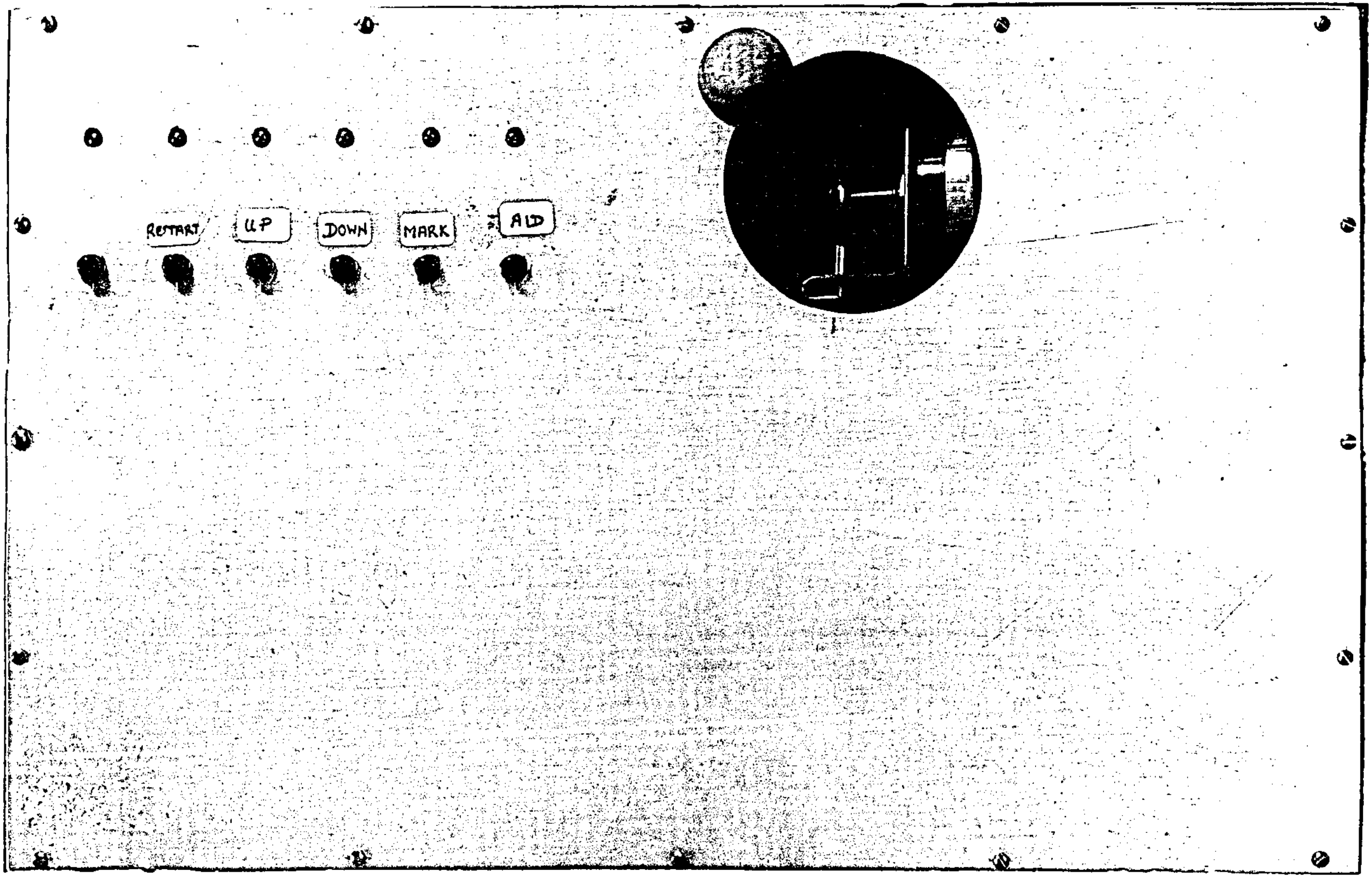


Plate 1 Joystick Control

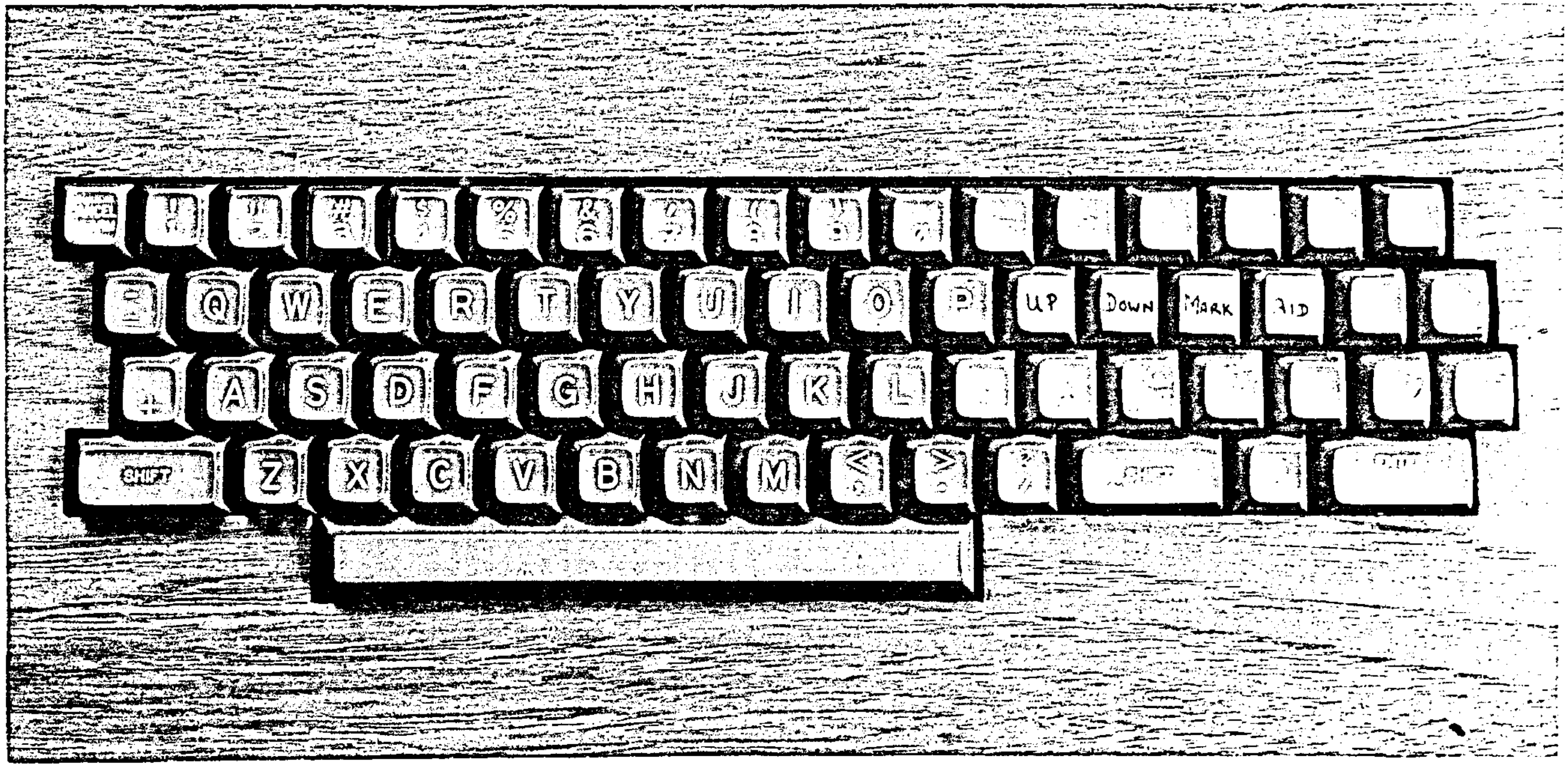


Plate 2 Cursor and Special Keyboard

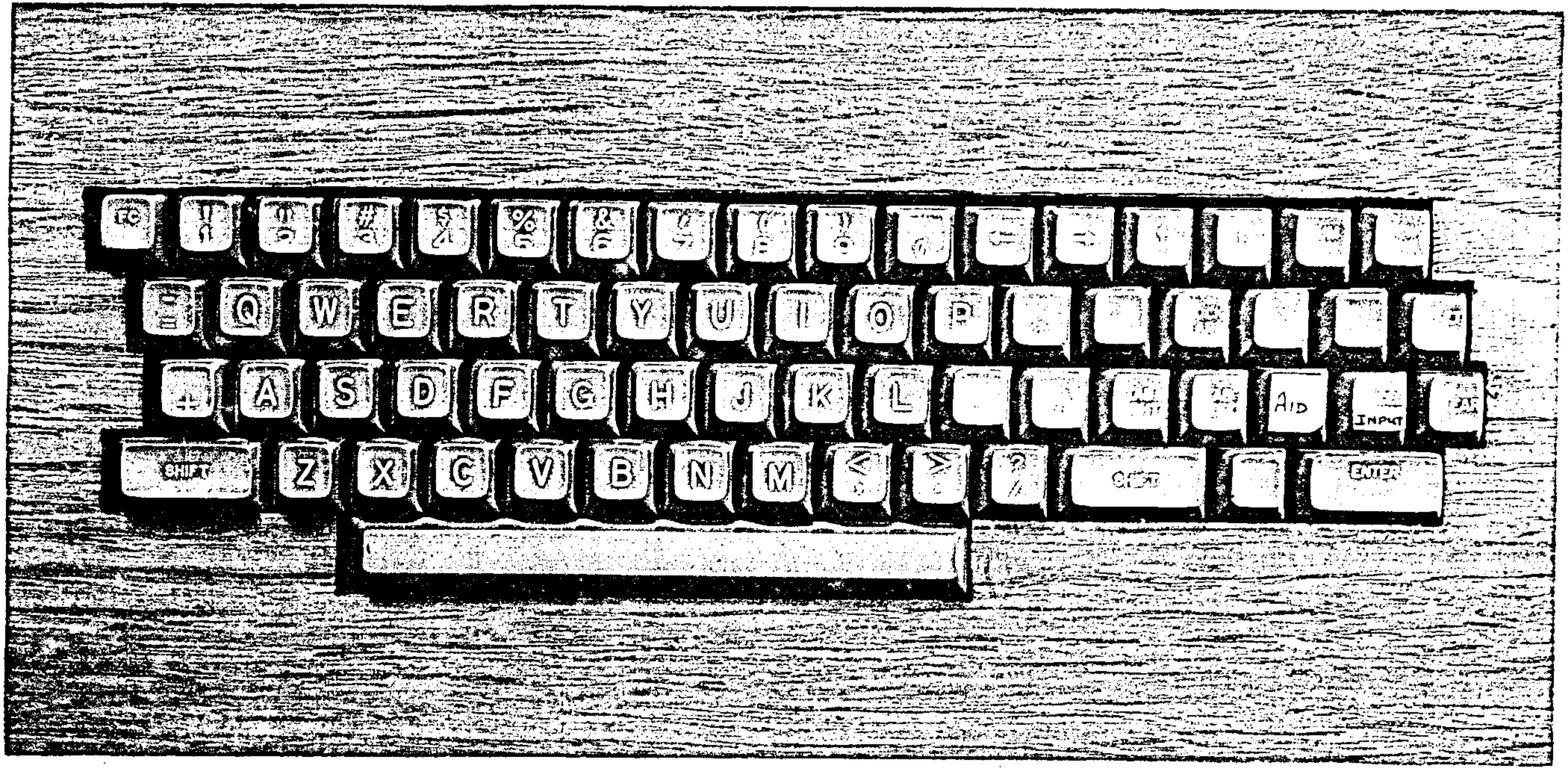
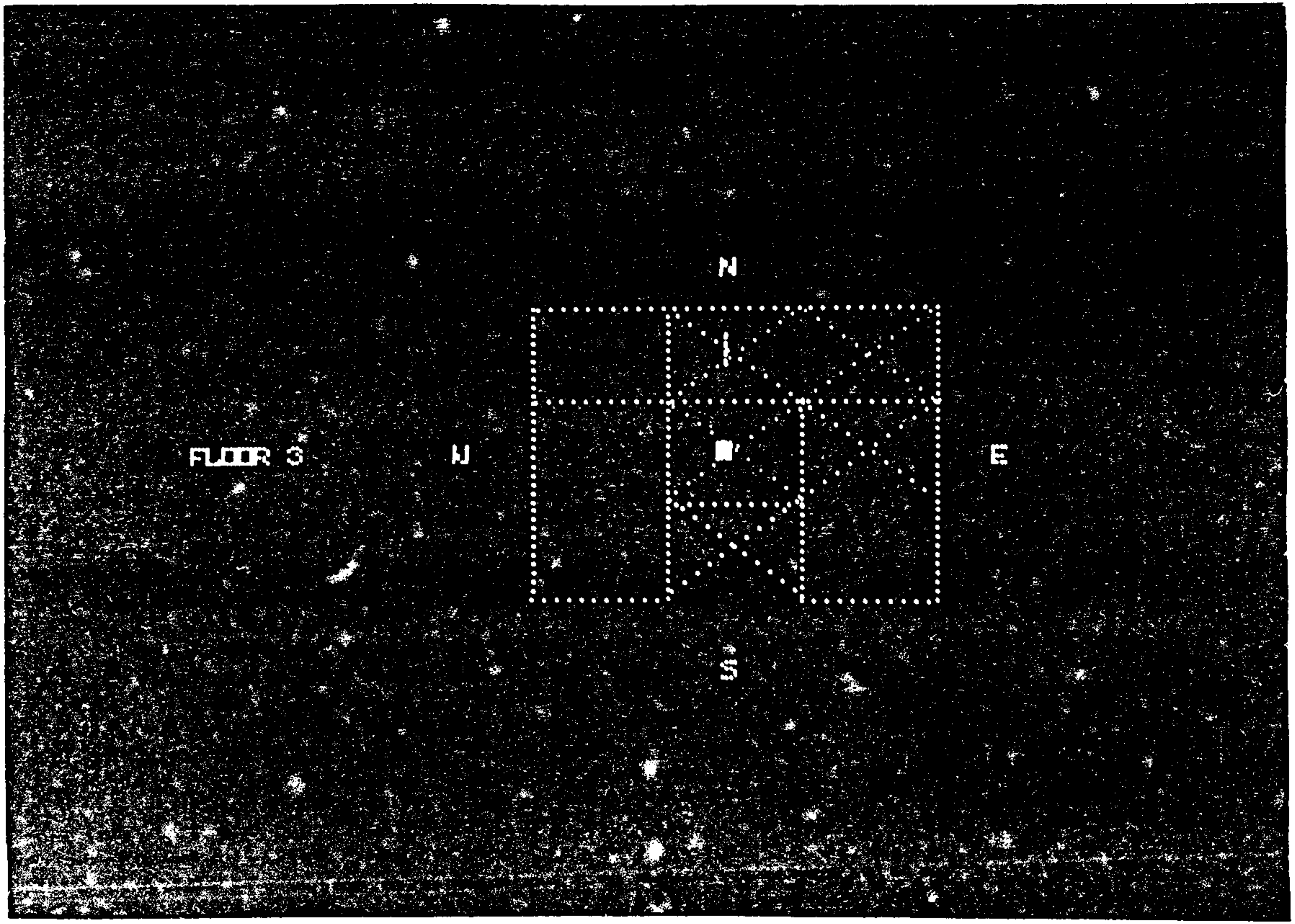


Plate 3 Single key/QWERTY keyboard with Special Keys

Plate 4 The Display as seen by the Subject



The four experimental input methods (referred to as A, B, C and D) are as follows:-

A. Joystick Input

A joystick and 5 special function push buttons marked <up> (go up to next floor), <down> (go down to next floor), < mark > (mark current position), <aid> (display all four floors), and <restart> (go back to beginning and cancel all errors), (see Plate 1).

B. Cursor Input

A QWERTY solid state keyboard using cursor keys (↑ → ↓ ←) for screen movements and five other specially labelled keys as for the joystick assembly. The nine keys were grouped together as 2 rows of 4 (with restart below them) on the right hand side of the keyboard (see Plate 2). No confirmation key was used and mis-typed characters were ignored, or if appropriate, executed. For this option, the current position was moved stepwise from region to region.

C. Alpha Input

A QWERTY solid state keyboard where the following keys were used for mmemonics:-

N - meaning move 'north' on the screen
S - meaning move 'south' on the screen
E - meaning move 'east' on the screen
W - meaning move 'west' on the screen
U - meaning go up to the next floor

D - meaning go down to the next floor
M - meaning mark the current position
A - meaning display all four floors simultaneously
R - meaning restart

(see Plate 3)

No confirmation was used for this input. Typed keys were either ignored or, if one of the above, executed.

D. Verbose Input

A QWERTY solid state keyboard where the following had to be typed as separate characters without error. They are self-explanatory:-

MOVE NORTH;
MOVE SOUTH;
MOVE EAST;
MOVE WEST;
GO UP;
GO DOWN;

(see Plate 4)

Special keys were used for:

<MARK>
<AID>
<RESTART>
<CANCEL INPUT>(for typing errors)
<ENTER>

In this group, a confirmation key (enter) was needed for commands but not special keys. Command Typing errors were corrected by deleting the whole of the input line, then re-typing.

The following should be noted:-

(i) For the C and D conditions, the floor was displayed with (N, S, E, W) in the appropriate sections of the display to remind the user of the convention.

(ii) In all cases there was no TTY or VDU 'echo' of input; the feedback was in terms of changes in the information display, i.e. current position in the maze.

(iii) Characters enclosed thus, < > , are single key operations.

(iv) '▿' means a space must be typed.

(c) Formal Description

(i) Source Primitive 'Change Position in the Maze to _ '

	INPUT DEVICE				
	Joystick with SF Buttons		Keyboard		
	Control	Buttons marked up down	← ↑ ↓ → ↓	N, S, E, W, U, D ↓ ↓	<GO> ∇ <[]> ↲
Syntactic	(p)d	(p)d	(p)d	(p)d	d...d d...p
Semantic	(p)d	(p)d	(p)d	(p)d	d...d d...p
Pragmatic	(p)d	(p)d	(p)d	(p)d	d d...p
	A		B	C	D

(ii) Source Primitives ('Mark current position' denoted by M

('Display aid' denoted by A

(Put current position at beginning of maze
(denoted by R

('Cancel input' (D only)

	INPUT DEVICE				
	Joystick with SF buttons	Keyboard			
	↓	M A R	M A R	<COMMAND>	<cancel input>
	buttons	grouped keys	QWERTY keys		
Syntactic	d(p)	d(p)	d(p)	d...p	p
Semantic	d(p)	d(p)	d(p)	d...p	
Pragmatic	p(d)	p(d)	p(d)	d...p	
	A	B	C	D	

KEY: p means prescriptive

d means descriptive

( ) means implicit processing at that level

< > encloses strings of characters (d...) which are receiver primitives

[ ] these enclose the strings as given in the previous description of D, i.e. UP, DOWN, GO, MOVE, etc.

∇ is a space character.

DATA STARTS IN DECIMAL BLOCK 24
TIME HISTORY OF DATA

LEGALITY	REGION	FLOOR	TIME (MINS)	U.R.T. (SECS)	VERBAL PROTOCOL
0	5	4	0		Trying to memorise what the floor looks like.
0		3	0.550	32.98	Down quick.
0		4	0.701	9.080	Up again.
0		3	0.866	9.880	
0	4		0.993	7.660	
0		2	1.094	6.030	Trial and error at moment.
0		3	1.190	5.770	Try again.
0	5		1.450	15.59	See exit on Floor 2.
0		4	1.474	1.450	
0		3	1.670	11.76	Top floor, very good.
0	4		1.913	14.57	Just repeating myself he
0	5		2.078	9.910	
0		4	2.139	3.650	Oh no, wonder why I am so slow.
0	4		2.661	31.36	Fourth floor again.
0	1		2.689	1.630	Can't find direction.
0		3	2.752	3.810	
0		4	3.018	15.95	Fatal mistake! I have to go on the floo to get out.
0	4		3.095	4.630	Lets get back to busines
0	5		3.129	2.030	
1		5	3.196	4.040	
0		3	3.393	11.83	
0	4		3.438	2.660	
0		2	3.517	4.760	
0		1	3.574	3.410	
0		2	3.694	7.220	Now to ground floor.
0	1		3.757	3.790	
0		1	3.817	3.550	
0	2		3.955	8.300	
0	5		3.993	2.270	
0		2	4.046	3.180	
0	8		4.103	3.450	Very silly to think ...
1		3	4.174	4.270	
1		3	4.292	7.080	Ridiculous OOPS! Can't go up
0	5		4.372	4.790	
1		3	4.407	2.090	I need these marks after all.
		<MARK>			
1	2		4.723	18.93	
0		1	4.794	4.300	
0	2		4.926	7.920	Hit brick wall there – stupid.
0	3		4.945	1.130	
1		2	4.970	1.480	
0	3		5.089	7.170	
1	0		5.124	2.110	
0	2		5.194	4.180	
1		2	5.238	2.630	
0	2	<MARK>	5.265	1.610	
0	5		5.356	5.500	
0		2	5.399	2.530	
0	8		5.458	3.570	
0		1	5.523	3.870	

TABLE 3 (CONTINUED)

```
0          7
0                 2       5.733       6.330       That's the right p
                                                  Back to beginning-
RESTART    5      4       5.773       2.370       I've made mistakes
                                                  off we go

0          4              6.010       14.23
0          1              6.091       4.890
0                 3       6.144       3.140
0                 4       6.374       13.81       May have forgotten
                                                  I've done

0          4              6.420       2.770
0          5              6.557       8.200       You know, I do bel
0                 3       6.625       4.100       I have.

0          4              6.707       4.910
0                 2       6.780       4.380
0          1              6.838       3.510
0                 1       6.919       4.850
0          2              6.990       4.230
0          5              7.041       3.110
0                 2       7.078       2.210       That's better. I
0          8              7.149       4.220       remember it now.
                                                  Nearly out.
0                 1       7.220       4.280
0          7              7.262       2.520
0                 2       7.301       2.340


NO OF TYPING ERRORS IS    1
FINISHED
```

A Time History of the Interaction.

TABLE 3

6. Dependent Variables and Measures

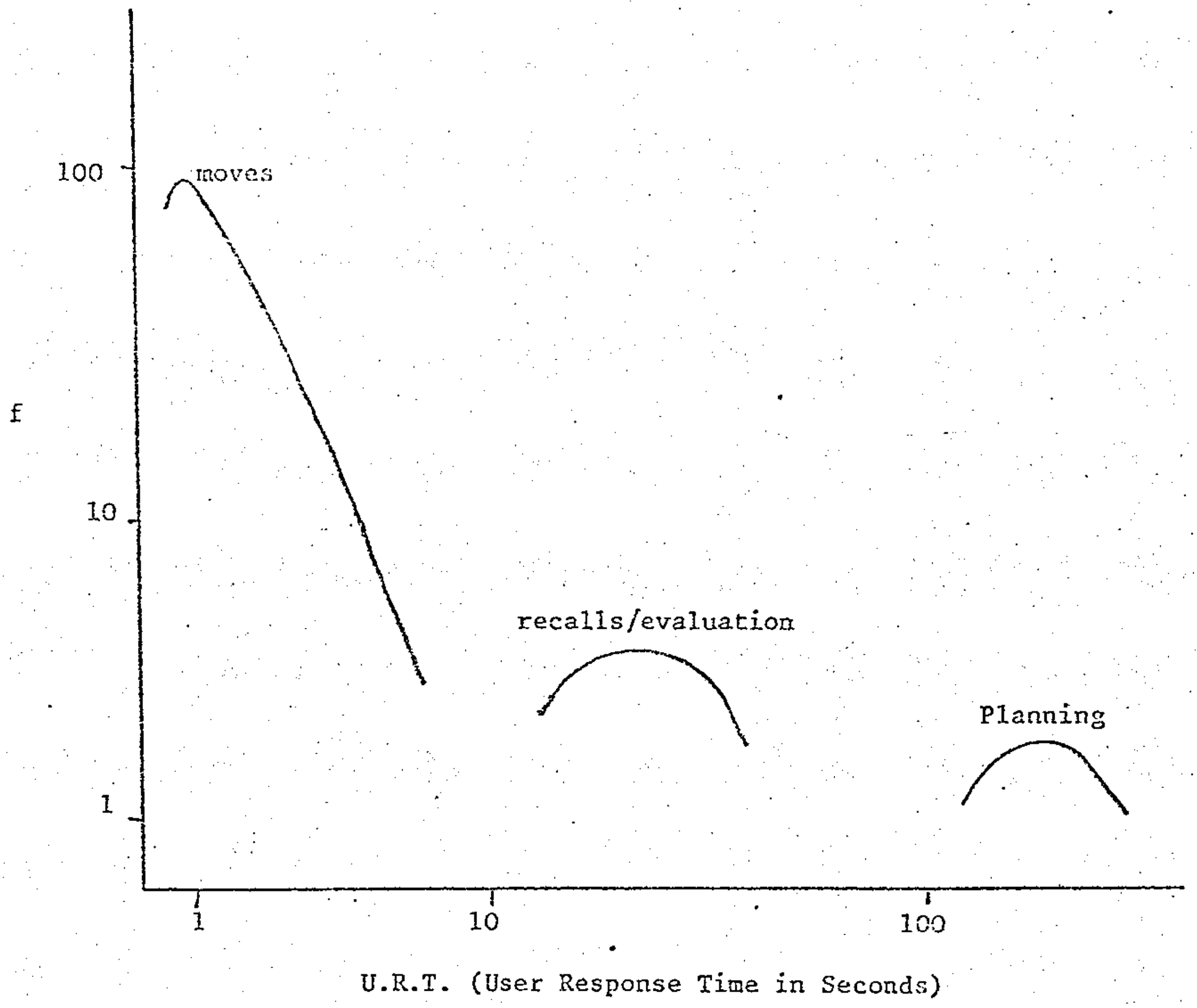
6.1 Problem Solving

Problem solving has been characterised by other workers using various measures on a variety of tasks. In order to determine which measures to use and their sensitivity to the independent variables, a number of measures were taken. Some were based on the number and time of individual steps to solution and others based on overall measures, such as the time to solve the problem.

Table 3 shoes typical data which was collected during the pilot experiments. The computer collected the activity and timing information while the verbal protocols were recorded on a small tape recorder.

The basic measure of subject's performance was the user response time (U.R.T.) which was the elapsed time between the time when the computer could accept information to the time when pragmatic processing requested by the subject, either implicitly or explicitly.

Graph 1 shows the distribution of URT's for particular conditions and a subject. There are three distinct peaks in the distribution and an examination of the verbal protocol showed that the first peak of the distribution may be referred to as 'pre-programmed' error-free input; the intermediate peak as recall or evaluation and the third peak as planning activity at a more strategic level.



Graph 1. Distribution of U.R.T. for S9, [illegible]
SOLVING MAZE 2 WITH INPUT
CHARACTERISTICS TYPE B.

The distributions are similar to those discussed by Bradley (1975) in the operation of a push button. He showed that skewness of the three peak distribution meant that parametric tests of differences were not adequate and led to false conclusions. It is also interesting that he found the second peak was for two- and more-errors. Thus an interpretation of the second peak found in the problem solving data is that it may be a consequence of input errors.

The third peak of long input times arose because subjects were using the aid to memorise a solution, rather than correcting input errors.

Derived measures of problem solving were based on:-

(i) the distribution of URT's
(ii) time and error scores for solving the problem, and
(iii) the number of trials needed to solve the problem.

These are labelled and listed in Table 4.

Of the 11 measures, those based on the URT distribution are dependent on the subjects particular URT distribution in the experimental condition. These are measures 5, 6, 7, 8 and 10 and were derived for each subject.

Identification	Name	Measure	Meaning
M1	PRT	Principle Reaction Time	Reaction time per move through the spiral maze 5 (wrt)
M2	NOTS	Number of Trials	Number of starts at the entrance of the maze before exiting
M3	OVTM	Overall Time	Time in minutes to get out of the maze
M4	TOTMV	Total number of Moves	Total number of moves to get out of maze
M5	NOPLANS	Number of Plans	Number of times the aid was used plus a long study of a floor
M6	NORCLS	Number of Recalls	Number of times S could be inferred to be recalling/evaluating his position
M7	MPT	Mean Planning Time	Self explanatory
M8	MMAA	Mean number of moves between one plan and next planning activity	Self explanatory
M9	ERRS	Number of errors	Total number of errors before successful exit is made
M10	TRM	Mean time of recalls	Self explanatory
M11	TMM	Mean time for move	Self explanatory

Table 4
Derived Measures

The URT's for the spiral maze is included in the list because it is used to correct the measures for individual differences. For example, if subject 5 was always faster using the joystick in the maze than subject 6, comparisons of the evaluation and planning times for these subjects and that input would be confounded. Therefore, the measures already described were corrected for this individual difference and new measures were derived. Similarly, the number of trials may confound the measures taken such as the total number of moves to get out of the maze, because they may be correlated. Therefore, the measures sensitive to the number of trials were normalised into further derived measures. Table 5 shows the additional derived measures.

Measure Identifier	Treatment	Meaning
M3,M4,M5,M6,M7,M8,M9	Divide by M2	All measures per trial
M10,M11	Subtract M1	Gives corrected planning and evaluation, decision time respectively
M10,M11	Subtract M1 and divide results by M2	As above per trial

Table 5: Additional Derived Measures

6.2 Acceptability

The meaning of acceptability in this investigation was taken to be a preference for particular input methods. Therefore, it was measured on completion of all the experimental conditions by asking each subject to rank the four input methods in order of preference.

7. Subsidiary Variables and Measures

These variables were introduced for the purpose of testing hypotheses outside of the experimental design. The independent variables were subject personality traits of introversion and neuroticism and were measured using the Eysenck Personality Inventory (EPI). The dependent variables were derived measures of performance as outlined in the previous section.

8. Analysis of Results

8.1 Work Production

The raw data shown in Tables 1-11 of Appendix 1 was subjected to a 2-way classification ANOVA - input method (treatments) versus subjects (replications). The measures which gave significant results using a fixed effect model are given in the Appendix. Table 6 shows a summary.

Measure	Treatment	F-value
PRT	I	260
NOPLANS*+	I	3.16
MMAA+	I	3.10
TRM^{o}	I	5.38
TMM	I	79.3
TRM	0	5.05

* This measure violated ANOVA assumptions
o This measure was associated with order effects
\+ This result fails with conservative F test; df = 1.15 but is significant using Hotelling's T^2 test ($T^2 > 16.3$)

Table 6: Significant Measures

Those measures showing effects due to I were re-tested for order effects (by casting the data in an "order" by "subjects" table and repeating the ANOVA). Order effects were found for TRM.

TMM and MMAA showed neither order effect nor effects depending on maze differences (tested by re-casting the ANOVA). t-tests were carried out on the significant results (excluding NOPLANS). The results are in Appendix 1 (Table 15).

Table 7 shows the inter-correlations between the paired derived measures across all subjects and conditions. It shows that, of the measures which are significant, only 2 (MPT and PRT) are independent and the others are related. Multi-variate analysis was not carried out on all the measures.

8.2 Acceptability

Table 8 shows the preferences for the input methods for each subject. The Friedman 2-way ANOVA shows that there was a significant difference ($p < .001$) between the ranks of the input methods. The overall ranking was B, A, C, D (i.e. B preferred most, D least preferred). This ranking agrees with the ranked performance measure M8 (MMAA) but not with any other measure.

8.3 Subsidiary Variables

The subsidiary variable was the personality trait of the subjects and these were characterised by two scores (shown in Table13 of Appendix 1). The neuroticism (N) and extroversion(E) scores were ranked and the Spearman Rank coefficient calculated to be -0.025. This was not significantly different from zero, i.e. the E and N dimensions are orthogonal.

The median scores for each personality trait were calculated and subjects cast into two groups for each trait according to whether their score was above or below the median. A similar process was carried out on the scores for each measure and a 2x2 contingency table constructed. This had the form shown in Table 9.

Model Y = BX + A							
Y		X		Correlation Coefficient R	A	B	% Variance acct'd for
IDENTIFIER	NAME	IDENTIFIER	NAME				
M3	OVTM	M5	NOPLANS	0.76	5.61	1.74	58
		M6	NORCLS	0.77	4.62	0.83	59
		M2	NOTS	0.74	4.92	1.21	54
		M9	ERRS	0.55	8.81	0.35	30
		M4	TOTMV	0.63	5.76	0.08	40
M2	NOTS	M5	NOPLANS	0.94	1.04	1.29	88
		M6	NORCLS	0.83	0.96	0.54	69
		M4	TOTMV	0.69	1.65	0.05	47
M4	TOTMV	M9	ERRS	0.92	36.99	4.81	85
		M8	MMAA	0.53	20.94	3.09	28
M5	NOPLANS	M9	ERRS	0.65	2.00	0.18	42
		M6	NORCLS	0.79	0.26	0.38	63
M9	ERRS	M1	NOTS	0.65	1.30	1.67	42
M10	TRM	M11	TMM	0.65	4.52	1.99	42
M1	PRT	M11	TMM	0.65	1.22	0.93	42

NOTE: All missing pairs had coefficients less than .5

Table 7

Intercorrelations Between Derived Measures

S. No.	Condition				
	A	B	C	D	
1	1	2	3	4	
2	3	2	1	4	
3	2	1	3	4	
4	1½	1½	3	4	
5	4	2	1	3	
6	2½	1	2½	4	
7	1	2	3	4	
8	3	1	2	4	
9	3	2	4	1	
10	2	1	3	4	
11	1	2	3	4	
12	1	2	3	4	
13	3	2	1	4	
14	2	1	3	4	
15	3	1½	1½	4	
16	3	2	1	4	
Rj	36	26	38	60	
$(Rj)^2$	1296 +	696 +	1444 +	3600	= 7016

$$\chi^2_r = \frac{12}{16x4x5} x \quad 7016 - 48x5 \quad \simeq \frac{3}{80} x \quad 7016 - 240$$

$$= \frac{3}{80} x \quad 6776 \qquad (df = 3)$$

$$\underline{\chi^2 = 253.1 \ (p < .001)}$$

Input Condition	A	B	C	D
Overall Rank	2	1	3	4

Table 8 Friedman 2-way ANOVA on Ranks of PREFERENCE FOR INPUT CHARACTERISTICS

		INPUT CHARACTERISTICS	
Trait		A	D
Group	N↑	a	b
	N↓	c	d

KEY: N = neuroticism

↓ = low; ↑ = higher than the median N score

a = the number of subjects with N↑ and with scores on the measure of performance using A above the median

Table 9

Showing the Construction of Fischer's Exact Probability Tables

Measure	E x I	N x I
M2	0.49	0.49
M3	0.28	0.37
M4	0.32	0.30
M5	0.47	0.16
M6	0.51	0.29
M7	0.51	0.49
M8	0.41	0.24
M9	0.39	0.39
M10	0.35	0.11
M11	0.36	0.24
M3/Trial	0.27	0.24
M4/Trial	0.38	0.36
M5/Trial	0.29	0.18
M6/Trial	0.42	0.42
M7/Trial	0.42	0.24
M8/Trial	0.35	0.13
M9/Trial	0.38	0.38
M10 corrected	0.38	0.22
M11 Corrected	0.51	0.46
M10 Corrected per Trial	0.42	0.45
M11 Corrected per Trial	0.45	0.30

KEY: I = Input Characteristics
N = Neuroticism
E = Extraversion

Table 10

Fischer Exact Probabilities for Input Characteristics A/D and Personality Traits E,N.

The Fischer exact probability test was then applied to the table which resulted in a probability for the distribution of the frequencies in the Table.

Table 10 shows the results over all the measures and for all combinations of input characteristics and personality trait.

9. Discussion of Results

9.1 Problem Solving Performance (Work Production)

(a) Measure Sensitive to Input Method (I)

The measures sensitive to input method are M8 (MMAA) and M10 (TRM-PRT) per trial, and the uncorrected M11 (TMM). All are significant at, at least, the 5% level.

Measure	Input (I)			
	A	B	C	D
M8 (MMAA)	24.22	22.2	14.8	10.95
M10 (TRM)	9.08	10.25	11.59	18.91
M11 (TMM)	2.09	2.73	3.08	8.74

Table 11: Table of Measures for Input Method

Table 11 shows that there was a trade-off between the mean number of moves between planning and the mean evaluation/recalling time per trial. Thus, the longer the evaluation time, the smaller was the number of moves between planning activities. A to D are ranked 1-4 on TMM; these data show that the lower this rank, the greater the amount of time spent evaluating/recalling and the smaller the number of moves per plan.

An explanation could be made in terms of 'cost' to the subject of using particular input methods. The higher the cost, the longer the evaluation/recall times and the smaller the number of moves per plan. The 'cost' may simply be in terms of the time taken to input information. This was measured by M11 (TMM) whose ranking corresponds with the ranking of A, B, C, D as 1, 2, 3, 4 (smallest time is ranked at 1). If so, then this association may possibly be explained by the effort needed to sustain short term memory of the solution information while coping with the input characteristics.

(b) Measures Sensitive to Problem Difficulty

Those measures which are task sensitive rather than input sensitive have been ranked in Table 12.

Table 12 shows that all measures do not agree on the rank order so that there was no absolute measure of difficulty. The ranks given by M7, M8 and M10 correspond to the rank of the minimum number of steps to solve the problem given in Table 2. Hence, the inherent task difficulty largely determined the number of steps in the attempt to solve the problem and the frequency and length of the longer subject response times.

Measures	Maze Label			
	1	2	3	4
M3 (OVTM)	1	3	2	4
M4 (TOTMV)	1	3	2	4
M5 (NOPLANS)	1	3	2	4
M6 (NORCLS)	1	3	2	4
M8 (MMAA)	1	2	3	4
Corrected M7 (MPT-PRT)	1	2	3	4
Corrected M10 (TRM-PRT)	1	2	3	4

NOTE: Rank 1 = smallest value

Table 12: Table of Ranks of Maze Difficulty on Measures

(c) Measures Sensitive to Order Effects

These were tested by recasting the ANOVA table into I x O order and M x O order and carrying out a 4x4 ANOVA. There were no significant order effects for any measure alone.

9.2 Subsidiary Measures

(a) Personality Data

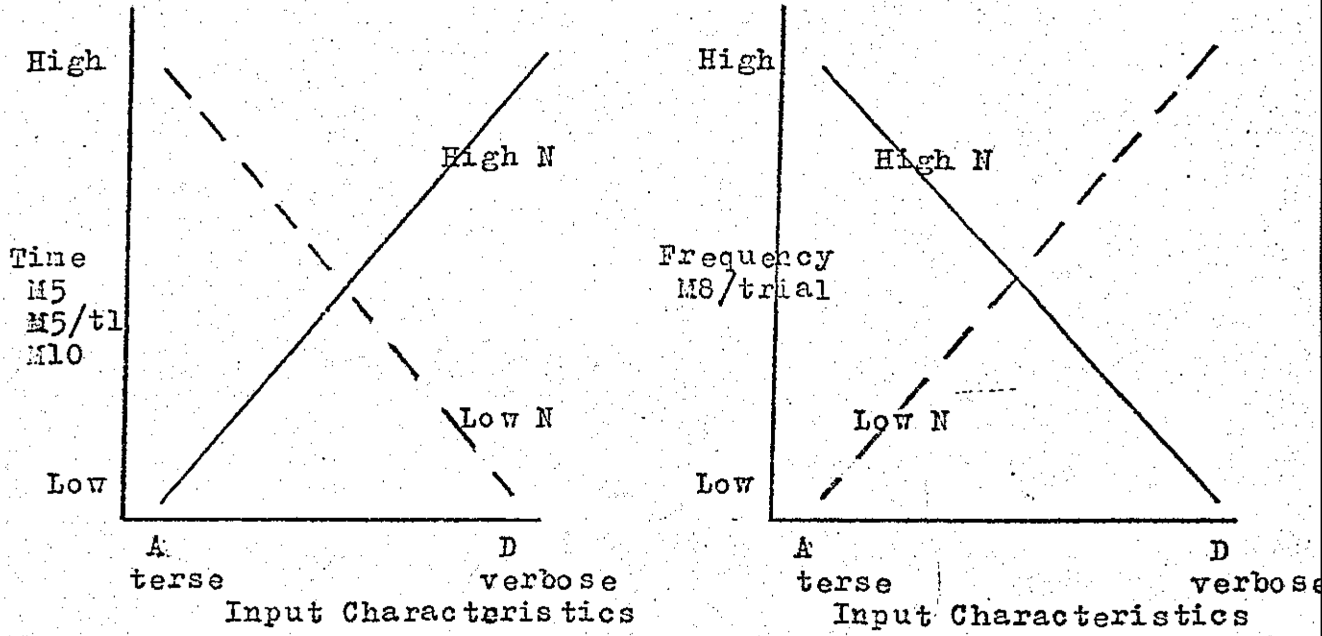
The experiment was not designed to be sensitive to personality factors and it was not surprising that none of the measures were significant at the 5% level. However, neuroticism appeared to be a likely factor to be considered for future research, particularly for the measures M5, M10, M5/Trial, M8/Trial. These were the measures concerned with planning, evaluation and the number of steps between each plan. Graph 2 shows the results for the measures mentioned in an approximate form.

Since these graphs are based on data which could have occurred by chance, it would be dangerous to make many inferences from them.

9.3 Acceptability - Preference and Performance

There are two sources of clues as to why the rank of the input methods (I) (in terms of preference for maze learning) should be B, A, C, D. First, the comments from each subject and second the only performance measures whose ranking with I agreed was M8 (MMAA).

All subjects commented during the trials about the sensitivity of the joystick and the need for maintaining its position with one hand. After the trials, S14 also commented on control-display incompatibility, because, as a trained pilot, the display rather than position was expected to be controlled. While the sensitivity of the joystick was an asset for well-



Key: N means Neuroticism
High means score above group median
Low " " below " "

Graph 2

Interactions between Input and Human Characteristics trained subjects, it was found to be too easy to make errors (by overshooting movements) by the experimental subjects who had one practice run through the spiral Maze 5. Typically, practice runs with Maze 5 using inputs B, C, D needed 1-3 trials before exiting the maze while with A (the joystick) the range was 3-10 trials. Hence, the lower ranking of A could be partially because of the ease of making errors by overshooting, i.e. purely a judgement on the sub-optimum control/display ratio. The error rates were B-2.67%, C-0.74%, D-1.09% and A-3.24%. These figures support the comments of subjects.

In the light of this explanation, the reason why M8 ranking agrees becomes more obvious; the joystick allowed easy movement within the maze but led to more overshoots between rooms of a floor - thus the time cost per move was low but the risk of error was high. With the first ranked input (B), the time cost was slightly higher but the risk of error was much lower than for the joystick. A similar argument was considered for the other conditions which led to the inference that subjects judged the preference of an input method to solve a maze by two combined criteria: the time cost per move and the risk of error.

9.4 Problem Solving Behaviour

Subject comments revealed a number of important aspects of the way in which subjects solved the problem. These were: the development of strategies; the special difficulties associated with the learning of the task; and the use and effects of particular input methods. The comments also indicated the degree of acceptability of the different input methods for learning the way through the maze.

Strategies were placed into one of three categories: first, the goal oriented minimum-difference strategy (e.g. subject number 4); second, trial-and-error (e.g. subject number 15); finally, a mixed strategy consisting of strategies 1 and 2 (e.g. subject number 11). The training tended to favour a goal-oriented strategy and it was surprising to find that some subjects (S5, S15) did not develop this successful strategy but relied on a simple trial-and-error basis. Most subjects (13/16) realised the strength of learning to define sub-goals so that associative 'chunks' of information were all that had to be recalled. The variations within this strategy were few and represent differences between remembering absolute (S4) or relative sequences (S11) with respect to the entrance. The odd-man-out in terms of strategies was S6 who developed a serial method of remembering the sequences rather than relying on visual pattern recognition and sub-goals. Surprisingly, this subject completed all the experimental tasks.

In conjunction with the reduction of memory load in the task by the development of coding systems, many subjects made use of internal visual models. Some of these models were said to be vivid and three dimensional (S2, S12, S14, S16) while others were two dimensional and referred to the plan views only (S1, S3, S4, S7, S8, S9, S12, S13). Since training encouraged the development of a 3-D model of a building, it was surprising that so many subjects did not do that. All of those that used a 3-D model (4/16) developed goal-oriented strategies. With the plan model, 2/7 did not develop goal-oriented strategies; of the remainder, 4/5 did not develop a goal-oriented strategy.

Difficulties commented on were in three distinct areas; task conceptualisation (S8, S10, S13), maze display of staircases (S2, S7), and input device compatibility (S14). The last of these arose because S14 was a trained pilot. The abstract nature of the up-down-up sequence was responsible for all the difficulties of task conceptualisation; this was expected, since it was the essential consequence of using an established concept and extending it into a new dimension. Training in this respect was successful for most (13/16) subjects.

The problem with the display of staircases was more difficult to explain since, in a binary situation (i.e. staircase or no staircase) it was expected that the display of crosses for solid floors into the display was the more logical choice. This seemed to be a secondary idea to the subjects in relation to the idea of presence or absence of a 'staircase' indicated by the cross, even though the former idea was presented first in training. One subject (S2) resolved the different strengths of the ideas by using the concept of 'lifts' rather than staircases. The reason for this was not made clear. Clearly, the ambiguous nature of the presentation of inter-floor passages was responsible for problems.

The problems of ambiguity in the problem solving situation were further emphasised by the comments on the use of the 'mark' facility. The convention for its use was open to the subject so that the 'staircase' problem (in that it was arbitrarily assigned as far as the subjects were concerned) could be avoided by using the mark. Of those subjects that used it, (S1, S4, S9) all commented on the difficulty of remembering which of the assigned meanings the mark had. This seems to indicate the overloading of the short term memory.

10. Conclusions

Within the constraints of this investigation the following conclusions are made.

10.1 Within Problem Solving

10.1.1 Information Rates

The time taken for a subject to make a step in interactive problem solving (URT or User Response Time) was distributed such that there were 3 peaks. By observing subjects' behaviour and analysing verbal protocols, it was apparent that the first peak was associated with previously decided solution steps; the second peak was associated with evaluation of progress towards the solution and the third with planning and remembering the general form of the solution. The following conclusions are based on this interpretation of the data.

As input methods varied from A to D, the mean number of steps in an attempted solution decreased but the time needed to input pre-programmed steps increased as well as the time spent during the solution evaluating stages of the solution and recalling appropriate information.

Differences between input methods were not reflected in measures of problem solving such as the overall time to solve, the planning time or the number of evaluations or steps in that solution. These measures were sensitive to problem difficulty.

10.1.2 Acceptability

The order of preference for the input devices was cursor input, joystick, alpha key and verbose input. The order did not correspond to the speed of input and the conclusion was reached that the error proneness of a device affected the preference ranking such that, if devices were judged comparably fast, then error proneness determined the rank.

10.1.3 Subsidiary Variables

The subsidiary variable was the personality trait of the subjects and it was concluded that the trait of neuroticism may interact with the input method. The interaction was such that for input methods requiring long input times for pre-learned or obvious sequences, neurotics made more plans, less moves per plan and took longer to evaluate than stable people. With input methods allowing fast but error prone input, this was reversed, i.e. neurotics make less plans with more moves per plan and spend less time in the evaluation of the solution than stable people. This conclusion is tentative because the results upon which it is based could have occurred by chance with a probability of about 0.11.

10.2 Non-Problem Solving

This refers to the situation where the solution was obvious and no planning or attempts to build a solution were necessary.

10.2.1 Work Production

The URT's were distributed with one peak and so were directly comparable. The joystick and cursor inputs were not significantly different but both were significantly faster than the alpha key input. The alpha key input was significantly faster than the verbose input. The differences between all the input methods were less in the task than in the problem solving task.

The error rates were different for each device; joystick (3.24%), cursor (2.67%), alpha key (0.25%) and verbose input (1.8%). The percentage of mis-typed commands (both entered and cancelled) was 11%. (This was for the verbose input only.)

10.3 Experimental Hypothesis 3

The hypotheses that (i) problem solving does not depend on the input method, and (ii) that problem solving does not depend on the way of use (coding) of an input device, must be rejected in the conditions in which they were tested. That is, the processes of problem solving are affected by subjects adopting strategies which cope with characteristic performance of input devices. With measures of overall problem solving performance, the hypothesis cannot be rejected. The system conditions were optimal; the distal pragmatic feedback being fast, and the system reliable.

10.4 Implications for the Model of Interactive Problem Solving

There was support for the idea of an evaluation process used during interaction for balancing central cognitive and peripheral processes. The relationship was such that the greater the time needed for accurate input of information, the less the number of steps per trial solution, and the greater the time to evaluate information. The basis of the relationship may be in the capacity of the short term memory for coping with solution and input information.

CHAPTER 6

EXPERIMENT 4

SUMMARY

This chapter describes the testing of the hypothesis that interactive problem solving does not depend on the interactions between the input method and the characteristics of systems and people. It was tested by using a balanced 3-factor experiment. Two sets of factors were used. Group one included the variables: (i) input method and distal pragmatic feedback, (ii) system and task, (iii) order of presentation. Group two included: (i) subject experience, (ii) subject personality and (iii) order of presentation. Twenty-four subjects completed the experiment.

Information input times were measured and a 3-peak distribution was found for input of the first part of a source message. The input of the remainder of the message had a single peak distribution. Unlike the 3-peak distribution, its characteristics were affected by the second group of factors but only for particular input methods which are error prone. The characteristics of the 3-peak distribution were affected by the first group of factors. The interactions between the factors are complex and no clear pattern can be seen, partially because the factors are groups of variables. However, the results are tentatively explained and these support the model of interactive problem solving described in Chapter 2. The hypothesis that problem solving does not depend on the interactions between the input method and the characteristics of systems and people is rejected. The relative importance of the input sub-system is indicated in the results of this experiment. A complementary experiment is described in the next Chapter which examines the importance in greater detail.

CONTENTS

TABLES

FIGURES

PLATES

GRAPHS

No. Title Page No.

1. Introduction

This Chapter describes testing the experimental hypothesis that problem solving does not depend on the relationships between the characteristics of people and systems and the method of input.

The previous Chapter described an experiment to test one aspect of this hypothesis and it was rejected within the conditions of that experiment. This Chapter describes an experimental test in different conditions which are intended to be nearer to those existing in 'field' experiments.

The conditions are so contrived that the major aspects of real time computer-aided problem solving are represented. These are system reliability and the characteristics of distal feedback at all levels (syntactic, semantic, pragmatic).

Previous experimental work in the thesis has shown computer experience and the personality traits of neuroticism may affect the use of the input sub-system in interactive problem-solving. The hypothesis is tested with these two main factors.

The investigation of this Chapter is largely exploratory in the sense that the overall purpose is to investigate a very complex situation rather than the simple well behaved conditions of the previous experiments.

2. Objectives

These are based on the experimental hypothesis described in the previous section. Specifically, they are to:-

(i) compare the use and acceptability of two alternative input methods when combined with two systems used for problem solving.

(ii) examine the effects of individual differences on the results of the comparison in (i).

3. Constraints

These were as for the menu selection experiments in Chapters 3 and 4 (described fully in Chapter 3). The most pertinent constraint is that the two available input devices were light pen and keyboard.

4. Experimental Details

4.1 Design

~~The de~~sign chosen for examining the hypotheses is considered in three parts; that concerned with (i) system characteristics and input method, (ii) subject characteristics within (i), and (iii) subject characteristics between (i).

(a) System Characteristics and Input Method

A 3-factor factorial design was chosen for the experiment. The factors were groups of variables combined in two ways. The first factor called 'inputs' (I) included the input method and the distal pragmatic feedback characteristics. The second factor called 'systems' (S) included the variable of system reliability, system response times and task difficulty. The third factor called 'order' (O) gives the order of presentation of each of the experimental conditions to each subject.

Two levels of each factor were used in the design which was balanced over all subjects, but not for order effects within factor one. This was done because, unlike in the previous experiment, the problem did not have flexibility in difficulty levels. That is, each problem had a unique difficulty level.

The factors and design are shown in Table 1. Subjects were allocated to conditions in a balanced way depending on their characteristics of personality and experience.

The levels of the factors were as follows:-

I -1 (light pen with distal feedback - called 'LP'
   2 (keyboard without distal feedback - called 'KYB'

S -1 (reliable system with fast, consistent response times - (called 'OPTIMUM'
   2 (unreliable system with slow, inconsistent response times - (called 'SUB-OPTIMUM'

O -1 (LP conditions presented first to the subject
   2 (KYB conditions presented first to the subject

(b) Subject Variables within System Variables

The two main subject variables under consideration are (i) the personality trait of neuroticism and (ii) experience of computers and systems. Experience may be further considered as specific and general experience. Specific experience refers to the experience gained on a particular system, which in this case is the system used in the experiment. In the design

Independant Variables	Input Characteristics	Light Pen				Keyboard			
	System and Task	Optimum and Easy		Sub-Optimum and Difficult		Optimum and Easy		Sub-Optimum and difficult	
	Order of Presentation	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Experimental Conditions	Set 1	A1	-	A2	-	-	B1	-	B2
	Set 2	-	B1	-	B2	A1	-	A2	-

Key: - means not used in the experiment
A's and B's were paired together. The order within pairs is given by the 1 (1st) or 2 (2nd)

Table 1.

Experimental Design
System Variables

there are two levels of the main system variable and each subject is required to solve the problem using both levels. The specific experience gained on the first trial (with level 1) may be useful in trial 2 (level 2). This may be termed an order effect and attributed to an order (0) factor described in the previous section.

The design chosen was a 3-factor ANOVA shown in Table 2.

Personality Trait	Neuroticism level 1				Neuroticism level 2			
General Experience	E1		E2		E1		E2	
Specific Experience (Order of Presentation)	01	02	01	02	01	02	01	02
Design Cell Number	1	2	3	4	5	6	7	8

Key: E = Experience

O = Order

Table 2: Experimental Design - Subject Variables

Each subject was allocated to a particular cell according to (i) general experience, (ii) personality, and (iii) order such that equal numbers of subjects are in each cell (i.e. balanced).

(c) Subject Variables between System Variables

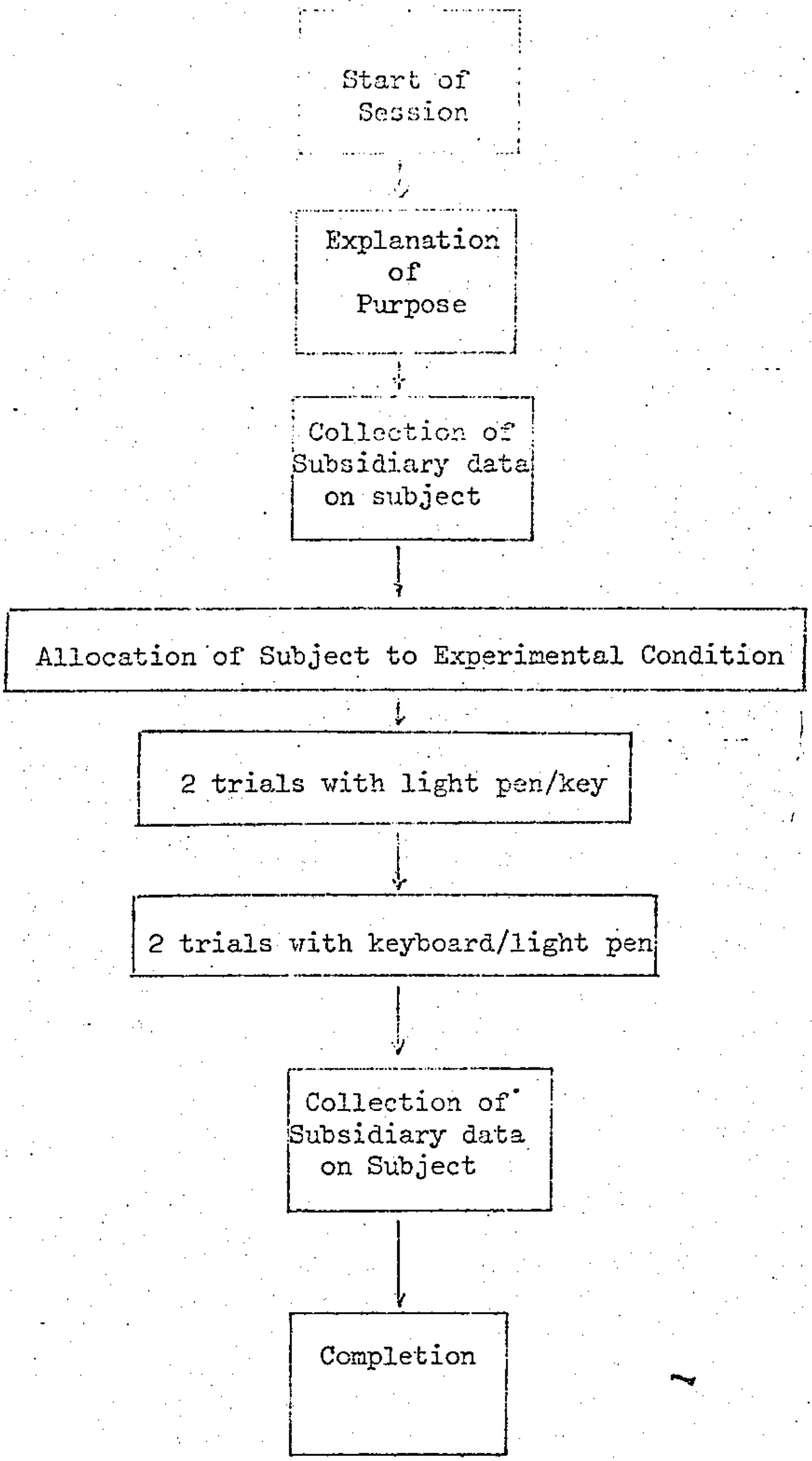
Provided that certain conditions are met concerning the balance of experimental design, the effects of subject variables may be examined between system factors using the factorial analysis technique. The balance of the design depends on the allocation of subjects to cells as they are appropriate and an equal number of subjects per cell. How well these may be met depends on the constraints of the experiment, and, in particular, the availability of subjects. These are discussed in section 4.5.

4.2 Sessions

The procedure for each subject in the main experimental session was as shown in Figure 1 and was based on a small pilot experiment using 2 subjects.

The subject was told that the purpose of the experiment was 'to investigate the relative merits of light pen and keyboard for solving a simple problem'. He was also informed that there was no time limit and that he may withdraw at any time for any reason which he was not required to give.

The details of the procedure within each experimental condition are shown in Figure 2.



Procedure of a Subject's Main Session
Figure 1

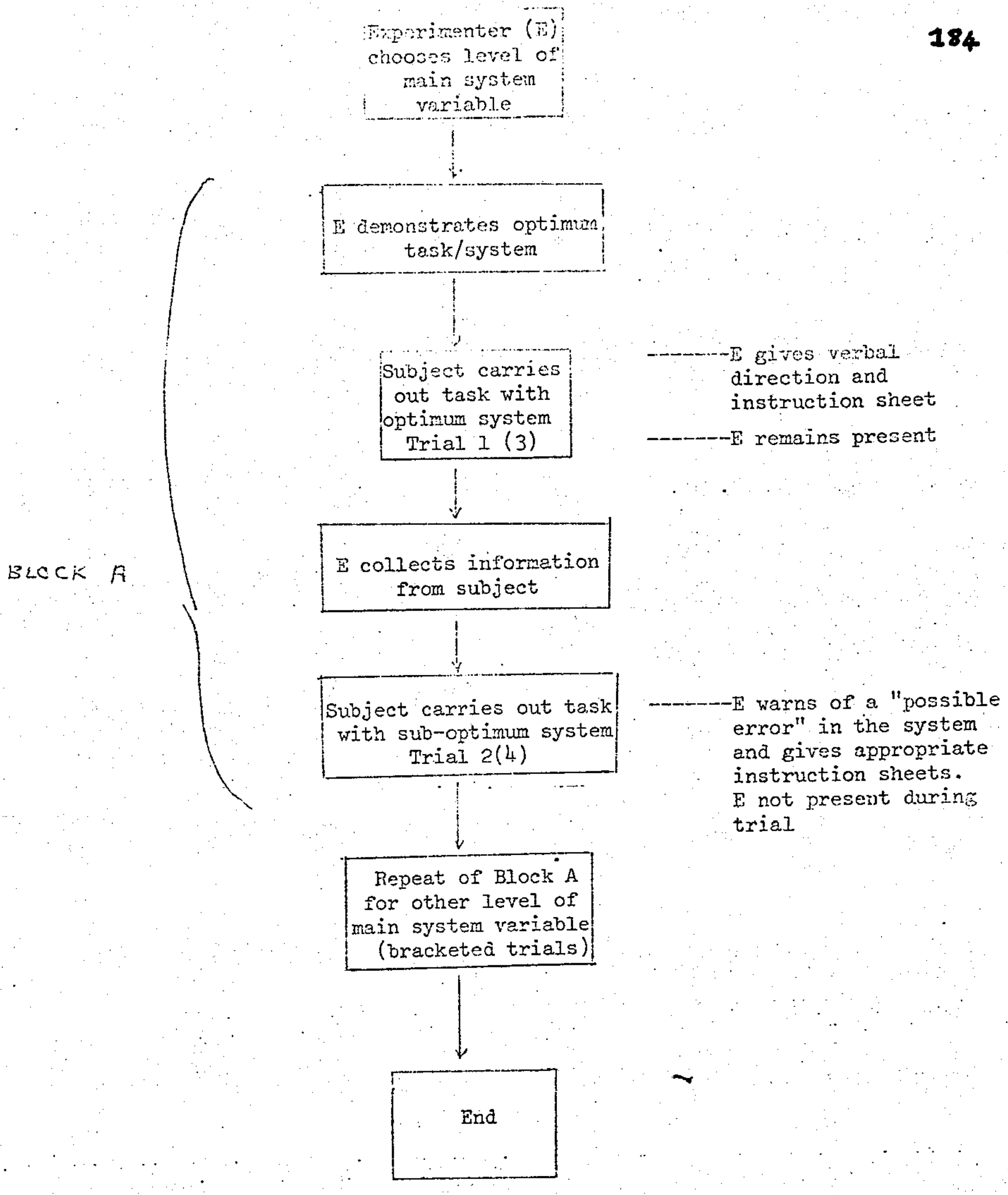


Figure 2

Subject Procedure within Experimental Condition

The subjects were unpaced in that they had as much time as necessary to solve the problem. Verbal directions were "Please complete the task if you can. There is no time limit. If, for any reason, you feel you cannot complete the problem, please quit and inform the experimenter. Thank You".

4.3 Task

4.3.1 Choice and Characteristics

The choice of a suitable task for this investigation was based on the following criteria:

(i) A graphic representation is desirable for reliable comparison with the results of the previous experiment.
(ii) The number of steps to a solution be controllable.
(iii) All steps should be of equal difficulty (homogeneity).

In a review of complex tasks for problem-solving research, Ray (1955) suggests that tasks may be divided into three parts (a) the given situation, (b) the desired situation, and (c) the method of proceeding from one to the other. A task proposed by Ray (op cit) is called the 'Disc Transfer' task and is alternatively known as the Towers of Hanoi problem. The basic materials are three identical pegs and a number of discs, each of a different diameter and with a central hole. The given situation is that all the discs are placed on one peg (the left most) in descending order of size. The desired situation is all the discs on the right most peg in the same order. The rules for proceeding from one to the other are that only one disc may be moved at a time from one peg to another and a larger disc must never be moved onto a smaller one.

The characteristics of the task are that the number of steps to a solution (degree of difficulty) are (2^n-1) where n is the number of discs. Wicklegren (1972) has shown how the solution may be broken down into the solution of (n-2) sub-problems each being a sub-goal of the overall problem.

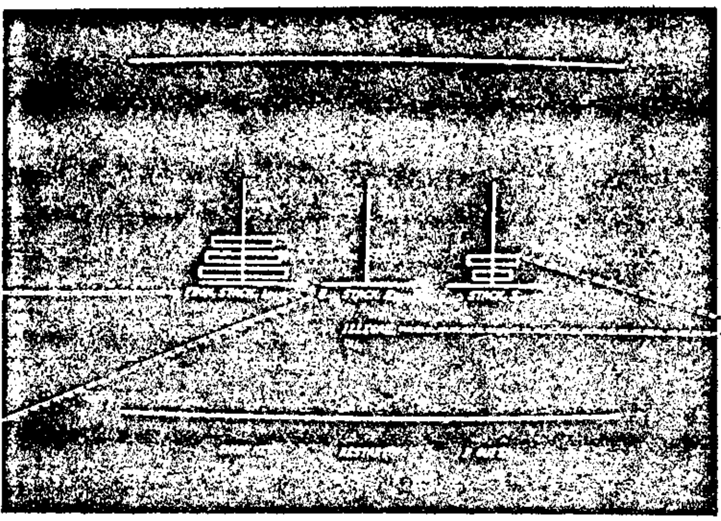
This may be understood by considering the solution of the 4-disc problem as follows. The overall objective is to place all the 4 discs on the right-most peg. The rules state that this must be achieved by moving one disc at a time. Therefore, the first objective must be to put the largest disc onto peg 3. This may only be achieved by removing all the discs above it and putting them onto peg 2. This is the first sub-goal. The next objective is to place the 2nd largest disc onto peg 3, which means moving the two discs above it onto peg 1. This is the second sub-goal. Having reached the position above, the 2 largest discs are correctly positioned on peg 3, and the two smallest are on peg 1 and peg 2 is empty. The solution is now trivial, i.e. for n=2, there are no sub-goals. Therefore, there are two sub-goals for the n=4 problem.

Each sub-goal has the same basic form of solution and therefore the general solution consists of a repeated hierarchical (recursive) use of the sub-goal method (Wicklegren (1972)).

4.3.2 Description of Computer-based Task

The usual form of the task is in blocks of wood or similar material. During a solution, a person may deliberately discard his previous efforts by placing all the discs back in the starting position (called 'restarting'). Alternatively, he may decide to give up altogether ('quit'). Thus any computer-based version must allow these functions as well as provide a means of symbolising the discs and pegs of the task and a way of moving them. Because of the nature of the solution it was also possible to program the computer to solve the problem itself and thus aid the problem solver by showing the solution. There was an additional function which was not usually possible in the wooden peg task which may be called 'policing'. If the disc rule was broken the computer informed the problem solver that the move was illegal.

The computer based task used in this experiment symbolised the discs and pegs by means of rectangles and vectors respectively. These were displayed graphically together with the options available as shown in Plates 1 and 2. Each peg was identified by a label 'stack' number 1, 2 or 3, so that the topmost disc on it could be selected by



The disc on top of Stack 1. has been selected

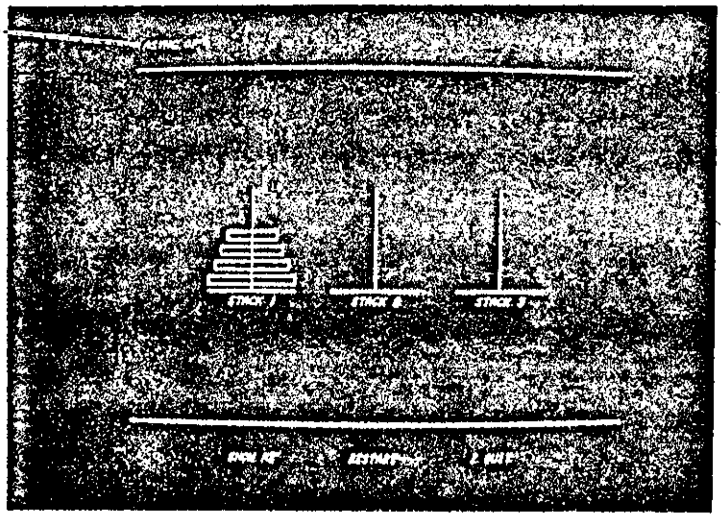
Confirm selection block

Placing a large block onto a smaller one breaks the game rules

LIGHT PEN HANOI TOWERS

PLATE 1

Blinking message indicating 'Show me' has been selected



LIGHT PEN AND KEYBOARD DISPLAY OF HANOI TOWERS
WHEN USER HAS ASKED TO BE SHOWN THE SOLUTION

PLATE 2

identifying the stack. The means of moving discs and selecting options to quit , restart or ask for computer aid were different for each input device.

4.4 Subject Information

The subjects were allowed the use of the information sheets shown in Figures 3 and 4. These could be placed where required for reference.

4.5 Subject Population and Context

All subjects were mature adults who had been well educated. These details are summarised in Table 3. The number, type and time of availability of the subjects in the experimental situation limited the number to 24.

5. Independent Variables and Measures

5.1 Input Method

5.1.1 Input Device

Two devices were used; the light pen of the GT42 and the QWERTY keyboard of an ASR 33 teletype.

5.1.2 Way of Use

(a) Light Pen

The method of selecting an option and moving a disc was to point at the option or stack label whereupon it blinked. A 'block' then appeared to the left of the label (see Plate 1) which, if pointed at, confirmed

Information for using Computer to solve the Towers of Hanoi Problem by Keyboard Command

When the symbol ">" is printed the computer is ready to receive instructions or cammands from the keyboard. Commands are single letters which are typed and entered by typing the return key. Mistyped letters which are not valid commands are ignored. The valid commands, their function and the action you should carry out after issuing the command are listed below:

Command (Letter)	Meaning	Computer Response	Your Action
M	Move a Disc	1. Types "OFF:" and waits. 2. Types "ON:" and waits. 3. If a legal move, a space is typed. If illegal, "ILLEGAL" is typed. In both cases move is done.	Type the peg number of where you want to take the disc off. Type the peg number of where you want to put the disc on.
D	Draw the current stacks	Peg display is updated for every disc.	None
R	Restart	Discs are all placed back in order on the leftmost peg and the display updated accordingly	None
S	Show me how it is done	Computer displays the solution by moving one disc at a time slowly. When finished you are restarted. (When there are 4 or more discs, the solution will not show for some time. You will see a message "ASYNC OPS" to remind you that the computer will interrupt you with the solution - until then you should proceed to try and solve it yourself).	None
Q	Quit	Display vanishes and an appropriate message is printed.	Please inform me.

If you succeed in solving the problem, please inform me.

Note: When there are 4 or more discs, there are
(i) Program errors and (ii) system delays.

PLEASE DO NOT TYPE A COMMAND OR STACK NUMBER UNLESS THE APPROPRIATE PROMPT HAS BEEN TYPED

FIGURE 3
KEYBOARD INFORMATION SHEET

FIGURE 4

LIGHT PEN INFORMATION SHEET

Subject Information

Your task is to transfer all the discs from the leftmost peg to the rightmost peg in the shortest time possible.

You may move only one disc at a time and it can only be moved onto a larger disc or an empty peg.

The procedure to move a disc is as follows:

(1) Point the light pen at the label of the peg where you want to move the disc from. A solid block will appear to the left indicating the computer has recognised your selection.

(2) Confirm your selection (or select another peg) by hitting the block with the light pen. The word "from" will appear next to the peg label you have selected.

(3) Point at the label of the peg where you want to move the disc to. A solid block will appear indicating the computer has recognised your selection.

(4) Confirm your selection (or select another stack) by hitting the block with the light pen. The block will disappear and, after a few moments, the topmost disc on the peg first picked will be moved to the topmost free position of the second peg selected. This completes the disc moving procedure.

Note that the computer will carry out a move even though you have asked for a move of a disc onto a smaller disc. In that event, the message 'illegal' will flash on the screen.

If you are experiencing difficulty in solving the problem you have three options which may be selected by pointing and confirming with the light pen. These are as follows:-

(1) 'Show me' On selecting this option, the computer will automatically and slowly show you the solution by doing the appropriate moves on the screen. On completion, it puts all the discs back on the leftmost peg. i.e. in the starting position. If there are more than four discs, there may be a delay from the time you request to be shown the solution to the time the computer presents it to you. In that time you may continue to try and solve the problem and the fact that the computer is working for you at the same time is indicated by a blinking message which says 'async op' which stands for 'asynchronous operations'. That message will disappear and the computer will show the solution as previously described after a short time.

(2) 'Restart' will place all the discs on the leftmost peg no matter how they appear when the option was selected.

(3) 'I Quit' is an escape route if you feel you cannot solve the problem. Selecting this terminates the experiment.

N.B. When there are five or more discs there is a possible program error whereby the 'message' 'illegal' is blinked onto the screen. Also, the computer may respond quite slowly because it is working harder.

Figure 4 (continued)

Subject Number	Sex (M=Male) (F=Female)	Age decade	Education Level*	Project Member+ (P)
1	F	20	A	P
2	M	30	U	N
3	F	30	U	N
4	M	20	U	P
5	F	20	U	P
6	M	40	A	P
7	F	20	U	N
8	M	30	U	P
9	M	20	U	P
10	M	20	U	P
11	M	40	A	P
12	M	20	U	P
13	M	20	U	N
14	F	20	A	N
15	M	20	U	N
16	M	30	A	P
17	F	20	U	P
18	M	30	U	P
19	G	20	U	P
20	G	20	O	P
21	M	20	U	P
22	M	30	U	P
23	M	20	U	P
24	M	30	U	P

* U - up to University level
A - up to A-level
O - up to O-level

\+ P - member of the project (SSDS) group; (implies CHD sympathies)
N - not a member of SSDS project.

Table 3

Details of Subject Population and Context that selection. Alternatively, the light pen may be pointed at any other option which then blinked and caused any other option to stop blinking, and the block appeared to the left of the new selection. This procedure was designed to offset the problem of using an 'unswitched' light pen.

On confirmation of a stack label, the word 'from' appeared and replaced the block. This signified that the topmost disc on that stack has been selected for moving. The problem solver now selects and confirms another option or stack. If a stack, the previously selected disc is moved to the topmost free position on it. If the option to quit or restart is selected and confirmed at any time, it is executed immediately. If 'show me' is selected and confirmed the computer shows the solution immediately or at some later time. If at some later time, the message 'ASYNCH OPS' was displayed from the time of request until the solution was present. No other action was taken by the system and the problem solver could proceed. On showing the solution, the discs were moved, one move at a time, each move being displayed for 2 seconds. After completion, all discs were placed on the left-most peg.

(b) Keyboard

The light pen version of the task was typical of interactive graphics systems in that the display file was actively updated during the interaction by the system and hence the disc moved on the screen (i.e. distal pragmatic feedback).

In many systems with teletype interaction and a graphic display, the system does not automatically update the display file during some interactions and the problem solver needs a command to enable this. In the keyboard version this command was 'DRAW'. All other options were similar to the light pen version.

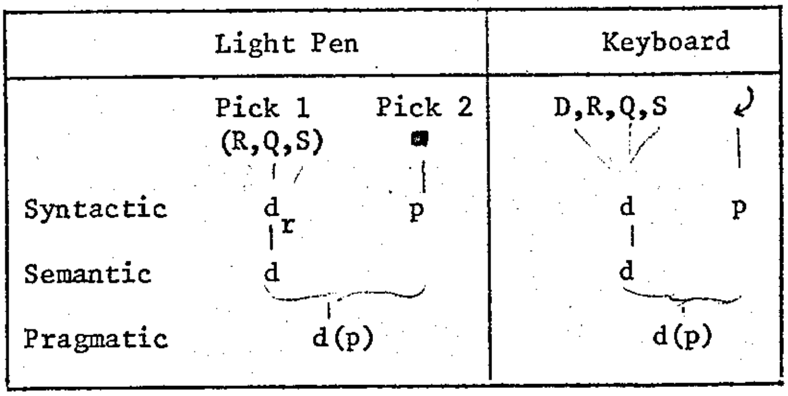
The form of interaction was that the teletype printed a prompt ('>') when ready for input and any command letter could be typed. If a mistake was made, the 'delete' key erased that character and another could be entered. Action was taken by the system when a command was confirmed by typing the carriage return key. The action depended on the command given. The command 'M' was used to move a disc, and the system typed 'OFF'. This signified a computer request for the stack number of the disc required to be moved. Again, a mis-typed stack number could be deleted and the system waited for a confirmed number. On receipt of such a number, the word 'ON' was typed by the system. The number of the stack where the disc was to be moved to was then typed and confirmed by the subject. The system then recorded that such a move had been made (but did not update the display) and typed the prompt for the next command. On receiving the draw command (D), each of the stored moves was carried out and the display was updated.

(c) Formal Description

The input methods may be more formally described as follows:-

Light Pen	Source Primitives	Move a Disc from Stack x to Stack y			
	Receiver Primitives	Pick 1	Pick 2	Pick 3	Pick 4
	Syntactic	d_r	p	d_r	p
	Semantic	d		d	
	Pragmatic	d		d(p)	
Keyboard Input	Source Primitives	Move a Disc from Stack x to Stack y			
	Receiver	'M'	↵	x ↵	y ↵
	Syntactic	d_r	p	d_{rp}	d_{rp}
	Semantic	d		d	d
	Pragmatic	d		d	d(p)

For the other source primitives, (draw, restart, show me, quit) the analysis is as follows:-



KEY: d = descriptive information (multiple inputs subscripted r).

p = Prescriptive information

■ or ↲ = end of input

5.2 Task Variable - Difficulty

A small pilot experiment was carried out using 3 subjects (who were thereby eliminated for consideration in the main experiment). Each was required to solve the problem using 3 and 5 discs (subject 1); 4 and 6 discs (subject 2) and 5 and 7 discs (subject 3). None of the subjects learned a general strategy of solution in the two trials and a retest of subject 1 showed that the solution had not been memorised to the extent that performance measured by overall time of solution had improved. The pilot experiment showed that the 3 discs problem was too trivial, and a 6 or 7 disc problem needed too many steps in its solution for motivation to be maintained. This meant that the 'easy' task was chosen to be

4 discs (requiring 15 steps and 2 sub-goals to solution) and the 'difficult' task would be 5 discs (needing 31 steps and 3 sub-goals).

5.3 Computer Task and System Variables

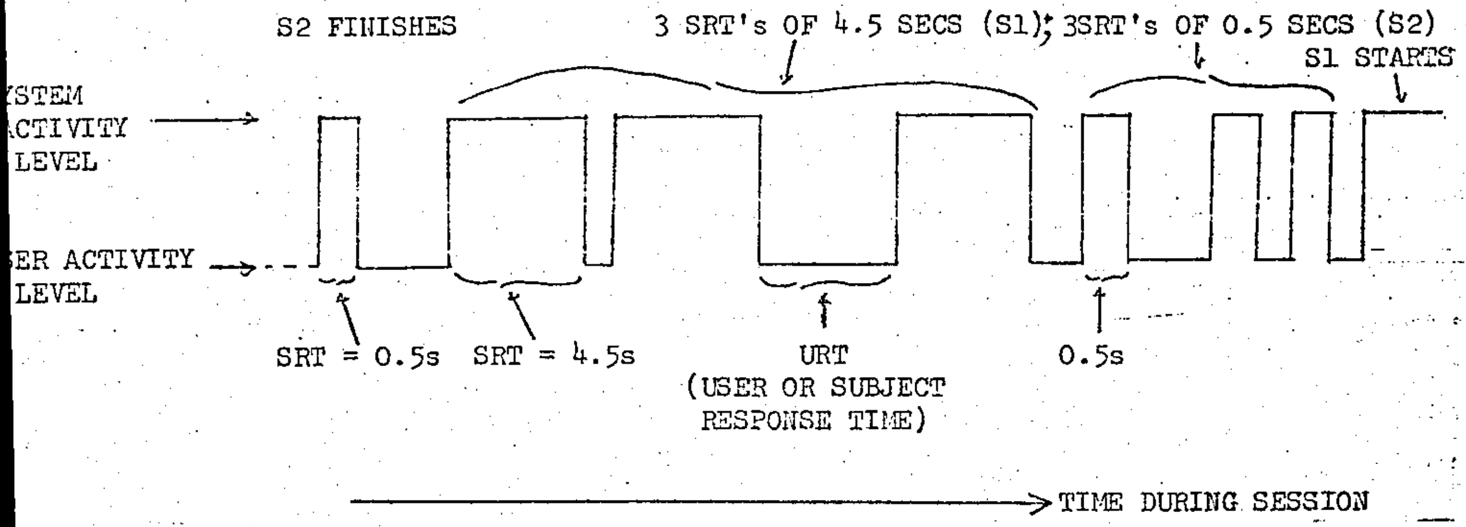
The two variables considered are the response time and reliability of the system. These are defined as follows:-

(a) System Response Time (SRT)

SRT is defined here from the subjects point of view, as the 'dead' time in which the system does not have the ability to be aware of (and therefore respond to) his actions. Every action of the subject of the input device had an associated SRT. Without any control of the SRT, the mean and standard deviation (SD) of the SRT for the 'optimum' of the experimental task was: mean 0.2 seconds; SD <.05 seconds. The uncontrolled SRT was therefore used as the fast and consistent condition of the experimental design.

The slow and inconsistent condition was generated by controlling the SRT using a continuously running clock in the software. There were few constraints on the control. Therefore, the choice was based on the following assumptions:-

(i) In conversation, people expect a response within 4 seconds (Miller (1969)). Any time outside this will be called 'slow' in relation to conversational rates.



GURE *6*

STEM RESPONSE TIME (SRT) SEQUENCE

Subject	Eysenck Personality Trait Score			Personality Group *	Experience Level+	Experience Group+
	I-Score (0.24)	N-Score (0.24)	Lies (0.9)			
1	6	21	2	3	3	1
2	2	14	3	3	5	2
3	12	15	0	3	2	1
4	13	15	0	2	1	1
5	19	6	5	4	5	2
6	16	16	2	2	1	1
7	13	8	3	4	1	1
8	10	8	2	1	3	2
9	13	9	5	4	3	1
10	7	10	1	1	4	2
11	13	10	2	4	1	1
12	10	5	2	1	3	1
13	18	15	0	2	4	2
14	11	7	4	1	1	1
15	12	10	2	3	4	2
16	16	18	0	2	2	1
17	14	18	4	2	3	2
18	7	12	3	3	3	2
19	5	13	4	3	4	2
20	17	14	2	2	3	1
211	13	7	3	4	4	2
22	9	2	2	1	5	2
22	9	2	2	1	5	2
23	13	9	2	4	3	1
24	8	4	1	1	4	2

Key: I - introversion/extroversion; N - neuroticism/stability dimension

* Refer to Table 3

\+ See Text and Table 3

TABLE 4

(ii) Experience of 3 successive similar events leads to a high expectation that the fourth will be similar. If it is not, then there is a high possibility that events will be construed as being 'inconsistent'.

Figure 5 shows the SRT sequence chosen for the sub-optimum system. Provided the number of such sequences is large (>20) the mean SRT is 2.5 seconds with an SD of 2 seconds.

Exceptions to this sequence were subject inputs not concerned with moving discs, such as requesting a restart (immediate response), requesting the computer to show the solution (delayed by a variable amount (TF)), and indicating that the subject wished to stop (immediate response). TF was zero in the optimum system and was given a value in the sub-optimum system such that a solution test would continue in the time after the request.

If the estimated total time of solution for the subject and N-disc task was TL, then TF = TL/(N-2) for (N-2) sub-goals. The method used for calculating TL is given in Appendix 2.1.

(b) System Reliability

The sub-optimum system was made unreliable in two ways; first, the subject was erroniously informed by the system that he had broken the game rules and secondly, the system would inform of a 'crash' and force a restart. The first case occurred wherever the system noted that a disc was being moved onto an empty peg. In an ideal solution of an N disc problem this occurred $(2^{N-2}+2)$ times. Crashes occurred after the estimated time of solution for that subject and N discs as given by TL (see Appendix 2.1). The optimum system was reliable in that these two conditions did not arise. The number and time of crashes were logged by the system.

5.4 Subject Variables and Measures

The subject variables were personality trait and experience. The personality trait in the design was neuroticism (stability) (trait N). This was measured using an Eysenck personality Inventory (EPI) which consists of 57 binary choice questions. The results were 3 scores; I-score (0 (introverted) - 24 (extroverted)); N-score, (0 (stable) - 24 (neurotic)) and a lie frequency (0-9). The EPI was implemented on an interactive computer system and the experimenter was not present when the subject completed it. After collection of personality trait data, the subjects were assigned to groups according to whether their score was above or below the group median for the neuroticism trait.

Subject experience was measured by means of a semi-structured interview using questions related to the use, context of use and period of use of computers by the subject. On this basis the subject was assigned to one of 5 categories of experience. These were: 1 - no experience; 2 - data input/output only; 3 - programs for self only as necessary; 4 - programs for others as required; 5 - systems programmer. For the purpose of the experimental design, group 1 included categories 1 and 2 and group 2 included the rest.

Table 4 shows the subject data and experimental group allocations.

6. Dependent Variables and Measures

6.1 Work Production (Problem Solving Performance)

The system collected data on the number and time for user entry of pragmatic information according to whether it was a disc move, request to restart, request to be shown the solution or a request to stop. Since the way of use is different for each device, the detailed procedure of using each must be considered in describing the measures that were used for the basis of comparison. The approach in deciding the measures of information flow is to assume that the first receiver primitives corresponding to a new source message have a different distribution from those which follow. The approach is based on inferences made from the previous experiments. Namely, that (i) the three peak distribution can be expected when conditions are similar to those of the previous experiment, and (ii) the single peak distribution may be expected when conditions are similar to those of the first two experiments.

(a) Light Pen Data

In order to move a disc, the light pen was pointed at the label under the peg holding the disc (action 1); a block appeared next to the label and the label blinked (action 2); the light pen was pointed at this block to confirm that the right peg has been selected (action 3); the word 'FROM' replaced the block and the label stopped blinking (action 4); the subject then pointed at the label of the disc destination peg (action 5); the label blinked and a block appeared next to it (action 6); action 3 was repeated (action 7) and the disc moved. The elapsed time between the system being able to respond to the light pen and action 3 is called the first subject-response-time (R1). The time between the system being able to respond after action 4 and action 7 is the second subject-response-time (R2).

(b) Keyboard Data

In order to move a disc, the command 'M' was typed (action 1'); action 1' was confirmed by typing the carriage return key (action 2'); the system printed 'OFF' the subject then typed the number of the peg with the disc on (action 3'); action 4' was confirmed by typing carriage return (action 5') - the system typed 'ON:'; the subject typed the number of the destination peg (action 6') and repeated action 4' (action 7'). In this case, R1' (the subject's response time) is the elapsed time between the system prompt being typed ('>' - indicating ready for input) and action 2'; R2' is the elapsed time between the system printing 'ON:' or 'OFF:' and action 5' or 7'.

In a small pilot experiment, R1, R2, and R1', R2' were typically distributed as shown in Graph 1. R2 is a single positively skewed distribution while R1 is a distribution with 3 peaks, only the first of which is positively skewed. These data were typical of that collected by the optimum system for all subjects and formed the basis of calculating the boundaries B1 and B2 for the sub-optimum system described in Appendix 2.1.

For ease of reference, R1 will be called 'command' entry time and R2 'data' entry time. These words suggest themselves by considering the relationship between the basis of the measures and the pragmatics of the source message.

(c) Errors

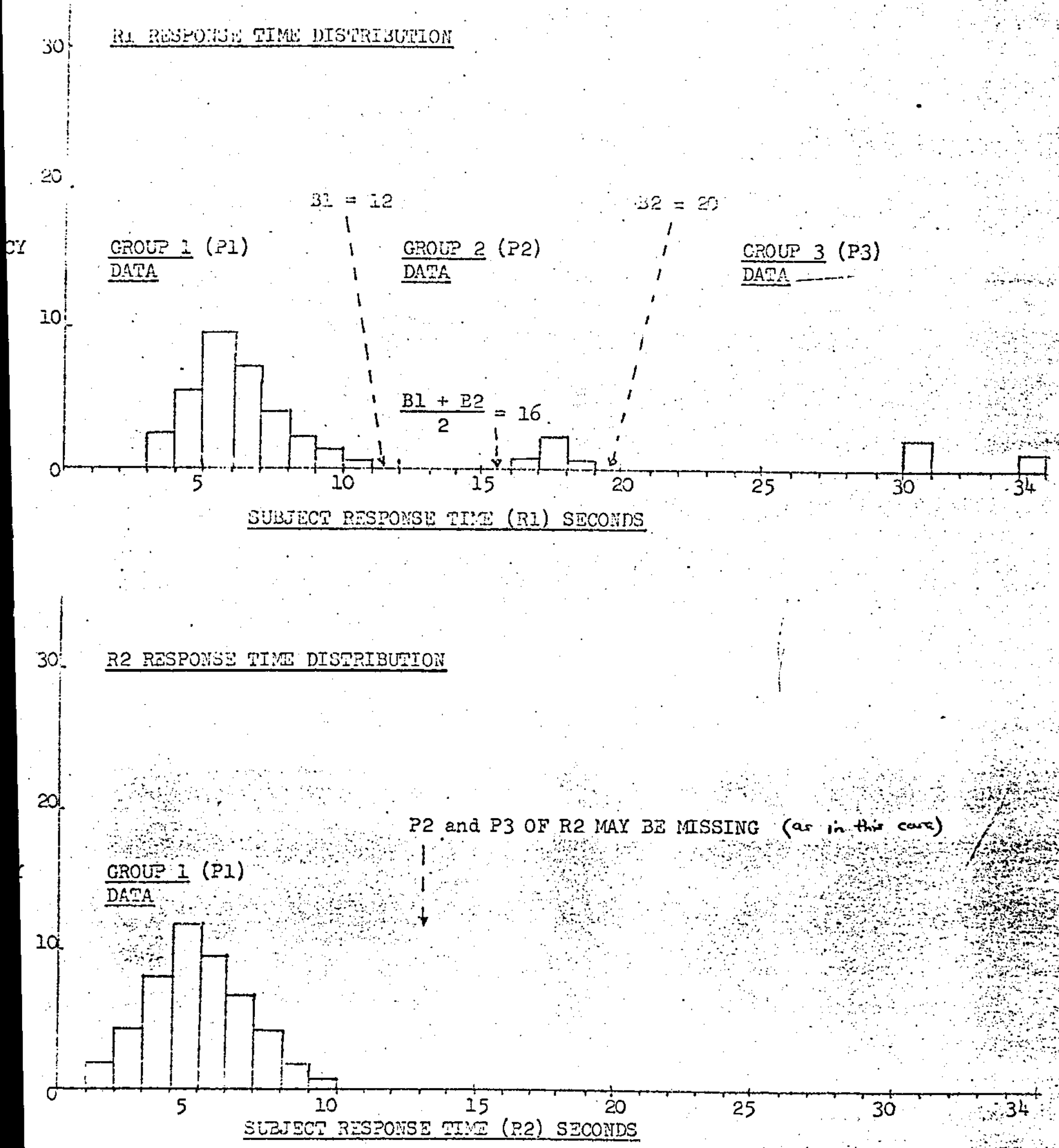
Apart from failure to solve the problem, errors were classified as 'logical errors' in which the task rule was broken or 'interface errors', in which typing errors (or mis-picks) occurred. All logical errors were logged but only typing errors (for both R1 and R2) and not mis-picks were logged.

6.2 Acceptability

(a) Relative Acceptability

In this experiment, each subject was required to rank the experimental trials differentiating between the optimum and sub-optimum systems as well as the input characteristics.

NOTE - SEE TEXT FOR EXPLANATION OF R1, R2
- SEE APPENDIX 4 FOR EXPLANATION
OF B1 and B2



PH 1 - AN EXAMPLE OF THE ENTRY TIMES USED
FOR DECIDING B1 and B2

(b) Acceptability of a System

A questionnaire was designed (shown in Table 5) for assessing the acceptability of the sub-optimum systems used. The same questionnaire was used for both input methods in the experiment and each subject was required to fill it in after completion of each experimental condition.

On the basis of the results of this questionnaire, the relationship between the subjective and objective measures of response time and reliability could be made. Within the context of this work, the system was considered acceptable if it is fast (Qu.6), reliable (Qu.2), easy to use (Qu.4) and helped the subject solve the problem (Qu.1,9).

The derived measure of acceptability (Q) used here is defined as the scores of the questions 6, 2, 4 and 1 added together. It is assumed that all factors represented by these scores are of equal importance to acceptability and that the measurement categories of each question are on an equal interval scale. With these assumptions and using the score values shown in Table 6, the acceptability score ranges between 0 (Acceptable) and 13 (Unacceptable).

associated item or box as quickly as possible.

How difficult was it for you to work out the solution to your problem?

very difficult	quite	quite easy	very easy
(3)	(2)	(1)	(0)

About how many times did the computer do or do not do something you did or did not expect?

very often	quite often	some-times	never
(3)	(2)	(1)	(0)

In relation to the previous question, whic aspect(s) were involved?

Program errors ...... ☐
Hardware " ...... ☐
Your misunderstanding ☐

How easy did you find you could tell the computer what to do next?

very easy	quite easy	quite difficult	very "
(0)	(1)	(2)	(2)

In relation to the previous question, which aspect(s) would you like improved?

		How?
Command names	☐	....
Input procedure ..	☐	....
Computer Response .	☐	....
Error handling	☐	....
Other (explain) ..	☐	....

hat did you think of the computers' esponse time in relation to:

	Consistent	Not Consistent	too long	just right	too fast
Inputting commands	(1)	(0)	(1)	(0)	(1)
Carrying out your command	(1)	(0)	(1)	(0)	(1)

n relation to the previous question, ndicate your preference for the computer esponse time.

	Seconds				
	1	1-5	6-10	11-20	20
Inputting Commands					
Carrying out commands					

ow often do you think you would have sked for help in

	Number of times							
	0	1	2	3	4	5	5-10	10
telling the computer what to do								
understanding what the computer has done								
trying to solve the problem ..								
other (explain)								

ccording to your previous experience in olving this or similar problems, did you nd using the computer

		How or Why?
helped you more	☐	....
was about as useful as other aids	☐	....
was less useful than other aids	☐	....
was a positive hinderance	☐	....

E THE SCORES IN THE BOXES FOR QUESTIONS 1, 2, 4, and 6 ARE USED FOR DATA ANALYSIS

TABLE 5

QUESTIONAIRE ON ACCEPTABILITY
USED FOR SUB-OPTIMUM SYSTEM

: bracketed Figures are assigned Scores for that response.

7. Subsidiary Variables and Measures

The purpose of using subsidiary variables was to allow testing of hypotheses in addition to those for which the experimental design was specifically chosen.

7.1 Independent Variables

These were the personality traits of anxiety and the occupation and typing ability of the subject. Anxiety trait refers to the emotional disposition of the person and was measured by using a standard questionnaire developed by C.D. Spielberger and his associates and called the 'STAI'. The score of the STAI lies on a range of 20 (not prone to anxiety) to 80 (anxiety prone). Appendix 2.2 describes the technique for collecting the data.

The subjects' occupation was categorised into four groups: draughtsman (D), Naval Architects (N), Computer Programmers (P) and the rest (O). 'O' was made up of typists and secretaries. The typing ability was assessed by means of a program which measured the inter-character typing time for a subject typing his/her own name and address on a teletype. The first and last characters were excluded from the assessment. The mean of the log transformed inter-character time was used to characterise typing ability. The transformed data was highly positively skewed and unsuitable for characterising typing ability. Table 6 shows the values of the subsidiary variables for the sample used in this experiment.

7.2 Dependent Variables

These were the same as for the previous Section (6).

Subject No.	Anxiety Trait Score*	Occupation+	Assigned Rank	Typing Ability Ø
1	48	P	1	4.34
2	46	P	1	3.32
3	49	O	4	2.75
4	45	D	3	6.32
5	32	P	1	3.83
6	46	D	3	6.83
7	49	O	4	2.97
8	39	N	2	5.46
9	35	D	3	5.50
10	34	P	1	2.33
11	35	D	3	7.43
12	31	N	2	6.12
13	34	P	1	3.61
14	30	O	4	3.36
15	43	P	1	5.50
16	58	D	3	6.62
17	34	N	2	3.24
18	49	D	3	5.51
19	45	P	1	3.91
20	37	O	4	2.91
21	34	P	1	3.59
22	38	P	1	3.90
23	37	N	2	5.56
24	34	N	2	3.71

KEY :

* Measured by Questionnaire – see Appendix 2.2
\+ P = Programmer; D = Draughtsman; N = Naval Architect; O = Other (secretarial; admin.)
Ø Mean of log-transformed distribution of inter-character times.

Table 6

Details of Subsidiary Subject Variables

8. Treatment of Data

The raw data on which the following analyses are based is given in Appendix 2.3. Readers not wishing to study details of the treatment and analysis may go to section 10 without penalty of misunderstanding the discussion therein. In the following analysis and that of the following chapter, all tests between means are two-tailed.

8.1 Work Production (WP)

8.1.1 Information Entry Time

Two measures were taken of each of the three peaks (P1,P2,P3) of the R1 and R2 distributions: these were the frequency (n) and mean (t). The number in P1 plus the number in P2 were added together to give the total number of logical solution steps used in solving the problem.

The mean times of the second and third peaks of R1 and R2 were corrected for each subject by subtracting the mean time of the appropriate peak 1. This was so that the values of peak 2 and 3 could be examined independently of peak 1. Peak 2 may then be labelled as mean time per 'recall' (evaluation) and peak 3 as the mean 'planning' time in accordance with the convention of the previous experiment (Chapter 5).

The eleven measures were cast into 2x2x2x24 tables and the paired t test was used to test for differences between mean values due to the system (S) and the input methods (I). The results are given in Table A26 of Appendix 2.

	N↓				N↑			
	X↓		X↑		X↓		X↑	
	L1	L2	L1	L2	L1	L2	L1	L2
Subject Numbers	12	14	22	10	20	16	2	15
	23	9	24	8	4	6	19	18
	7	11	5	21	1	3	13	17

TABLE 7

Allocation of Subjects to Design Cells

Key: N - neuroticism trait

X - general computer experience

L - order of presentation of experimental conditions

↑,↓ - high and low levels of the factors

Numbers in table are subject numbers (see Tables 3 and 4).

8.1.2 Overall Time to Solve

For the optimum system there is a significant difference ($t = 2.91$ with $df = 22$) between the means of the overall times to solve the problem using the light pen and keyboard.

In the sub-optimum system, six out of 24 subjects failed to solve with the light pen whereas 18 out of 24 failed with the keyboard. The failure times were treated as censored data in comparing the overall times using Cox's test Cox (1972)).

The resulting value of Z (3.39) shows that, at the 5% level, there is a difference between the mean overall times to solve the problem using the light pen and keyboard.

The overall-time-to-solve data were analysed for each optimum system condition between people using the t-test. In the data for the sub-optimum system, the failure-to-solve times were taken to be the time-to-solve.

8.1.3 Failure to Solve

The McNemar test for related samples was used to test for differences due to input methods on the sub-optimum systems data.

For within the sub-optimum systems data, the Fisher exact probability test for 2 independent samples was used to test the hypothesis that user characteristics influenced the failure rates. For the optimum systems, 2x2 contingency tables were constructed for user experience, neuroticism, and order of presentation within input methods. Interactions between the variables were not examined. The level of significance chosen was $p \leqslant .05$ (i.e. 5%).

8.1.4 Logical and Typing Errors

The frequencies of these errors were cast into appropriate tables for the ANOVA technique. The typing errors only applied for the keyboard input. The resulting data were subjected to a paired t-test.

8.2 Acceptability (A)

8.2.1 Rank of Acceptability

The Table of Ranks was subjected to a 2-way Friedman ANOVA. The Table was split into groups according to input, order and system characteristics and the sign test used to test for differences between ranks. The within systems groups were tested by selecting from the rank table according to the characteristics of the subjects and again applying the sign test.

8.2.2 Score of Acceptability

The score was treated as meeting the requirements of the t technique (i.e. equal interval data) and analysed in the same way as the work production entry time data.

8.3 Relationships between Acceptability and Work Production

8.3.1 Work Production Data (Excluding Errors) and Acceptability Scores

Subjects were ranked in order of the number of responses and mean information entry time for Peaks 1 and 2 and for R1 and R2. Peak 3 was not used because of the low number of responses. The subjects were ranked on acceptability score.

The Spearman rank correlation coefficient was calculated for each pair of measures within systems (S) and input method (I). Correlations between systems and inputs were calculated for those measures which showed significant differences for I and S.

8.3.2 Work Production Errors and Acceptability

The subjects were grouped according to whether they were above or below the median acceptability score for each system. 2x2 contingency tables were constructed using error/no-error groups for solution and logical errors with each system and input type. The Fisher exact probability was calculated for each table.

For the typing errors, the subjects were ranked on acceptability for the keyboard only and the Spearman rank correlation coefficient calculated.

8.4 Relationships Within Work Production Measures

Subjects were ranked according to their response for Peaks 1, 2 and 3 of R1 and R2 for each input type and system. The Spearman rank correlation coefficient was calculated for each pair and tested for significance using the t-test.

8.5 Effects of Subsidiary Variables

8.5.1 Work Production

For each input condition, subjects were ranked on the measures of work production (R1, R2, peaks 1 and 2, overall time and typing errors). The Spearman rank correlation coefficient was calculated between each R1, R2 rank set and each rank set corresponding to the 4 subsidiary variables. Table 6 gives the assigned rank of the occupation groups used in this analysis. The failure-to-solve and logical errors data were treated as binary data and subjects were grouped according to whether they were above or below the group median score, (except Occupation groups which were 'programmers' and 'non-programmers'). 2x2 contingency tables were constructed and the Fisher exact probability calculated.

8.5.2 Acceptability

Subjects were ranked on acceptability scores and the correlation coefficients calculated as for the entry time data.

9. Analysis of Results

The data presented in Appendix 2.3 were appropriately analysed and Table 8 shows a summary of the results. The 5% level of significance was the criterion used for inclusion of factors in this Table.

The general description 'system' (S) refers to the combination of task difficulty and system characteristics (i.e. fast and reliable or slow and unreliable). 'Input' (I) refers to the input method and associated pragmatic feedback. For simplicity, levels of this variable are referred to as light pen and keyboard.

9.1 Work Production

9.1.1 Between Systems and Inputs

Tables 9 and 10 show the measures sensitive to I,S and interactions between them.

9.1.2 Within Systems and Inputs

Tables 11 and 12 show those measures sensitive to differences between people within both system and input conditions.

TABLE 8

Summary of Significant Measures (5% level)

Measure				Between Systems Input(**I**), System(**S**) (Ref. TABLE A25 (Appendix 2))	Within Systems (Between People). Order(**O**), Experience(**E**), Neuroticism (**N**) Light Pen Optimum *System*	 Light Pen Sub-opt.	 Keyboard Optimum *System*	 Keyboard Sub-opt.
VP	R1*	P1	n	**S**,(logical moves used)ie,corrected for n,P_2&n of P_3	**NSD**	**NSD**	**NSD**	**NSD**
			$\bar{t}$	**I**,**S**,	"	"	"	"
		P2	n	**I**,**S**,	"	"	"	"
			$\bar{t}_c$	**I** (corrected by subtracting t of R1,P1)	"	"	"	"
		P3	n	**I**,**S**	"	"	"	"
			$\bar{t}_c$	**S** " " " "	"	"	"	"
	R2*	P1	n	NT	"	"	"	"
			t	NSD	"	**NXE**	"	**E**
		P2	n	**S**	"	**NSD**	"	**NSD**
			$\bar{t}_c$	**S** (corrected by subtracting t of R2 P1)	**EXO**	**E**	"	"
		P3	n	**S**	**EXO**	**NSD**	"	"
			$\bar{t}_c$	**S** " " " "	**NSD**	**EXO**	"	"
			OVTM+	**I**,**S**	**NSD**	**E**	**NSD**	**NSD**
		Errors	Failure to solve	**I** within **S**	**NT**(all solved)	**E**,**O**	**NT**(all solved	**O**
			Logical*	NSD	**N**	**O**	**NSD**	**NSD**
			Typing*	**S** within **I**	**NT**	**NSD**	**NT**	**NSD**
θ			Ranks*	**I**,**S**	**NT**	**NXE**	**NT**	**NSD**
			Score*	NSD between **I** within **S**	**NT**	**NXE**	**UT**	**NSD**

Key
(NSD - No Significant differences at 5% level for any variables and that measure
(θ - Acceptability
(WP - Work Production
(NT - Not tested
(* - See Text for explanation

TABLE 9

System Measures Sensitive to I,S

Measure	System/Task	Light Pen	Keyboard	
R1,P1,$\bar{t}$	Optimum/easy (O) Sub-Optimum/hard (S)	4.80 4.40	4.65 2.67	I,S,
R2,P2,n^*	O – S	0.54 1.42		S
R2,P2,$\bar{t}_c^*$	O S	2.94 1.51		S
R1,P2,$\bar{t}_c$	–	10.20	13.75	I
R1,P3,n^*	O S	0.38 1.30	0.65	I,S
Measure	System Task	LP and Keyboard		
R1,P1,n	O S	28 39		S
R1,P3,$\bar{t}_c^*$	O S	21.5 58.5		S
R1,P3,n^*	O S	0.12 0.75		S

* Signifies low numbers in Sample : (non-normal)

TABLE 10

Overall Measures Sensitive I.S.

Measures	Light Pen		Keyboard		
(all per subject)	Optimum	Sub-optimum	Optimum	Sub-optimum	Significant for
Mean overall time to solve (failures excluded) in Minutes	5.813	15.91	10.24	24.26	I,S
Percentage failure to solve	–	25	–	75	I within S
ercentage Typing rrors (100xtyping rror/total number f typed inputs)	–	–	6.26	1.65	S within I
cceptability Rank	1	3	2	4	I and S

TABLE 11

Work Production Measures Between People

<table>
<tr><td></td><td colspan="8">Light Pen</td><td colspan="3">Keyboard</td><td></td></tr>
<tr><td>Measure</td><td colspan="4">Optimum</td><td colspan="4">Sub-optimum</td><td>Optimum</td><td colspan="2">Sub-optimum</td><td>Significant for</td></tr>
<tr><td></td><td colspan="2">E↓</td><td colspan="2">E↑</td><td colspan="2">E↓</td><td colspan="2">E↑</td><td></td><td colspan="2"></td><td></td></tr>
<tr><td></td><td>01</td><td>02</td><td>01</td><td>02</td><td colspan="2"></td><td colspan="2"></td><td></td><td colspan="2"></td><td></td></tr>
<tr><td>R2 P2 $\bar{t}_c$</td><td>9.20</td><td>1.61</td><td>0.91</td><td>4.25</td><td colspan="2">18.5</td><td colspan="2">5.81</td><td>-</td><td colspan="2"></td><td>E X O,E</td></tr>
<tr><td>R2 P3 N</td><td>3</td><td>0</td><td>0</td><td>2</td><td colspan="2">-</td><td colspan="2">-</td><td>-</td><td colspan="2"></td><td>E X O,</td></tr>
<tr><td rowspan="3">R2 P1 $\bar{b}$</td><td rowspan="3" colspan="4"></td><td colspan="2">N↓</td><td colspan="2">N↑</td><td rowspan="3">-</td><td>E1</td><td>E2</td><td rowspan="3">N X E,
E</td></tr>
<tr><td>E↓</td><td>E↑</td><td>E↓</td><td>E↑</td><td rowspan="2">3.42</td><td rowspan="2">2.52</td></tr>
<tr><td>3.91</td><td>2.69</td><td>3.00</td><td>3.85</td></tr>
<tr><td rowspan="3">R2 P3 $\bar{t}_c$</td><td rowspan="3" colspan="4"></td><td colspan="2">E↓</td><td colspan="2">E↑</td><td rowspan="3">-</td><td rowspan="3" colspan="2">-</td><td rowspan="3">E X O</td></tr>
<tr><td>01</td><td>02</td><td>01</td><td>02</td></tr>
<tr><td>11.60</td><td>5.06</td><td>13.50</td><td>0.00</td></tr>
</table>

TABLE 12

Acceptability and Overall Work Production

Measures between People.

Measure	Light Pen		Keyboard		Significant for
	Optimum	Sub-optimum	Optimum	Sub-optimum	
Overall time to solve(or quit)	–	E ↑ 12.85 E ↓ 18.70	–	–	E –
Failure to Solve (%)	–	E ↑ 58.3 E ↓ 91.6	–	–	E
		01 58.3 02 91.6	–	41.6 8.3	0
Frequency of Logical Error	N ↑ 9 N ↓ 0	01 6 02 0	–	–	N,0
Score of Acceptability (0 = acceptable 13 = unacceptable	–	N↓: E↓ 5.0, E↑ 6.3 N↑: E↓ 7.0, E↑ 5.83	–	–	NXE

9.2 Acceptability (O)

9.2.1 Between Systems and Inputs

Table 10 shows the relative acceptability of the different system and inputs across all subjects.

9.2.2 Within Systems and Inputs

Table 12 shows the differences in acceptability acores due to individual differences between the subjects.

10. Discussion of Results

This section makes reference to Tables in Section 9 and Tables in the Appendices containing analysis data. (Table numbers in Appendix 2.3 are preceded by 'A').

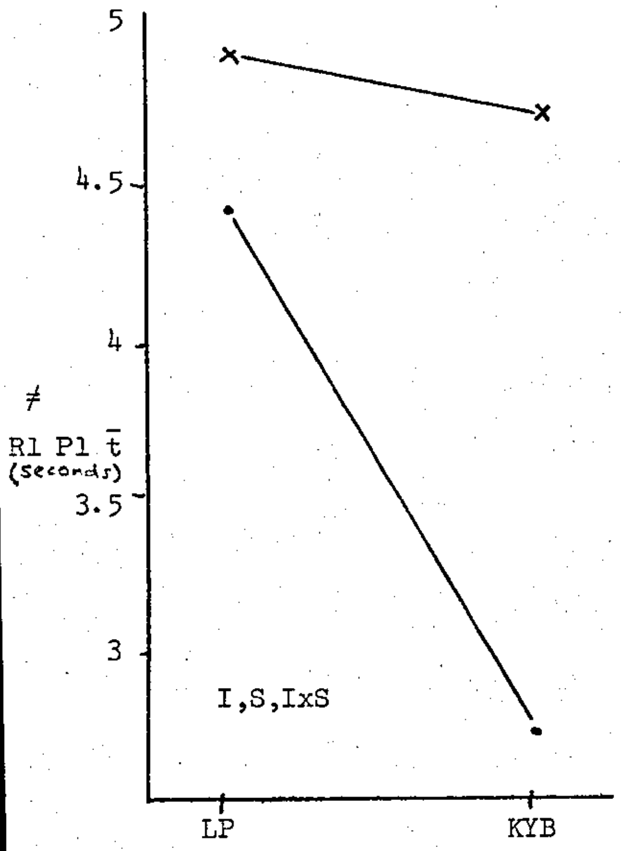
10.1 Work Production

10.1.1 Between Systems (S) and Inputs (I)

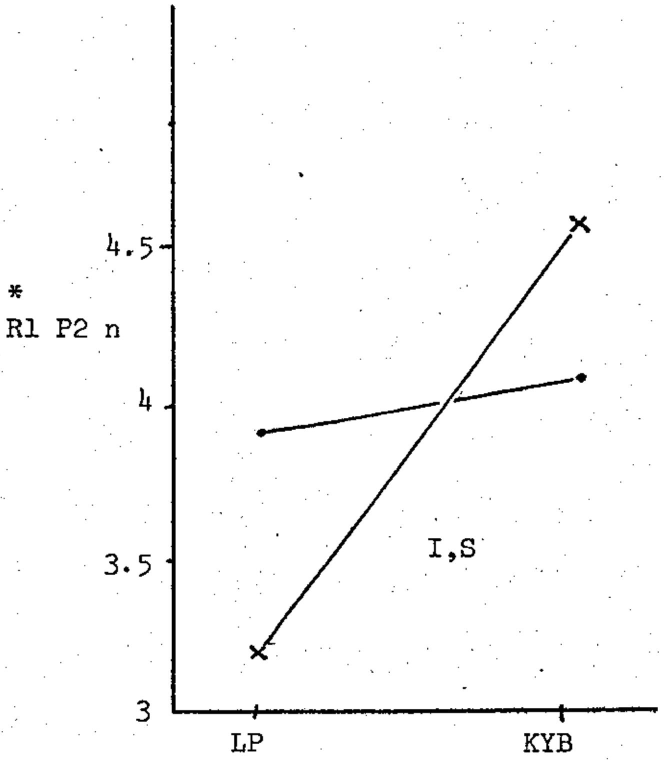
Differences due to I are shown in the measures associated with the first (R1) information input rather than the second (R2) input. The converse is true for differences between people. Differences between systems appears in most measures.

Graph 2 shows that subjects took longer to make a command response with the light pen (~ 4.92) than the keyboard (~ 4.80) and that this difference was amplified in the sub-optimum system. An explanation of this may be made if it is assumed that the subjects' attention remained mainly on the keyboard and single letters of commands were typed

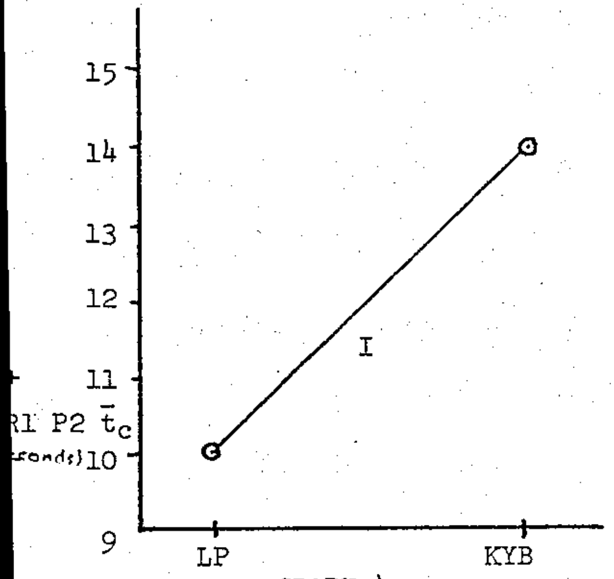
INPUT DEVICE VS WORK PRODUCTION MEASURES



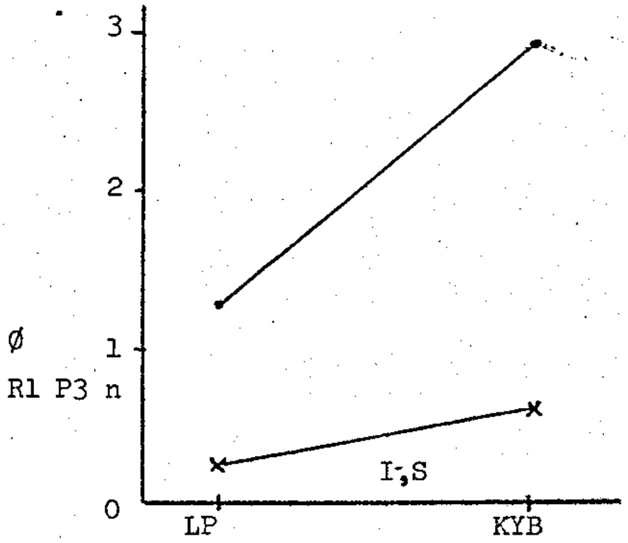
GRAPH 2



GRAPH 3



GRAPH 4



GRAPH 5

Key:

≠ Mean Time to enter a Command per Subject (Seconds) [R1, P1, Mean]
* Mean Number of Recalls per Subject for R1
\+ Mean Time per Recall per Subject for R1 (Seconds)
∅ Mean Number of Plans per Subject for R1

•—• Sub Optimum/Difficult System
×—× Optimum/Easy System
⊙—⊙ For Both Systems

before a computer prompt had been received. The analogous situation (of queueing a command) was not possible with the light pen input.

Subjects carried out more than two disc moves before requiring a screen update. Hence it may be assumed that their command sequence was constructed at a particular time. The input method affected the rate of input of this sequence according to its error proneness. The keyboard was less error prone than the light pen. This is particularly true in the sub-optimum system - hence the amplification of the effect.

Graphs 3 and 4 show how the number of and mean time per 'recall/evaluation' vary as a function of the input method. Graph 3 shows that less recalls were made using the light pen than using the keyboard.

Graph 4 shows that the mean time per recall was less with the light pen than the keyboard. An explanation of this is as follows. With the light pen, every disc move was confirmed by pragmatic feedback without the subject needing to change attention from the display. With the keyboard the subject could carry out as many moves as wanted, then request pragmatic feedback. On such a request, the subject would check that the computer had carried out his requests. At particular points in the solution process, this activity could compete with others for the use of short-term memory resulting in an increase both in the number of recalls and the time per recall. Thus the cost of the less constrained input without pragmatic feedback to the user was an increase in the number of recalls and possible interference with the solution process.

On the basis of this explanation, the difference shown on Graph 3 between the systems may be expected. Task difficulty is directly related to the number of moves and sub-goals to solution and (for the keyboard) the number of requests to update the screen. The longer the computers speed of response, the greater the possibility of forgetting a solution goal or sequence and hence more recalls are needed.

More subjects failed to solve the problem with the keyboard than with the light pen and each failure resulted in a planning activity before quitting, More plans would be expected with the keyboard than with the light pen. Graph 5 shows this.

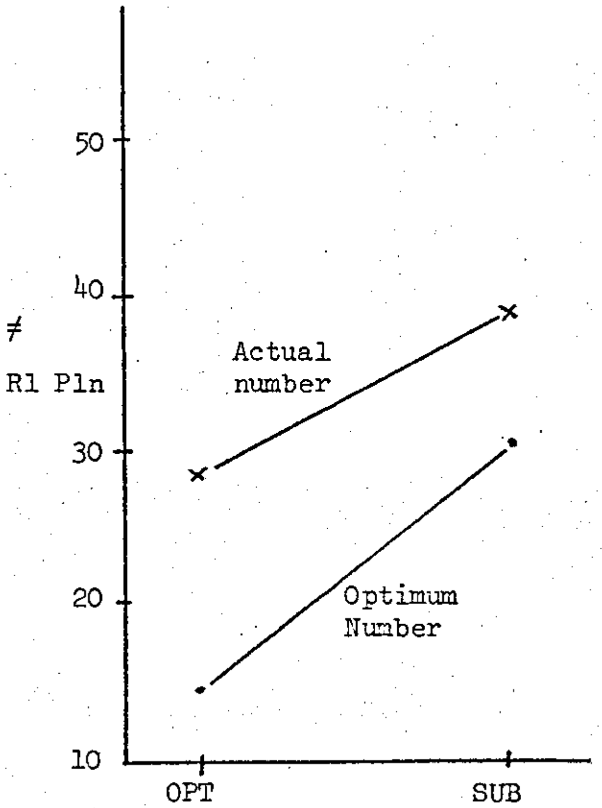
Graphs 6-8 are of those measures which show differences due to the systems variables. Graph 6 shows the difference in the number of moves to solve which is mainly due to the task difficulty rather than the speed of system response or system reliability. Differences shown in Graph 7 may not be attributed to any particular system variable, but to the combination of task difficulty, system reliability and system speed of response.

In Graph 8 the measure is concerned with data (R2 distribution) rather than command (R1 distribution) input. The planning activity is a major break in data entry which could be caused by system unreliability, rather than the task difficulty or slow speed of response.

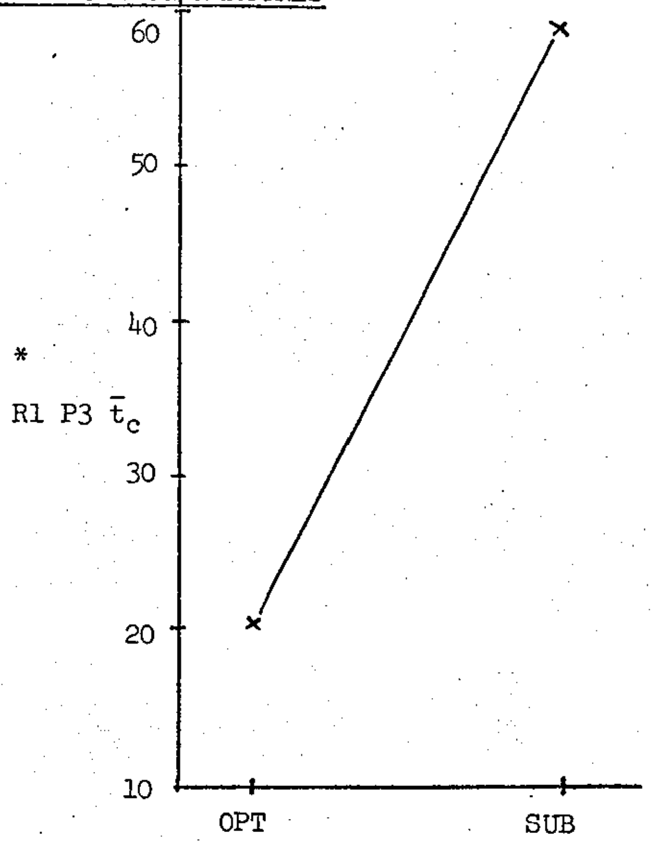
The overall time to solve using the keyboard was longer than for the light pen for both the optimum and sub-optimum systems. This may be attributed to the extra time needed to request pragmatic feedback with the keyboard input.

With the keyboard, the typing errors in Table 10 show that more mis-typed inputs occurred on the optimum system than the sub-optimum system. This may be interpreted as a strong learning effect of the input commands.

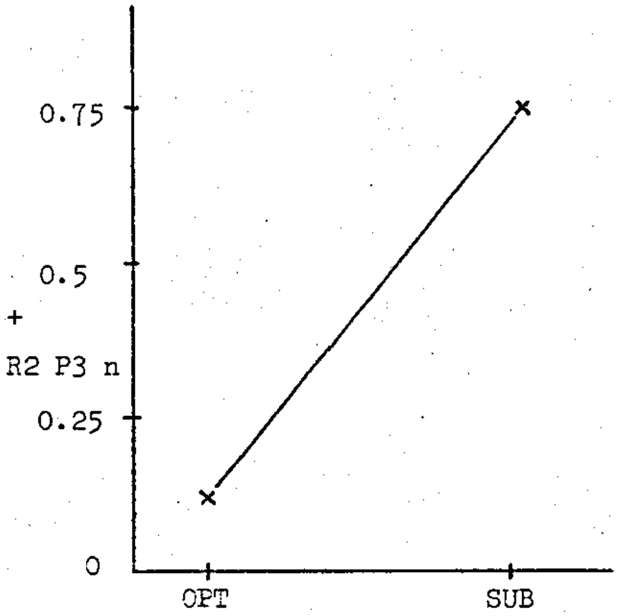
SYSTEM/TASK LEVELS VS WORK PRODUCTION MEASURES



GRAPH 6



GRAPH 7



GRAPH 8

Key:

≠ Measure Number of Logical Steps to Solve

* Corrected Mean Planning Time of R1 (Seconds)

\+ Mean Number of Plans per Subject in R2

10.1.2 Within Systems and Inputs

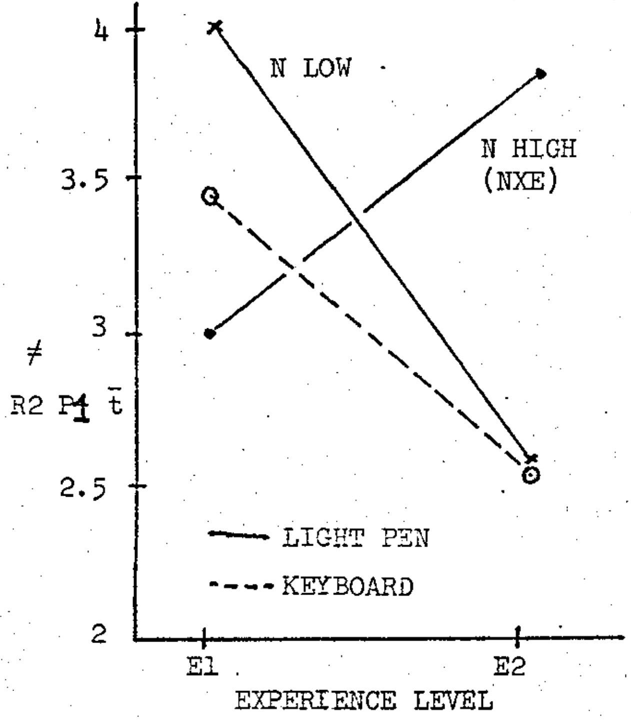
(a) Light Pen Optimum System

Graphs 10, 11 and 12 all show that there is an interactive effect between the subjects general and specific computer experience. Generally experienced subjects made less plans but longer thought out plans and took less time evaluating their solution than generally inexperienced subjects when first presented with the problem.

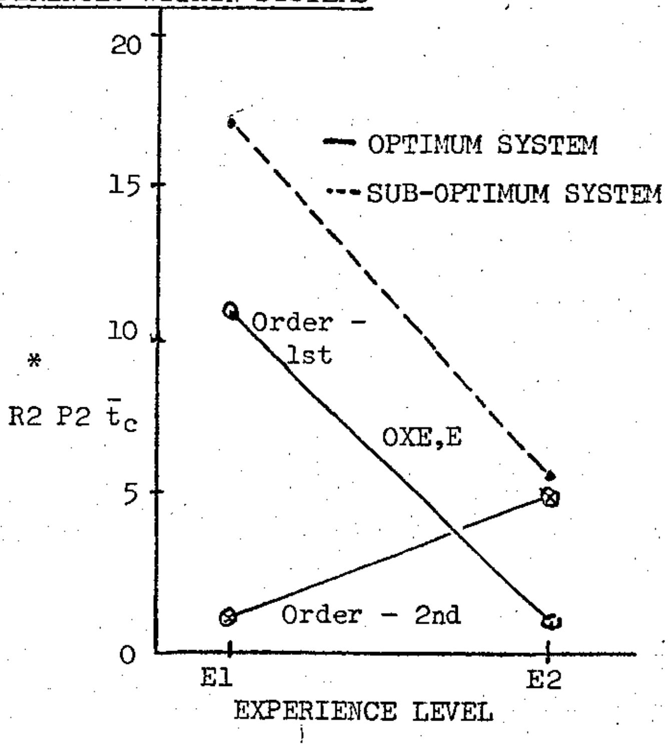
The converse is true once specific experience has been gained on the problem using the keyboard. A possible explanation is that, having just completed the sub-optimum difficult keyboard task, experienced subjects modified their behaviour according to a new criterion of penalty-for-error (input errors) such that they were more cautious. Hence, the larger number of plans (but less time per plan) and longer evaluation times during solution than the inexperienced subjects who did not adapt but merely used their specific experience in a more positive but reckless way.

Generally, experienced subjects took less time recalling or evaluating their solutions than inexperienced subjects, regardless of the specific experience gained in this experiment.

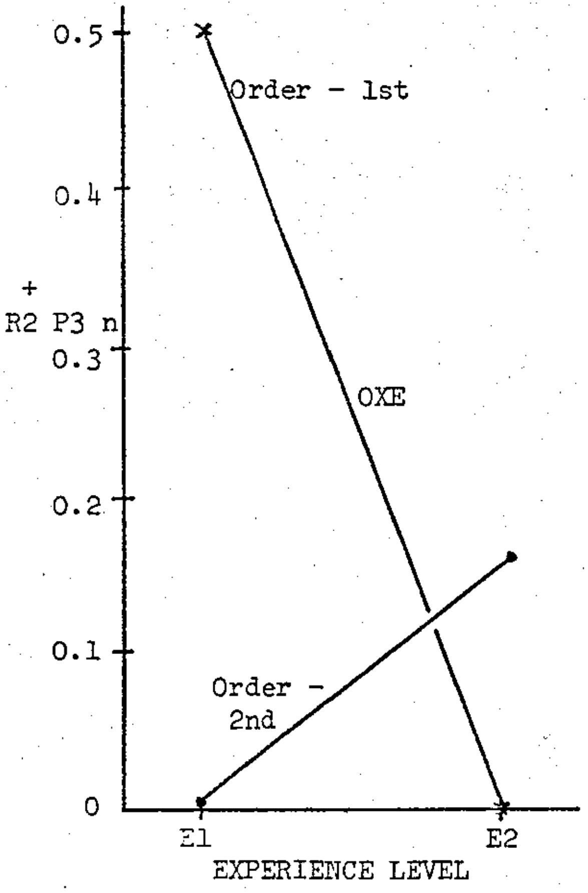
MEASURES SENSITIVE TO INDIVUAL DIFFERENCES WITHIN SYSTEMS



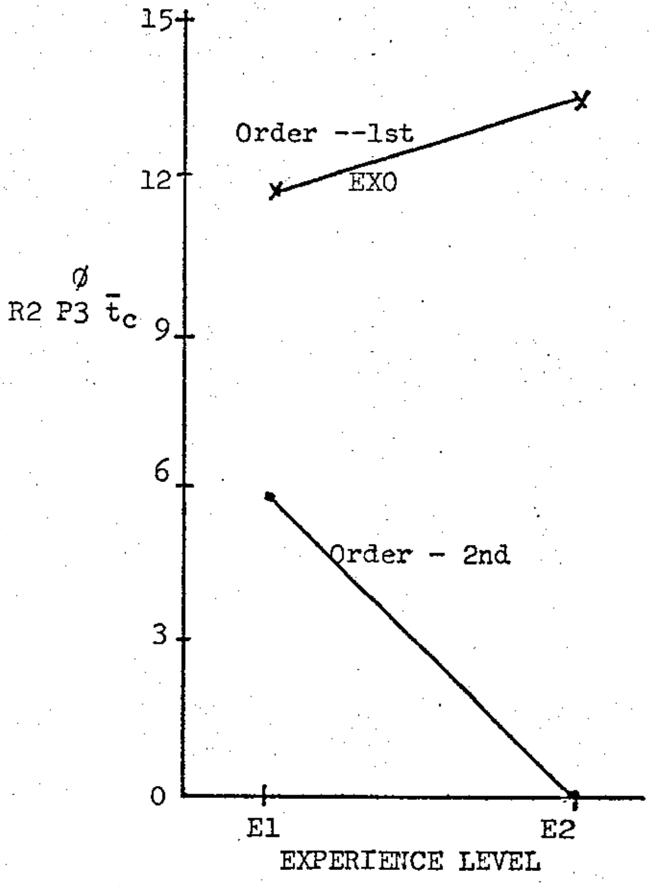
GRAPH 9



GRAPH 10



GRAPH 11



GRAPH 12

Key:

≠ Mean Time per Subject to Input Data using Sub-optimum System

* Mean Time per Recall on "Data" Entry for Light Pen

\+ Mean Number of Plans per Subject on R2 for Light Pen Optimum System

Ø Mean Planning Time per Subject in "Data" Entry using Light Pen Optimum System (Seconds)

Finally, the frequency of logical input errors (illegal moves) depended on the neuroticism of the subjects such that unstable subjects made more errors than stable people.

(b) Light Pen - Sub-Optimum System

Graph 10 shows that the advantage afforded by general experience mentioned in the previous section was sustained when the sub-optimum system is considered. Graph 9 shows that the trait of neuroticism interacts with general experience. Stable experienced subjects took about 2.7 seconds to input the data (R2, P1) whereas stable inexperienced subjects took about 4 seconds. This trend is reversed for the unstable (neurotic) subjects so that experienced unstable subjects took longer than inexperienced unstable subjects. An explanation is that unstable inexperienced subjects proceeded recklessly compared with stable inexperienced subjects. After gaining experience in which the consequences of this recklessness became apparent, unstable subjects adapted their behaviour to be slower. This happened to a much lesser extent with stable subjects who therefore were able to proceed quickly with confidence. Different personality traits did not show differences in the numbers who solved the problem (Table A15).

More generally experienced subjects took significantly less time to solve the problem, and more of the generally experienced subjects solved the problem than inexperienced subjects. The effects of specific experience were similar in that those who had already used the keyboard made less logical errors in solution and more of them solved it.

(c) Keyboard - Optimum System

Measures on this were not significant for any of the inter-personal factors.

(d) Keyboard - Sub-Optimum System

Graph 9 shows that general computer experience was associated with higher rates of typing data into the terminals. This did not, however, mean that more experienced subjects solved the problem (Table A15). Observation showed that fast rates were mainly due to anticipation of the computer output of a prompt and premature input, (i.e. typing before the prompt had been printed).

Table A14 shows that specific experience gained on the problem due to the order of presentation was important in that more subjects solved the problem after having tried before (using the light pen).

10.2 Acceptability

10.2.1 Between Systems and Inputs

Table 10 shows the overall ranking of the conditions and that the system characteristics were more important than the types of input device. Within both the optimum and sub-optimum systems the light pen was preferred. The acceptability scores were not significantly different (Table A19) for the sub-optimum systems. The rank order of the conditions agrees with the rank order of the number of plans (R1, P3) shown in Graph 5.

10.2.2 Within Systems and Inputs

Table 12 shows that the acceptability score for the light pen sub-optimum system was a function of the interaction between general computer experience and the neuroticism of the subject. Stable experienced subjects found the sub-optimum conditions less acceptable than stable inexperienced subjects. This was reversed for unstable subjects. Graph 9 shows that this would be expected if the basis for the acceptability judgement was the mean time to input data (R2, P1 mean).

10.3 Relationships Between Measures

10.3.1 Work Production

(a) Within Systems and Inputs

The 8 measures which were found significant (Table 8) were those of R2, overall time to solve, and errors. Some of these have weak correlations (Tables A20, A21) with R1 measures and, for the light pen input on the optimum system (R1- P2-mean), correlates with overall time to solve. Hence, the more evaluations needed while entering commands, the longer the time taken to solve the problem. Other interpretations of these correlations are also possible.

(b) Between Systems and Inputs

The 11 significant measures of Table 8 are some of R1, R2, overall times and errors. Table A22 shows that with the same input conditions, the measures of R1-P1-mean and R2-P1-mean are related between systems. That is, these measures of performance may be used in a predictive way across systems provided the input characteristics, feedback and task are similar.
In optimum systems, the predictive value is also preserved across light pen and keyboard for R1-P1-mean i.e. the time for entry of a command may be predicted from one set of input characteristics to another.

10.3.2 Acceptability and Work Production

Tables A20 and A21 show that there is no simple correlation between acceptability score and the R1, R2 or overall time to solve measures of work production for the sub-optimum system.

11. Effects of Subsidiary Variables

Table A23 shows the results of the tests between the four variables; anxiety trait score (A), extraversion score (E), occupation (O) and typing ability (T); and the measures of work production and acceptability. Typing ability (as measured by rate of typing) appears most often as being significantly related to other measures. A fast rate of typing outside of a problem solving context is a skill carried over appropriately to interactive problem solving resulting in more and faster inputs. However, this skill is not significantly related to the overall time-to-solve. For the sub-optimum keyboard system, typing ability relates to the failure to solve such that more skilled typists solved the problem than unskilled typists. An explanation is that for the unskilled typists, the capacity to cope with the situation was exceeded because of the need to perform more 'search and locate keys' tasks. This was not true for the skilled typists.

The explanation of other relationships between T and the other measures is less obvious. For example, it is reasonable to expect that on the optimum system a higher typing rate results in more typing errors, but the reason why T is related to frequency of light pen picks (in plans and recalls) and overall solution time with the optimum system is not clear. In the table there are 128 correlation coefficients of which 6 would be expected to occur by chance at the 5% level of significance. Perhaps these may be in that category.

Table A24 shows that of the 6 subsidiary measures only two are related across subjects. These are occupation and experience. This is as expected since a system programmer obviously has a general computer experience compared with a draughtsman. Personality traits, typing skills and occupation were not expected to be related within the small sample size of this experiment.

12. Conclusions

12.1 Work Production and Problem Solving

A number of measures of work production were taken. Depending on the measurements, particular factors and their interactions were significant. Two groups of measures were identified; one at a command level (first input of information) and the other at a data input level (a following sequence). Within each level, three further sub-divisions were identified according to whether the subject was following a solution plan, evaluating (or recalling) a solution plan or planning a solution. The frequency and time of each activity in each group was used as a measure.

In general, the command group did not show differences between people but did show differences due to input methods and systems. Conversely, the data group showed differences between people rather than systems and inputs.

Measures of problem solving performance showed differences between systems, inputs and people. There was little correlation between the measures. However, some measures could be used to predict results between systems with the same input methods. In optimum systems, some measures could be used to predict results that would be obtained with different systems and input methods.

The differences between light pen and keyboard are summarised as follows:-

(i) less recalls and plans were made with the light pen.
(ii) the recall time was less with the light pen.
(iii) the solution step time was greater with the light pen and was increased by poor system characteristics.
(iv) the overall time to solve was less with the light pen.
(v) more people solved the problem with the light pen.
(vi) the light pen was preferred to the keyboard at the same system level.

An explanation was suggested that these differences arose because of the differences in input method and associated pragmatic feedback. These led to a difference in ease of learning between the two conditions.

The differences between systems are summarised as follows:-

(i) more solution steps were taken with the sub-optimum system (difficult task).

(ii) more time was needed for planning with the sub-optimum system (difficult task).

(iii) less typing errors were made with the sub-optimum keyboard system.

(iv) system differences increased the effect of differences between the input characteristics.

The main explanation for these differences was that task difficulty increased in the sub-optimum system, although poor system reliability influenced (ii). Speed of response interacted with the light pen input device which was responsible for (iv) and ease of learning was responsible for (iii).

The differences between people were mainly shown in the time for light pen data input. An explanation is that the light pen was error prone whereas this was not the case with the keyboard inputs; hence individual differences were not as apparent in the keyboard data. The data showed that this difference became more important when the system was sub-optimal and the task was more difficult.

The differences between people may be summarised as follows.

(a) Light Pen Optimum System

(i) With no specific experience, generally experienced subjects made less plans and took longer to do so than inexperienced subjects. Also, they took less time evaluating their input. This was reversed when specific experience had been gained.

(ii) Regardless of specific experience, experienced subjects took less time recalling or evaluating their input than generally inexperienced subjects.

(iii) Unstable (neurotic) subjects made more logical errors than stable subjects.

(b) Light Pen Sub-Optimum System

(i) The advantages of general experience are as for (a) above.

(ii) Stable experienced subjects were faster to input descriptive data with the light pen than stable inexperienced subjects. This was reversed for unstable (neurotic) subjects.

(iii) More of the generally experienced subjects solved the problem and in a faster time than inexperienced subjects.

(iv) Less logical errors were made after specific experience on the task.

(c) Keyboard Sub-Optimum System

Subjects with general experience had confidence to anticipate the computer prompt and so input information faster than inexperienced subjects.

12.2 Acceptability

Preferences for the light pen were dominant at a particular system level, but the keyboard optimum system condition was preferred to the light pen sub-optimum system. Hence, the differences between input methods were subjectively less important than task and system differences.

The acceptability scores of particular sub-optimum system conditions did not reveal differences between input methods. Differences between people were apparent in the scores for the light pen characteristics. Stable experienced subjects judged the sub-optimum conditions to be less acceptable than stable inexperienced subjects. The opposite was true for unstable subjects. This is what would be expected if the basis of the judgement was the time taken to input data rather than commands.

12.3 Work Production and Acceptability

Apart from the relationships mentioned in 12.2, no simple relationships were found.

12 .4 Subsidiary Variables

These related to differences between people and the most significant was typing ability. Typing skill allowed more and faster input and, with the sub-optimum system was helpful in that it did not interfere with coping with the problem. No variables (except occupation and experience) were related.

12.5 Experimental Hypothesis

The hypothesis that problem solving does not depend on the relationships between the characteristics of people, systems and input method is rejected. The dependency is complex and needs further investigation, particularly the relationship between input method and distal pragmatic feedback.

12.6 Implications for the Model

The distribution of the input times for the different types of information supports the hypothetically proposed operation of the model described in Chapter 2. Namely, that logical solution steps are derived and converted into source messages held in memory which are progressively transferred to the computer. The first part of the message transferred has a frequency distribution that is expected from the model of human problem solving; the remainder of the message is transferred as a simple sequence and therefore has a single peak distribution.

The results support the view that people adapt to different sets of circumstances dynamically through evaluative decisions as suggested in the model.

CHAPTER 7

EXPERIMENT 5

SUMMARY

This experiment was designed to test hypotheses concerning the effects of dialogue differences in relation to the effects of different input methods. Two dialogues were examined; one without automatic distal pragmatic feedback and the other with. Individual differences were not designed to be examined but taken into account in the selection process for allocation to design cells. Sixteen subjects carried out the trials with each dialogue/input method combination solving the 4 disc problem as defined in the previous Chapter.

A 3-factor ANOVA was carried out on the results using a random effects model. Problem solving performance was measured by the overall time to solve and the various measures of information quantity and rates of input defined in the previous Chapter.

The results showed that input method interacted with dialogue type to affect the input rates of pre-programmed command information, the frequency of plans, and the frequency of evaluations. The results were compared with the results of the previous Chapter. Hypothetical explanations are presented for the effects and related to the model of interactive problem solving.

The main conclusion is that the effect of not providing adequate distal feedback on performance depends on the input method according to how cognitive processes are loaded.

CONTENTS Page No.

TABLES

GRAPHS

DIAGRAM

1. Introduction

In the previous experiment, the form of the dialogue (particularly provision of pragmatic feedback) was different for each type of input method. This was deliberate in the context of that experiment in that each dialogue was chosen to be typical for those input methods. However, it may be argued that dialogue differences are also important in relation to the measures taken. This Chapter describes an experiment based on the previous one to examine the effects of dialogue differences in relation to different input methods. The experimental hypothesis is that problem solving does not depend on input and dialogue differences.

2. Objectives

The objective is to compare the effects of different dialogues (implicit vs explicit pragmatic distal feedback) and input method (as described in the previous Chapter) on interactive problem solving. This is to be achieved with the same problem as the previous Chapter with the optimum conditions as described in that Chapter.

3. Experimental Details

The following description assumes that the relevant sections of the previous Chapter describing the task and input method have been read. The experiment described in this Chapter used the same task and input method of the previous experiment.

3.1 Design

A 2-factor design was used wherein factor one was the input method (light pen or keyboard), factor two was the form of dialogue (D type 1 and 2). The order of presentation was systematically controlled. D type 1 was the form of dialogue not needing an explicit command to move a disc and not needing a draw command for pragmatic feedback.

D type 2 needed both commands. D type 1 with the light pen involved putting menu choices on the display (see Diagram 1) for moving and drawing. The dialogue then proceeded by pointing at MOVE; this would blink and a confirmation block appear next to it. On confirmation of selection, 'M' replaces the block and a stack label can be pointed at. On selecting a stack, the word 'FROM' appeared in place of the block. Then, the destination stack was pointed at.

On selection, the system noted the move but did not update the display accordingly. Pointing at and confirming the 'DRAW' option caused the screen to be updated. Measures of input times of the commands, (move, draw, restart, show me and quit) were grouped together (R1). Measures of time of other information input (stack label choices) were grouped together for analysis (R2).

D type 2 with the keyboard was such that input of the number of a disc to be moved caused an 'F' to be printed together with a prompt. On entry of another stack number, the disc was transferred on the display. No 'draw' command was necessary.

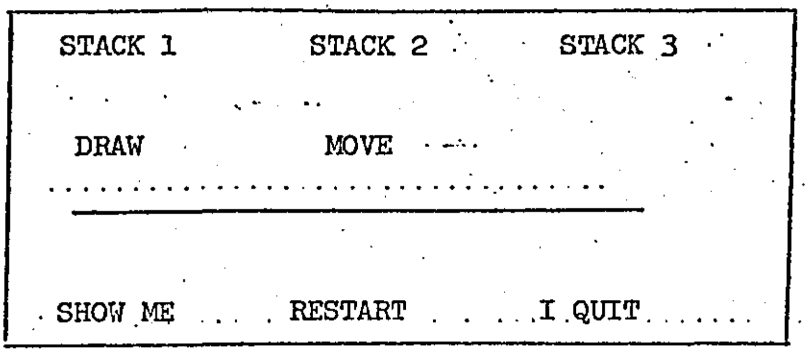


Diagram 1

Format of Screen for Light Pen Dialogue Type II

Measures of response time to the first entry of a stack number (implicit move), R for restart, S for show me and Q for quit were grouped together in R1. Measures of time for the entry of the second stack number were grouped together in R2 for analysis.

Measures of R1 and R2 for D1 with the light pen were therefore directly comparable with R1 and R2 for D2 with the keyboard. Similarly, R1 and R2 for D1 on the keyboard were directly comparable with the light pen, D2 version. Cross comparisons were assumed to be valid on the same basis as described in the previous experiment, i.e. the first entry of a new goal-directed transaction has a characteristic time distribution. System differences were not examined. An optimum system was used, (fast, reliable and consistent). Sixteen subjects were selected from the same group of twenty-four subjects used in the previous experiment. The subjects were selected such that they formed a balanced group on personality (neuroticism), experience and order of presentation of the conditions.

Table 1 shows the allocation of subjects to design cells. Complete replication of all combinations of input method and dialogue type was not necessary since the assigned order of presentation of input conditions for each subject could be used which was opposite to that received by the subject in the previous experiment. In order to do this, the task difficulty was similar (i.e. 4 discs) for each condition.

Subject No.	N	E	Order in Previous Expt.	Order in this Expt.
1	↓	↓	Light Pen First (L)	K
2	↓	↓	Keyboard " (K)	L
3	↓	↓	L	K
4	↓	↓	K	L
5	↓	↑	L	K
6	↓	↑	K	L
7	↓	↑	L	K
8	↓	↑	K	L
9	↑	↓	L	K
10	↑	↓	K	L
11	↑	↓	L	K
12	↑	↓	K	L
13	↑	↑	L	K
14	↑	↑	K	L
15	↑	↑	L	K
16	↑	↑	K	L

KEY :
Light Pen version in previous expt. had dialogue type 1 (no draw, move)
" " " " this " has " " 2 (with " ")
Keyboard Pen version in this expt had dialogue type 1 (no " ")
" " " " previous " has " " 2 (with " ")
↑ above ↓ below median score

Table 1

Allocation of Subjects to Experimental Conditions

3.2 Sessions

Each subject had one session wherein two trials were completed; each trial consisted of solving the 4 disc problem with the appropriate dialogue/input characteristics combination following training with 3 discs. The order of trials was determined by the experimental design. On completion of the session, the subject was asked which version was preferred.

The sessions were carried out approximately 1 year after those of the previous experiment. All subjects commented that they did not remember their previous sessions.

3.3 Other Details

The task was as described in the previous experiment and subject instructions were modified to take account of the dialogue differences.

The dependent variable was work production as defined and measured in the previous experiment. Similarly, treatment and analysis of the data was as for the previous experiment, except that not all the measures were tested for sensitivity to individual differences. Only those found to be sensitive in the previous experiment were tested. Furthermore, intercorrelations between different measures of work production were not examined.

4. Results

Table 2 shows those measures which were found to be sensitive to different factors at the 5% level using the F test on the ANOVA results presented in Appendix 3. In comparison with Table 8 of the previous chapter, there are more effects related to the measures of data information input. An inspection of the raw data shows that in this experiment the number of plans and recalls was very low and their distribution non-normal. Therefore, confidence in the results of this experiment particularly for comparison with previous results, is not great. This is a consequence of the small sample size arising from a low tas difficulty and a small number of subjects.

5. Discussion of Results

The results have been plotted on graphs which also include (where appropriate) comparative results from the previous experiment. The results are grouped into two categories; those which are affected by input methods (I) and dialogue differences (D), and those to D only.

5.1 Measures Sensitive to D x I

Graphs 1 to 3 show that of the measures, the mean time for inputting pre-programmed commands, the mean time for evaluation and the number of plans are sensitive to I x D. In all these measures there was no significant difference (t - test between means at the 5% level) between the means obtained by using sixteen rather than all twenty-four data points of the last experiment.

			Sensitive For		Chapter 6 Table 8
Measure			Input x Dialogue	Neuroticism x Experience x Order	Input x System
R1	P1	n	NT	NT	S
		t	I x D	NT	I,S
	P2	n	D	NT	I,S
		t_c	I x D	NT	I
	P3	n	I x D	NT	I
		t_c	NSD	NT	S
R2	P1	n	NT	NS	NT
		t_c	I x D	NS	NSD
	P2	n	I,D,D x I	NS	S
		t_c	D	NS	S
	P3	n	NSD	NS	S
		t_c	NSD	N x E (Light Pen)	S
Overall time to solve			D, I x D	All subjects preferred Keyboard because of dialogue	I,S

<u>Key</u> D - Dialogue type

I - Input Method

O - Order

N - Neuroticism

E - Experience

n - frequency of responses

t - mean time per response

t_c - corrected t for peak 1

P1,2,3 - Peaks 1,2,3 of response time distribution

R1 - response distribution of first user input after last pragmatic process

R2 - response distribution of other user inputs after R1

NT - Not tested in design

NS - No Significant Difference found at 5% level with F - test in design

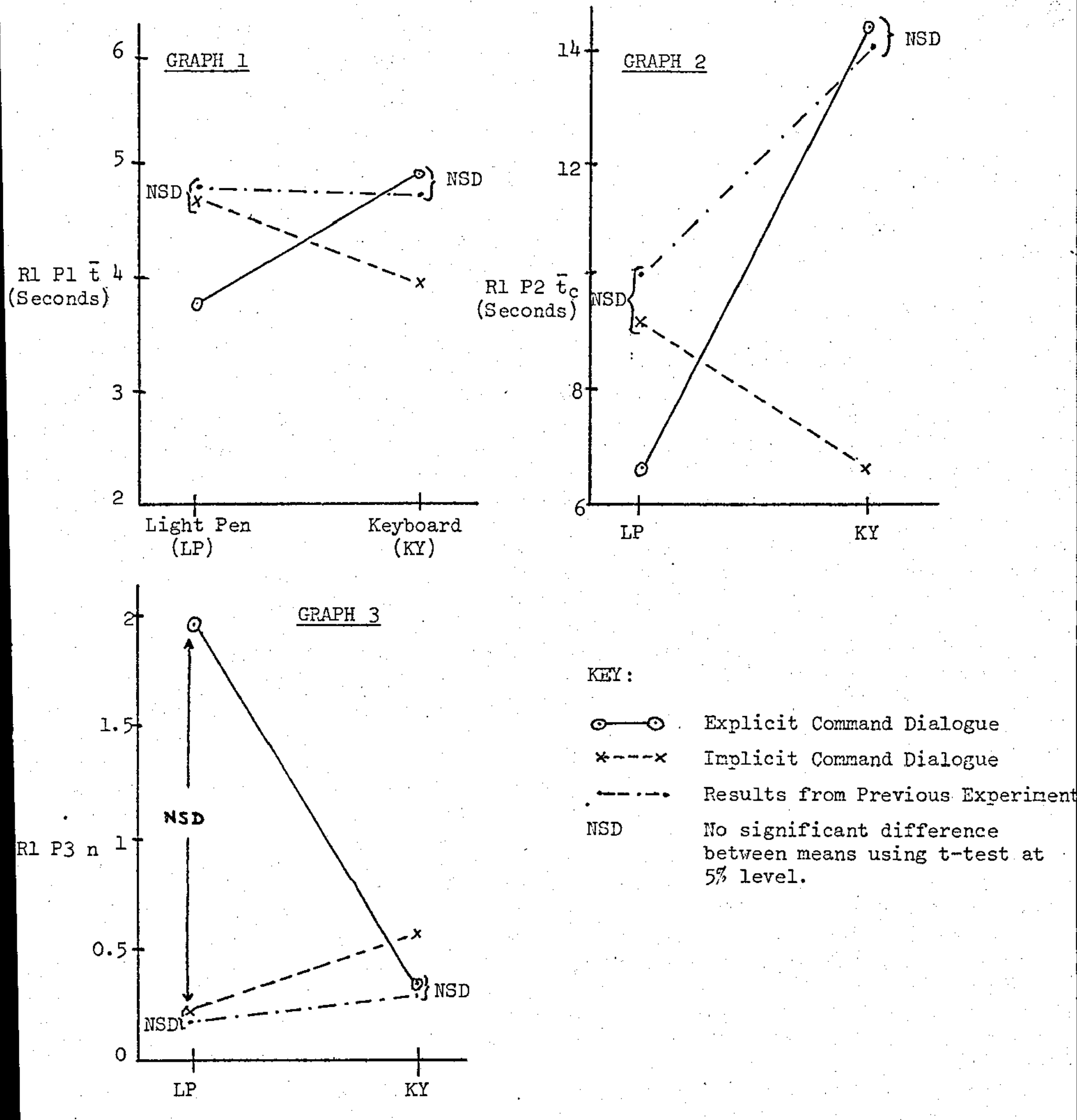
I x D - interactive effects found

Note: Individual differences tests were carried out on light pen data only.

Table 2

Significant Measures at 5% level in Dialogue Experiment

(See Tables 15 and 16 of Appendix 3)



MEASURES SENSITIVE TO INPUT CHARACTERSITICS
AND DIALOGUE DIFFERENCES

All the graphs show that the interactive effect between D and I was such that a particular combination altered the input times. Graph 2 shows that for D1 (implicit commands) the light pen was slower than the keyboard. The main explanation for this effect is that the movement time of the subject to activate the light pen was longer than for the keyboard. In comparison with the movement time difference, other differences (such as time to switch attention from display to keyboard) were insignificant overall. Observation of behaviour showed that most subjects (14/16) maintained both hands on the keyboard with fingers over the numbered keys (1 to 3) and the confirmation key during pre-programmed inputs. The amount of attention switching was minor. With dialogue type 2 (explicit commands) the light pen was faster than the keyboard. This was because attention had to be switched for every typed command from the display to the keyboard, then the command key located and typed. With the light pen, no change of attention was necessary and location and movement times were small (because of the display formatting of only 5 command options) compared with the keyboard (normal keyboard layout).

There was no overall difference in command input times due to D or I independently. However, a t-test was used to examine differences between the mean command time for each input characteristic within dialogue. There was a significant (5%) difference between the means for both the light pen (t = 2.15) and keyboard (t = 2.13). For the light pen, the mean command response time was about 1s faster for explicit commands than for implicit commands. For the keyboard, the converse was the case.

The interactive effect is also found in the data input times shown in Graph 7. With the implicit dialogue and the keyboard, it was possible for subjects to minimise movement and attention shifts by keeping one hand over the data input keys. Hence, input times are faster than those for the light pen which required movement across the screen. With the implicit dialogue, time had to be taken with the keyboard to change attention to the keyboard, locate and move towards the key. This time was longer than that required to simply move the light pen to pick the explicit command.

The 'evaluation' times (shown in Graph 2) indicate a similar trend to the 'pre-programmed' command entry time, and the differences in times are amplified. With explicit commands, the mean evaluation time is considerably greater for the keyboard than the light pen but the converse is true when implicit commands are used. An over-simplified explanation is that the evaluation time is an error correction of input of messages. The trends are similar to those in Graph 1. Thus, with the keyboard and explicit commands, the time to carry out a solution step is greater than for the light pen. Hence, in evaluating progress and working out the next series of moves, more time may be spent ensuring that an attempted solution is correct with the keyboard than with the light pen. Graph 5 shows that the number of evaluations (rather than the time per evaluation) depends on the dialogue type as suggested in the previous Chapter. This may arise because of the extra cognitive strain imposed by the need for updating the screen contents so that it shows the current state of the problem.

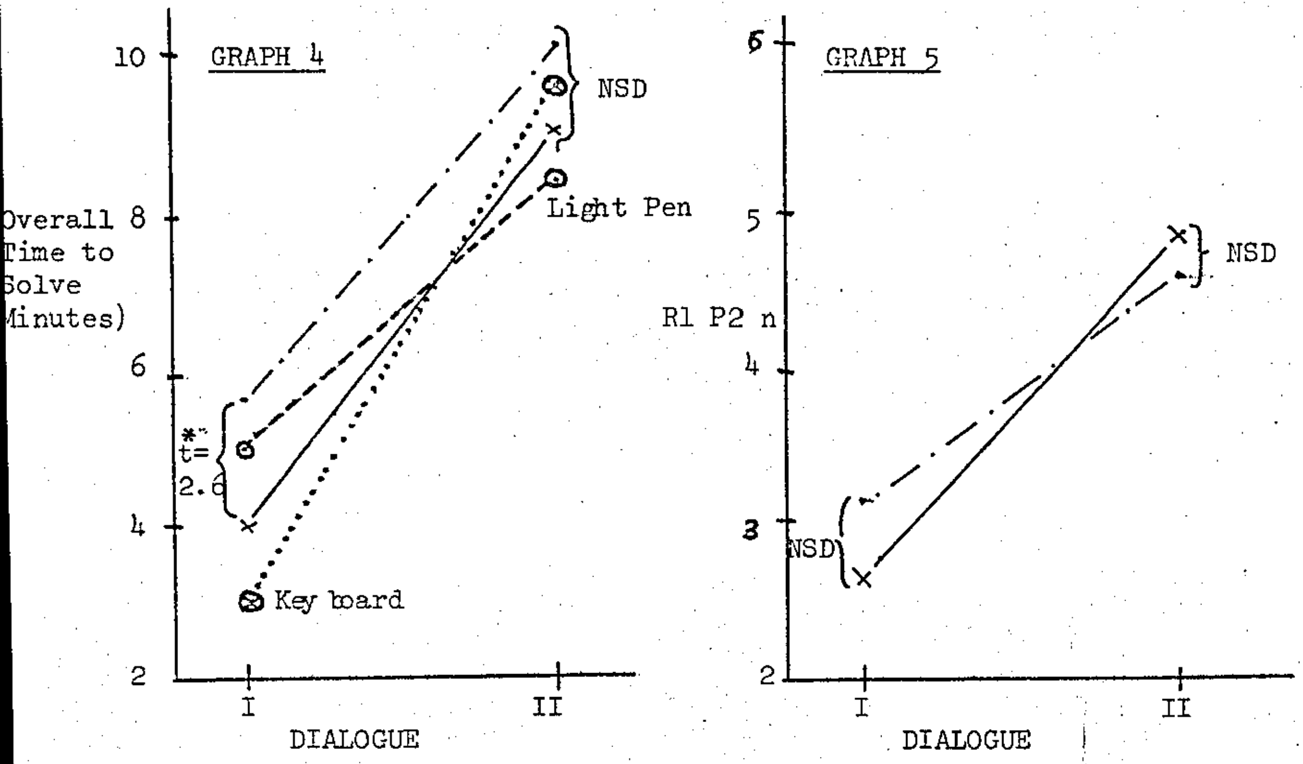
Graph 3 shows that dialogue differences need to be taken into account when the input methods are compared (as in the previous Chapter) since they affect the planning frequency measured in command input times.

With explicit commands the number of plans was greater for the light pen than the keyboard; with implicit commands, the converse was true. This implies that the frequency of planning varies inversely as a function of the time taken to externally test solutions.

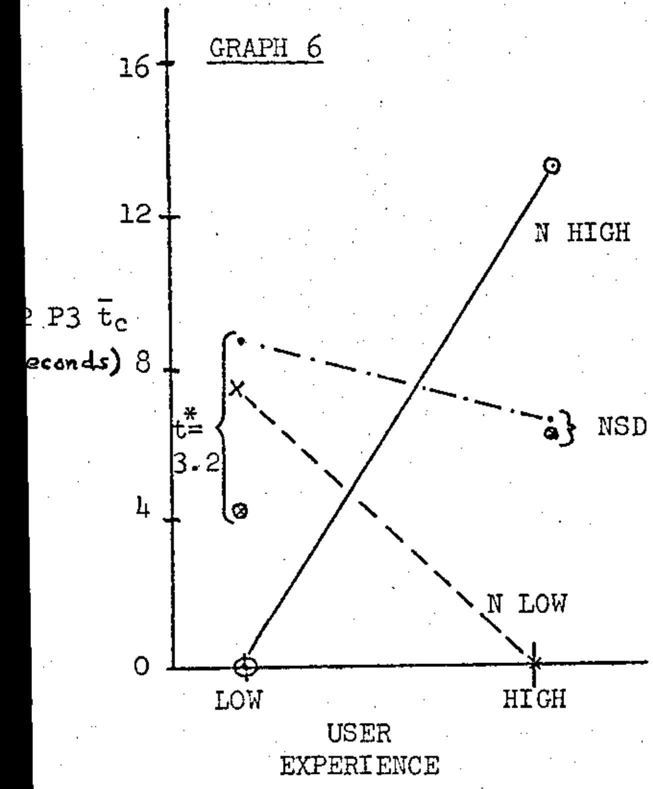
5.2 Measures Sensitive to Dialogue Differences

Graphs 4, 5, 7, 8 and 9 show measures sensitive to dialogue differences. Graphs 5 and 7 have been discussed in the previous sections. Conclusions drawn in the following discussion concerning Graphs 8 and 9 should be treated with caution since the sample size was small and extreme values are involved.

Graphs 8 and 9 concern the number of and time of evaluations as measured in data input. Graph 8 shows that the number of evaluations with explicit commands was greater than for implicit commands and Graph 9 shows that the mean time per evaluation also increased. The number of evaluations increased because of the light pen interaction with dialogue. Observation of subjects indicated that a main reason for this was that errors were made in using the light pen for selecting the second (destination) label. Because the light pen was unswitched, spurious hits on other option labels for explicit commands caused them to blink and so grab the subjects attention. In particular instances, the confirmation block would also be spuriously hit and so cause the subject to not only have to recall the move currently being carried out, but to draw the current state of the problem as well. Both extra actions increased the number of evaluations above that



MEASURES SENSITIVE TO DIALOGUE DIFFERENCES



MEASURE SENSITIVE TO

INDIVIDUAL DIFFERENCES

KEY:

I - Dialogue with implicit Commands

II - Dialogue with explicit Commands

-·-·-· Results from previous experiment

* - Significant difference between means at 5% level using t-test

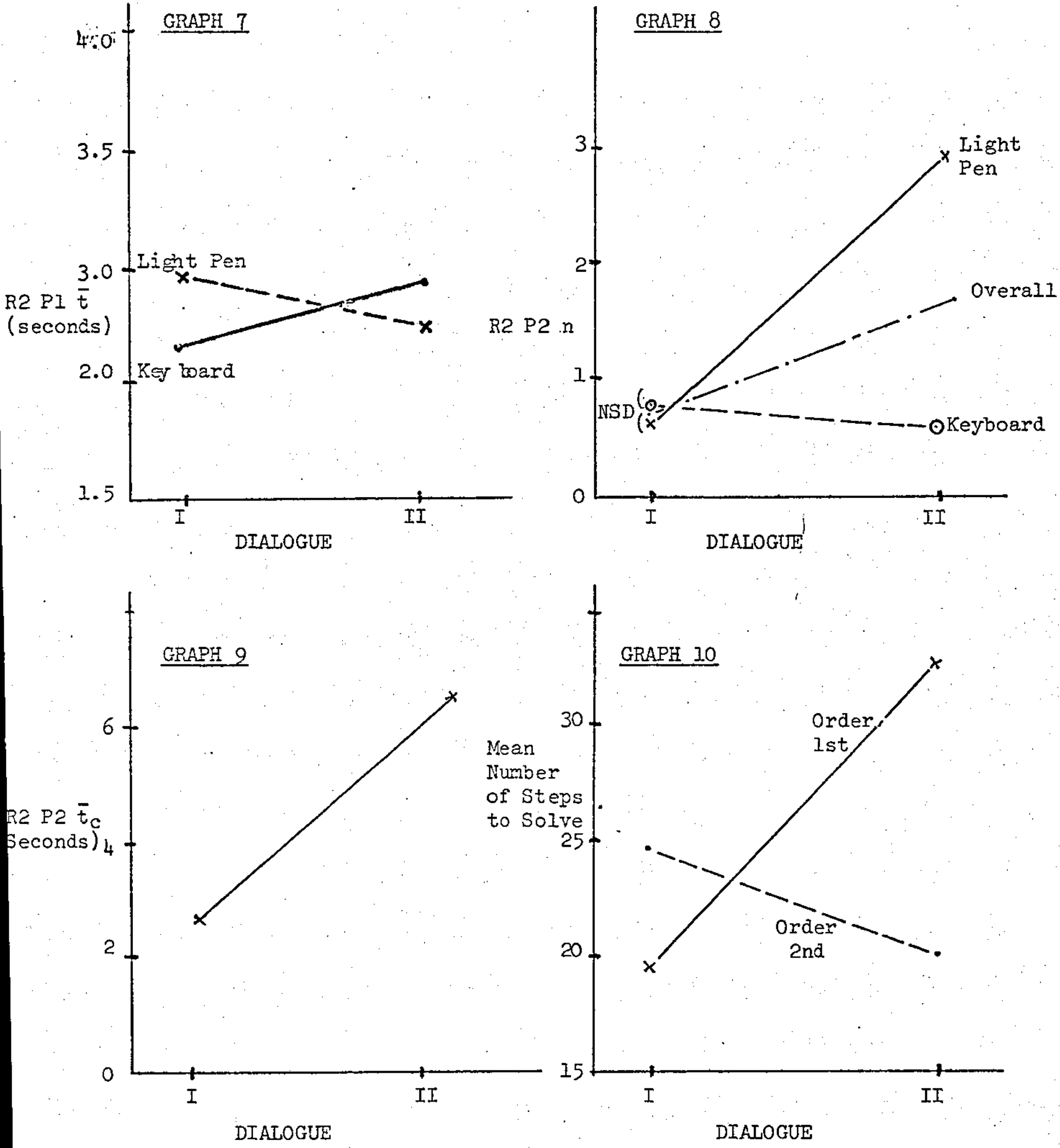
NSD - No significant difference

N - Neuroticism

High - Above medium score for group

Low - Below medium score for group

⊗ - Group Mean of all subjects



MEASURES SENSITIVE TO "DIALOGUE"

Key:
I - Implicit command for Drawing, moving
II - Explicit command for Drawing, moving
NSD - No significant difference (t-test @ 5%)

to be expected from, for example, the less error-prone keyboard. This did not occur with the implicit command dialogue because the extra command menu options were not displayed and so did not contribute to error proneness of the light pen (cf. Expt. 1).

The longer time for evaluation arises because, in the light pen error case, the time to work out what action to take added to an interfered with the time taken to correct input errors.

Graph 4 shows that differences in the overall time to solve the problem depended on the interaction between the input methods and the type of dialogue. The provision of implicit commands reduced the solution time from ~9 to ~5 minutes on average. The fast input times which Graphs 7 and 1 show for data and command input are reflected in the overall times for solution.

5.3 Other Measures

Graph 6 shows the interactive effects of neuroticism and experience on the mean planning time as measured in the data entry times. However, it would be misleading to base a discussion on this graph since the number of samples per cell is only two.

Therefore, it is presented here as a possible item for further research.

Graph 10 shows the interactive effects of dialogue and order of presentation on the number of steps taken to solve the problem. It shows that there is a difference in the ease of use of the dialogue such that more solution steps were taken when the explicit dialogue was used first then after the problem had been attempted using the implicit command dialogue. A possible reason for this is that, with explicit commands, subjects made, on average, 2 logical moves before updating the screen. When meeting the problem for the first time with this dialogue, subjects had to learn to do this with such a frequency. In the course of this learning, errors were made and logical solution steps had to be repeated. After having solved the problem using the implicit command dialogue, the effect of learning the strategy with the explicit command dialogue was not so great.

When the implicit command dialogue was used after the explicit dialogue, the relative ease of use was such that redundant testing of solutions was attempted compared with when the implicit dialogue was used first.

Finally, all subjects preferred the keyboard version and when asked for reasons, said that they found the dialogue for the light pen (explicit commands) cumbersome. The preference order agrees with the rank order of the number of plans (rank 1 = lowest) as measured in the command response time data. This is as observed in the previous experiment.

6. Conclusions

Within the constraints of this investigation the conclusion is that provision of distal pragmatic feedback at each step considerably relieves the cognitive strain of the problem solver. The effects of not providing this to the problem solver depends on the input method according to how it loads cognitive processes.

Input methods with implicit input relieves cognitive strain but the effects of this on problem solving behaviour is not as great as that due to sub-optimal distal feedback.

6.1 Experimental Hypothesis

The hypothesis that problem solving does not depend on input and system differences is rejected.

6.2 Implications for the Model

The provision of adequate distal feedback is essential to the model. The model should be extended to include the effects of timing and control of distal feedback.

CHAPTER 8

FURTHER ANALYSIS OF EXPERIMENTAL DATA

CONTENTS

TABLES

GRAPHS

1. Introduction

The purpose of this chapter is to present and briefly discuss further analyses of the experimental results. These are used and referred to in the discussion chapters which follow. Two sets of analyses are presented. The first set is directed towards comparative analysis of the results in the thesis so as to aid their interpretation; the second analysis is an example of the use of the results by system designers.

2. Comparative Analyses

The selection of the particular comparative analysis is based on the similarity of both the input method and the task being carried out. Two task levels have been used in this thesis: simple menu selection and problem solving.

2.1 Menu Selection and Input Method

Experiments 1 and 2 were comparative studies of alternative input methods for menu selection. Comparisons between the results across these experiments may be carried out using the time-to-select an item as the measure for comparisons. The selection time is the time to strike the first key or point at an item following a cue.

Table 1 shows the comparative results which indicate that for both speed and accuracy there was no significant difference between the keyboard techniques, but that there was a difference between light pen and keyboard. The light pen was considerably

Technique = / Measure	A Numeric Identification (1-9)	B Alpha ID (A-I)	C Light Pen
No of correct responses	118	99	73
Mean time to select (Seconds)	3.487	3.531	2.069
Standard deviation of time to Select (S)	1.389	0.760	0.988
Error Rate (%)	0.67	1.67	27.00

TABLE 1a

Comparative Data for Light Pen and Keyboard Techniques

		A	B	C
INPUT METHOD (SEE TABLE ABOVE)	A	▨	0.297	6.967*
	B	0.732	▨	9.339*
	C	0.379	0.372	▨

t-values

c'-values

* Significant at 1% level

TABLE 1b

Welch test on difference between Means

more error prone (possibly because it was unswitched) but was used faster than the keyboard regardless of the keyboard input method. Although the differences between the keyboard input methods was not significant, it is proposed that when the number of items in a menu list is greater than 9 this would not be the case. This is because after 9 but up to 26 items, 2 keys must be pressed to identify a menu item (e.g. '25') whereas only one is necessary with the alphabetic identifier (e.g. 'Y'). Thus even in this simple task, the generality of results is limited, although this may be extended by identifying and accounting for the sources of variations in the data.

Table 2 shows the estimated relative contributions to the variability between and within the results of experiments 1 and 2. These are given in seconds and as percentages of the average maximum select and confirm time. Inspection of Table 2 shows that in the simple menu selection task the time taken for list searching is relatively independent of differences between input methods. Also, for the keyboard input methods, other 'constants' may be identified such as the average time to select an item and the time to change attention from the keyboard to the display and vice-versa. Hence, performance of particular component processes within particular constraints may be generalisable.

The conclusion of this section is that in simple menu selection tasks the range of variation in performance due to different input methods is large and depends on a number of contributory factors.

INPUT METHOD			TASK COMPONENTS			SUBJECT COMPONENTS		
DEVICE	WAY OF USE	RANGE OF CONTRIBUTION OF INPUT METHOD AS % OF MAXIMUM Entry TIME	LIST SEARCH	DEVICE ACTIVATION		DIVISION OF ATTENTION	TASK EXPERIENCE	SYSTEM EXPERIENCE
				SELECT TIME	ENTRY TIME			
LIGHT PEN	POINT	$\frac{2}{7.54}\% \rightarrow \frac{2.3}{7.54}\%$ = 26% → 30%	2→2.33(P)s		2.83(P) (40% →20% errors)	–	Pointing Technique -0.75s	–
KEYBOARD	ALPHA-KEY	44.8%	2.33s	–	0.85s (1.67% errors)	+0.2s	–	–
KEYBOARD	DIGIT-KEY	53% → 61%	2.33s	0.85s	0.85 (0.67% errors)	+0.2s	Technique (T) -0.2s	Penalty for errors (E) +2 x 0.2s
KEYBOARD	STEP	48% → 100%	2.33s	0.85s	N x 0.44(P)	+0.2s(P)	Parallel (P) Device Activation -0.1s	Penalty for errors + 2 x 0.2s
RANGE OF CONTRIBUTION IN SECONDS		2s→7.34s	2 → 2.33s	0.85s	0.85s → N x 0.44s (0.67 → 1.67% errors) 1< N < 9	0.1 → 0.2s	-0.7s→ -0.2s	2 x 0.2s
RANGE OF CONTRIBUTION AS A % OF MAXIMUM RANGE OF ENTRY TIMES		26% → 100%	26 → 30%	11.2%	5.8 → 52%	0 → 2.6%	-9.3 → -2.6%	5.3%

Notes: Times are approximate to about 5 %
N = number of steps down the list

The figures are added across rows to give total time in seconds for entering a menu item.
Negative values are shown where appropriate.

TABLE 2

Response time and Error contributions of Different Factors in Menu Selection

These are to do with the interactions between component processes which make up the task and depend on the input method. The component processes may be placed into task and subject categories. Task components are list searching and device activation; subject components are experience and division of attention. For the particular experiments of this thesis, relatively constant times were assumed for particular processes. These were list searching, striking the entry key, changing attention from display to keyboard and vice-versa. These are postulated to be generalisable results for similar menu selection tasks and input methods. However, different tasks and different input methods affect the generalisation of the results.

2.2 Problem Solving and Input Method

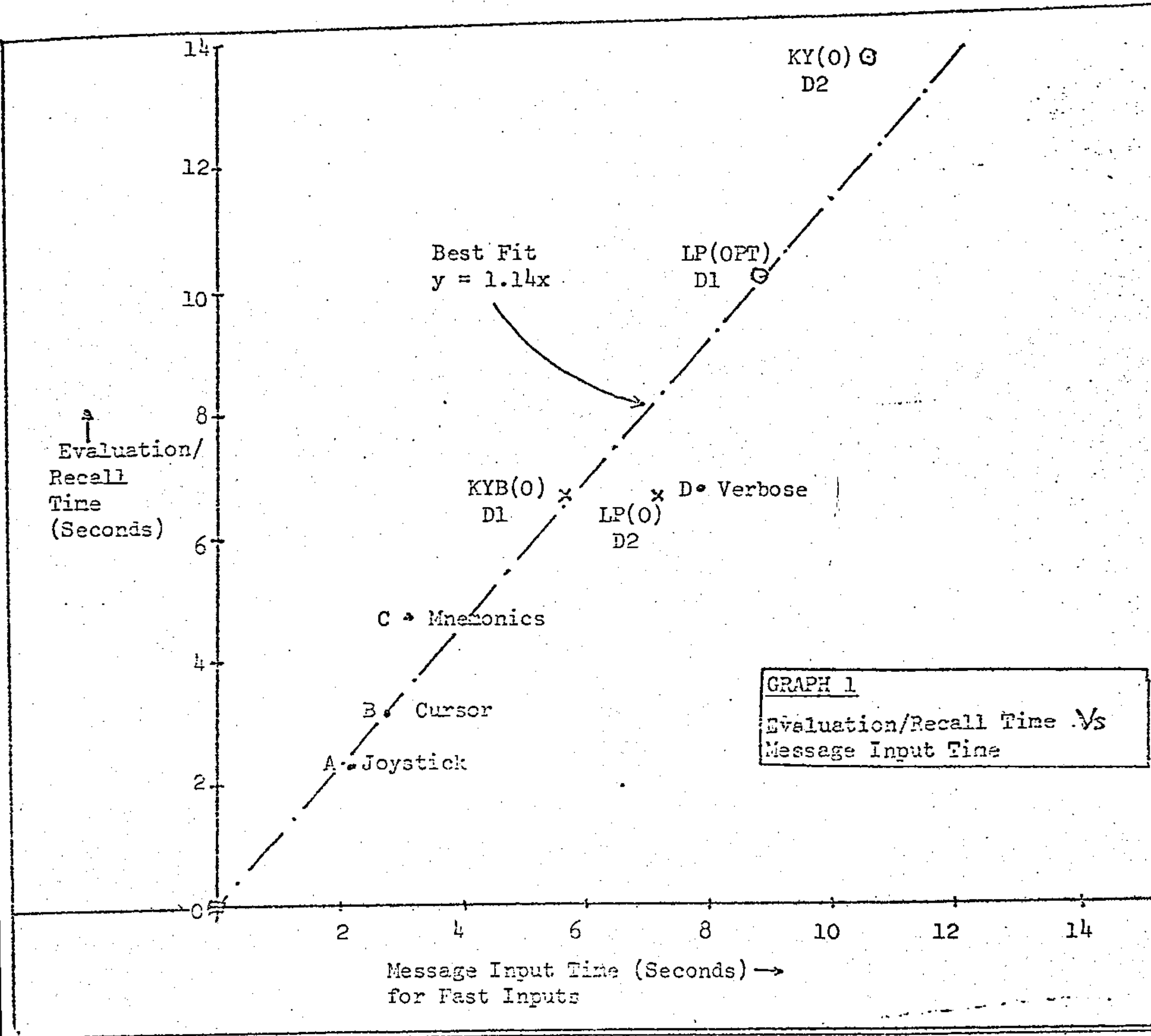
Experiments 3, 4 and 5 investigated alternative input methods used for problem solving. They used a wide range of tasks and input methods and, consequently, there is little that is generalisable from experiment to experiment.

The main general observation is that the distribution of input times for messages in each experiment had 3 peaks.

The first peak was inferred to be due to 'pre-programmed' or obvious sequences of input; the second was called 'evaluation/recall' and the third called 'planning'. While the first and third peaks were reliably interpreted using verbal protocol, there is some ambiguity about Peak 2.

Graph 1 shows the relationship between evaluation/recall times (peak 2) and message input time (at the fastest rate) across experiments 3, 4 and 5. An interpretation of the assumed simple linear relationship is that the delay in input time called 'evaluation/recall' may arise from time needed for error correction during message input with that input method (as suggested in discussion of experiment 3). Thus, if no error is made on message input, no 'evaluation/recall' time is necessary; hence the line is assumed to pass through the origin. The slope of the line is slightly larger than 1 which may indicate that more effort is needed to correct input errors with input methods which need longer times for message input than for input methods which do not. This effort may hypothetically be used for making sure information is not forgotten during error correction of input messages. The hypothesis is suggested by the observation that the interaction of dialogue (distal feedback and redundancy of input) and input method strongly affected the amount of information that was held in short-term memory for input in experiment 4.

The conclusion is that the second peak (called 'Evaluation/Recall' times) should be interpreted in discussion as being due to error correction of input.



GRAPH 1

Evaluation/Recall Time Vs Message Input Time

Key: A→D Results of Experiment 3

LP(O) D1 } Results of Experiment 4 Optimum System
KYB(O) D2

LP(O) D2 } Results of Experiment 5 Optimum System
KYB(O) D1

Note: For D1, Message Input Time = $(R1\ P1\bar{t}) + 2x(R2\ P1\bar{t})$
D2, Message Input Time = $(R1\ P1\bar{t}) + (R2\ P1\bar{t})$

3. An Analysis of Some Results for Real Time Systems Development

There are a number of possible ways for the results of this thesis to be used by systems designers. The purpose of this section is to describe one particular way which is aimed at aiding the designer of a real time multi-user command system. In these systems the users and computer are co-operating in sharing resources. The designer has the general problem of deciding how resources should be shared at any time.

3.1 Sharing Resources Between Users

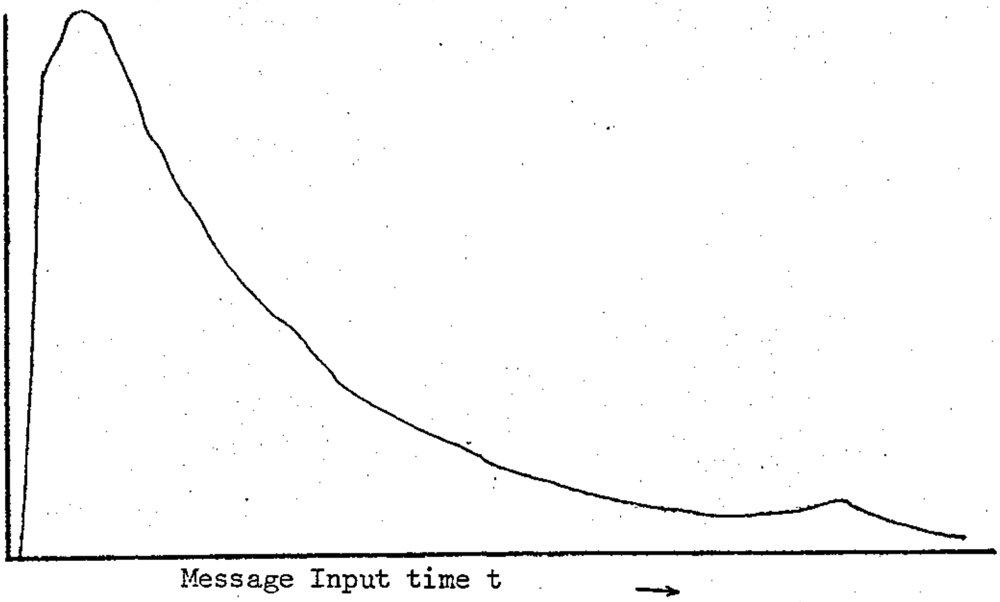
How resources should be shared depends on many factors and these have been discussed in many places (e.g. Nickerson et al. (1968)). Among the most important resources suggested by Carbonell et al (1968) is the time available to users. Hence the general problem is to obtain dynamic characteristics that are acceptable and convenient to users.

However, user behaviour and computer system performance are interdependent. That is, system response time depends on the number of users and the operations they are carrying out and, possibly, vice-versa. The use of particular hardware configurations (e.g. intelligent terminals) may allow the optimisation of dynamic characteristics within a band defined by the overall system, but the main constraints apply.

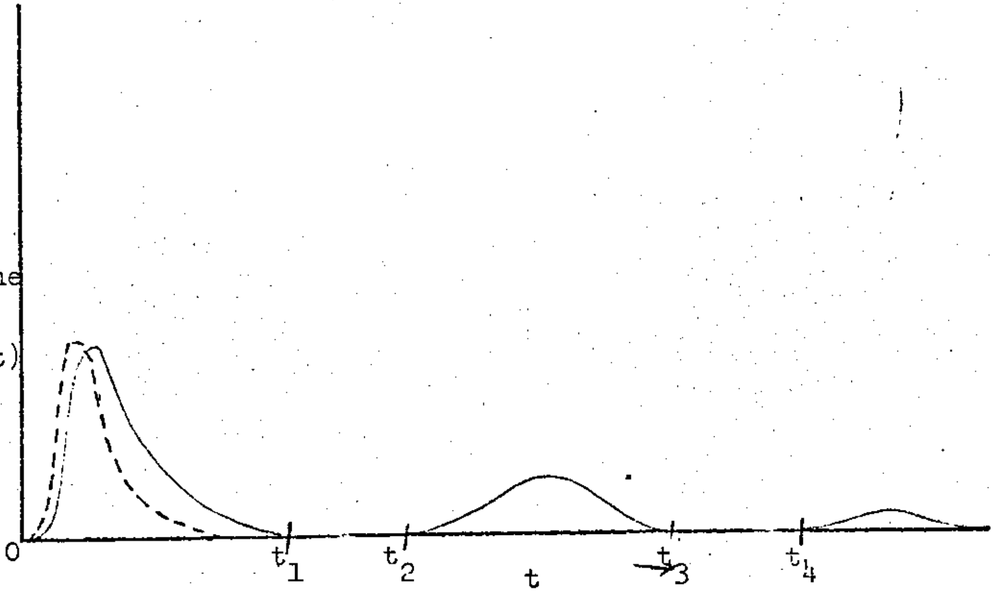
In central computer time sharing systems, research was suggested by Simon (1966) for the design of time scheduling algorithms based on the users point of view. In that paper seven classes of user were identified based on tasks carried out with systems. The class to which the research of this thesis is pertinent is real time operation at human conversational rates. Simon's early suggestion was reinforced by Sackman (1970) in a proposed study on human engineered scheduling algorithms. The basic idea is that these algorithms take into account the dynamic behaviour of users of the system in some way. This behaviour is partially determined by the input method as indicated in this thesis, and is characterised by the distribution of input times for messages sent to the system by a user(s).

Graph 2 shows data which was collected on message input times in the CERN interactive graphics system, (Yule (1972)) using a light pen and keyboard for problem solving. It is typical of data found in many systems (see Sackman (1970)) and there is no discrimination in the data between users, task or other characteristics. The research of this thesis has made such discriminations and collected data on message input times.

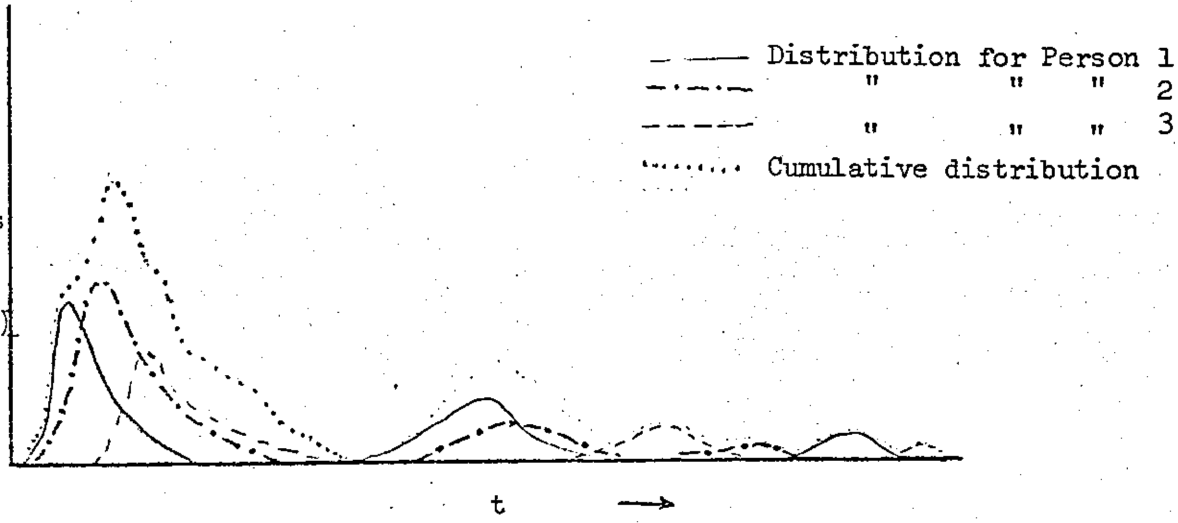
A typical distribution of input times for messages in this thesis is shown in Graph 3. The three peaks may be used in a scheduling algorithm, for example, in the following way. The times between t_1 and t_2, and t_3 and t_4, may be taken as 'dead' user times wherein the system does work for other users. In order to do this, the system needs to be aware of particular users and usage.



raph 2
istribution
ound in
ime-sharing
ystems for
essage input



aph 3
equency
stribution
und in
esis for one
rson



aph 4
ationship
tween Graphs
nd 3 above

However, the cost of identifying every user and adapting to his behaviour may be prohibitive where there are many users and wide variety in task and input characteristics. Therefore a compromise may be to group data together (as shown in Graph 4) so as to minimise the cost yet retain some adaptive flexibility which can be used to enhance system performance. In this case, grouping users to obtain grouped data may only be achieved after checking that different users input times and type of usage are statistically compatable rather than interfere with each other. This process may be far less costly than maintaining active profiles of each user's data.

As one example of how human engineering may aid total system performance in this way, the results of experiment 3 (maze problem solving) were used in a method defined by Yule (1972). (I am indebted to Dr. Yule for giving permission to use his results and method in this thesis.)

3.2 Yules Method and the Maze Experiment

In an interactive graphic system, a program uses a small amount of central processing time in comparison to the real time that the job remains in the computer. All the time the program is running it uses some central memory and affects system throughput. Therefore, the job is 'rolled out' into back up store until an input message is received from the user. However, if this is done too soon, and a message comes in just as the program has been rolled out, the program must be rolled back in immediately; this implies higher system cost because core use and some channel usage has been lost. The question arises therefore, "What is the best time to wait before rolling out a program?".

Yule attempts to answer the question by calculating a quantity, C_s, which is the reduced throughput of jobs per second because of a program being in core. To simplify the theory, C_s is assumed to be independent of program size. Let Cd be the reduction in throughput per second caused by taking a channel away from the system and assume that when a program is being rolled, the cost is C_s + Cd, (only if C_s and Cd are not near to 1).

In developing a cost function, C (T), various cases arise according to whether the best waiting time (T) before a job is rolled out is less or greater than a user's input time (t).

If $t < T$ then the job will not be rolled out and the cost C_1 will be $C_1 = \frac{C_s\ t}{t} + C_s$

If $T < t \leq T + q$ (where q is the time required to roll out the program) then the program will be rolled out only to find an input message is causing it to be rolled back in. The cost for this condition C_2 is $C_2 = (C_5\ T + 2\ (Cd + C_5).q)/(T + 2q)$.

If $T + q < t \leq T + 2q$, then the message arrives when a new program is just being rolled in and the cost C_3 is

$$C_3 = (C_s\ .\ T + 4\ (Cd + C_s).q)/(T + 4q)$$

Finally, if $t > T + 2q$, the cost C_4 is

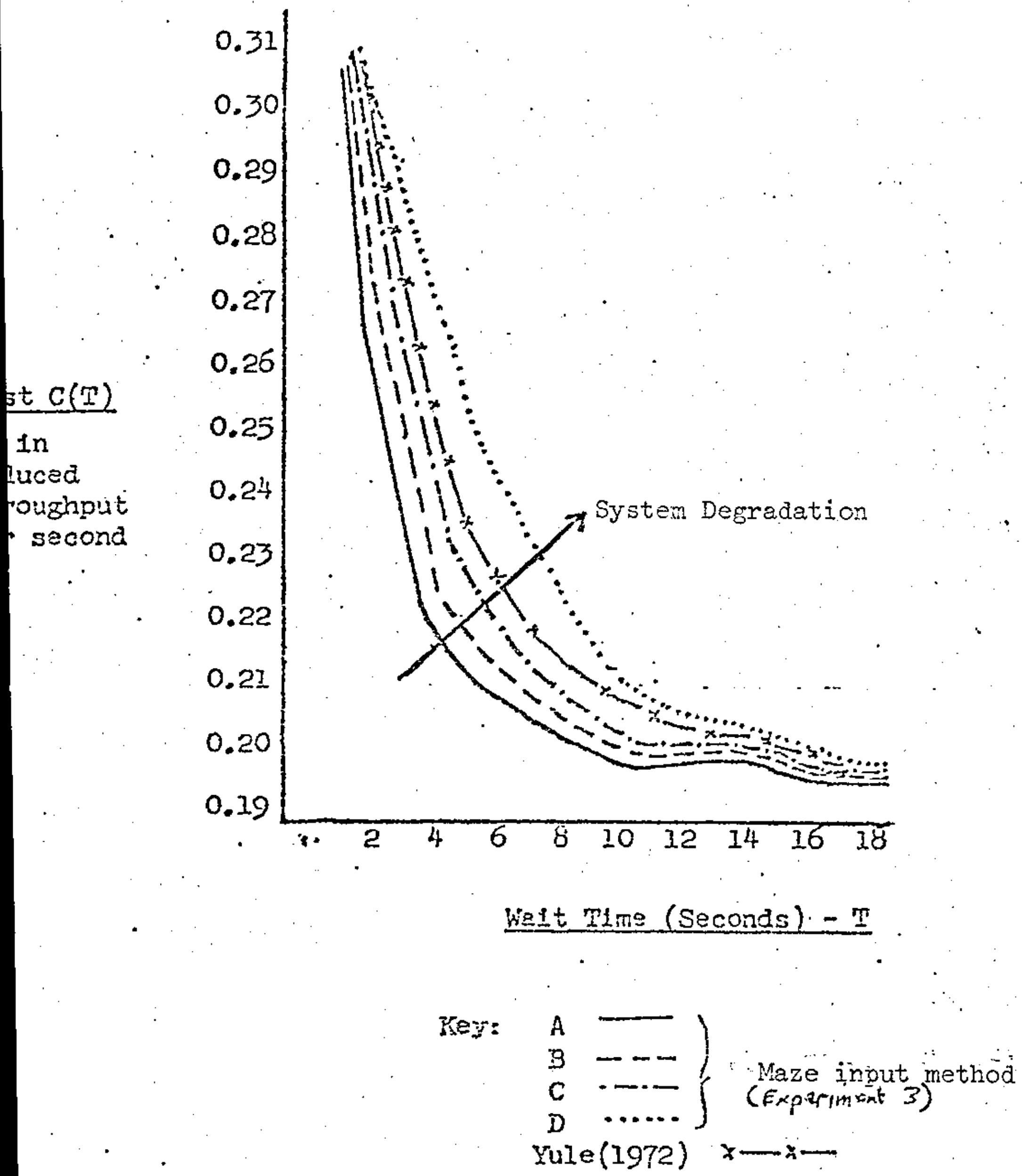
$$C_4 = (C_s.T + 4\ (Cd + C_s).q)/(t + 2q)$$

Thus the total cost C(T) in terms of the probability function $\emptyset\ (t)$ that an input message will be present at time t is

$$C(T) = \int_{o}^{T} C_1.\emptyset\ (t)\ dt + \int_{T}^{T+q} C_2.\emptyset(t)dt + \int_{T+q}^{T+2q} C_3.\emptyset(t)dt + \int_{T+q}^{\infty} C_4.\emptyset(t)dt$$

CERN values for the constants are $C_s = 0.2$, $C_d = 0.25$ for 25 K word programs and, for this, $q \doteqdot 1$ second.

Thus, Yule has derived an analytical function which takes into account the dynamic behaviour of users in order to optimise the use of the computers resources. In using this function with fixed values in this way, the underlying assumption is that the dynamic behaviour (as reflected by $\emptyset(t)$) does not change with waiting time. This is likely to be true only within certain limits. Assuming that this is so, Yule has answered the question of 'best waiting time' by plotting the cost function against waiting time for his $\emptyset(t)$ data. One curve on Graph 5 shows the results and that the reduced throughput of jobs per second increases dramatically with a wait time of about 7 or less seconds. Since the wait time affects the users response time, then 7 seconds is an acceptable minimum which must be weighed against user acceptability.



Graph 5 Cost curves For all Subjects Averaged together on Mazes 1-4

In order to see how the human engineered algorithm may aid user and system together, the message probability distribution of the maze experiment was taken for each input method averaged across all subjects. With the assumption that the CERN values are appropriate, the resulting cost functions are shown in four curves of Graph 5. For all the inputs the minimum wait time before the cost increases dramatically is about 4-10 seconds. But there is less penalty for reducing this time in a trade-off with other factors (e.g. acceptability) with A than with D. Thus, not only did subjects find A a more acceptable input method than D, but, in a time-sharing system like CERN, this may be reinforced by an associated improvement in system response time and system throughput. This is a question for future research.

4. Conclusion

This chapter has presented further analyses of the results to review the generality of some of the results and also their relevance to the design of real time systems. The chapter concentrated on a small number of analyses which are pertinent to the following discussion chapters and which complement the analyses in previous chapters.

CHAPTER 9

DISCUSSION OF GENERAL HYPOTHESES

CONTENTS

1. Introduction

The purpose of this chapter is to review the experimental results in relation to the general hypotheses presented in Chapter 2. Some general qualifications should first be noted.

(i) The particular nomenclature is as defined in Chapter 2. Terms such as 'problem solving' and 'input method' are used for particular meanings and should not be interpreted otherwise.

(ii) In general, the reader should refer to the summary or main body of the chapter describing a particular experiment for quantitative validation of statements made in this chapter concerning experimental results.

(iii) The generality of the results is of limited scope. Chapanis (1967) emphasises that caution is necessary in generalising results from laboratory experiments. Since the experiments in this thesis are largely exploratory and deal with a wide range of variables, this is particularly true of the present results. Therefore, the reader should consider generalisation and proposed explanations as tentative rather than conclusive.

(iv) Where the discussion refers to comparisons of results from different experiments and no supporting data are given, the previous chapter on further analysis should be referred to.

(v) Unless otherwise stated, references to input times for information applies (in appropriate circumstances) to pragmatic information and so the times are for inputting full messages.

In the following discussion, each general hypothesis is considered as if it were independent of the others.

2. General Hypothesis 1

'Problem solving processes are affected by the input method required by the computer'.

Experimental hypothesis 3 (Chapter 5) was directly concerned with this general hypothesis. The hypothesis that problem solving does not depend on the input method was rejected (in the Maze experiment).

In the experiment three rates of transfer of information from man to computer were identified: a fast rate wherein a predefined sequence was being followed; a slower rate (referred to as evaluation or recall) wherein transfer was delayed; a very slow rate called planning. It was found that the number of steps in an attempted solution decreased as the time needed to input information increased.

This was true despite general differences between people and differences between their solution strategies (e.g. means-end versus goal oriented strategy). The time to make a step in Experiment 3 was the same as the time to input a single message to the computer at the fastest rate. Therefore, the result may generalise into 'the number of steps in an attempted solution decreases as the time needed to make a step increases'. This is a question for future research.

To summarise it is tentatively concluded that: regardless of differences between people and between problem solving strategies in a problem used in this thesis, problem solving processes were affected by the input method for the computer such that the number of steps in an attempted solution decreased as the time needed to make a step increased. The input method determined the time needed to make a step.

3. General Hypothesis 2

'The input method determines the transfer of information between man and computer depending on the interactions between the characteristics of computers, people and problems'.

Unlike the first general hypothesis, this one was not directly tested via a single experimental hypothesis. All the experimental results contribute to the following discussion. For ease of reference, the experiments will be referred to as experiments 1 to 5, i.e. in their order of appearance in the thesis.

3.1 Input Method in Relation to Differences Between People, Tasks and Computer Characteristics

Experiment 1 showed that the range of variation in speed and accuracy of information transfer which arose from differences between input devices is comparable with that due to the differences between people. Further, the variation in speed of information transfer for a light pen due to changes in display characteristics were comparable with the variation due to the differences between people, although this was not true for accuracy. The general conclusion was that subjects accepted the error-proneness of the input method because the emphasis of the experiment was on speed rather than accuracy. That is, the light pen was used as fast as possible for selection and the consequent error rate was accepted by the subject who did not adapt his behaviour since there was no penalty for error.

The results showed that, if the need to adapt was present, then the characteristics of the light pen allowed a wider range than the keyboard. The results agree with those of English et al. (1967) who found that adaptation to input characteristics occurred only with particular input devices which exhibited large error rates when used quickly.

In the problem solving of Experiment 4, the information input times with the light pen were a function of individual differences; this was not true for the keyboard. That is, the effects of individual differences in input rates were shown especially with particular input methods and devices which were amenable to a time/error trade off.

Experiment 1 showed that large individual differences existed in information input times, even in a simple task. Experiment 2 attempted to take some of these into account by postulating that a mixture of general computer experience and specific experience with the task was important. It was suggested from the results that, on a simple menu selection task, experience was used to generate motor plan 'subroutines' (see Newell and Simon(1972), p.803), which used an internalised process for locating keys rather than using visual scanning. This allowed subjects to maximise their input rates by 'overlapping' processes (Long, (1976)). The degree to which such plans were made was found by Long to be a function of the input method which affected the compatability between the stimulus and the required response.

In Experiment 2, the input device remained the same and two ways of using it were compared. The effects of experience on information input times weremuch less than the effects due to differences between ways of using the input device for selecting and entering a menu item. This was because the techniques differed largely in terms of, for example, penalty for error and the number of keypresses needed to select an item.

3.1.1 Individual Differences

Although Experiment 3 was designed to examine the effects of different input methods and problem solving, data from that experiment was used to test hypotheses related to individual differences based on work by Eysenck.

Eysenck (1952) suggested that the personality dimensions of neuroticism-stability and extraversion-introversion may be significant parameters in understanding individual differences. In his book, Eysenck presents many hypotheses predicting general behaviour patterns of people. For example, he provides evidence for the view that neurotics under emotional stress exhibit disorganised motor responses. In a report of studies made with pilots, Eysenck refers to two types of error due to motor disorganisation. These were errors of overaction and errors of inertia. Overaction refers to larger control movements than required and inertia refers to under-activity. Taking Eysenck's view, it may be expected that neurotics under stress in problem solving from using error prone input methods would tend to make errors through motor disorganisation. This hypothesis was tested in Experiment 3 as a subsidiary hypothesis.

The results were inconclusive but did support the view that the personality trait of neuroticism may be a factor in some way related to the variety in input times that can arise from differences between people.

With this in mind, Experiment 4 was designed to take into account the personality trait of neuroticism as well as the general and specific experience of individuals.

The results of Experiment 4 showed that neuroticism and general computer experience interactively affected the rate at which information was input to the computer using the light pen with an unreliable and slow system on a difficult task. That is, the effect of personality was only brought out under stress conditions which were exacerbated by an error prone input device.

Further, these effects were shown only for 'data' entry; that is, for entry of the last part of a message rather than the first part. (The first part being 'command' entry. Together, command and data make up a message which defines a problem solving step.)

The effects were such that stable, experienced subjects were inputting data faster than stable inexperienced subjects, but neurotic (unstable) experienced subjects were slower than neurotic inexperienced subjects. These effects may be hypothetically explained if an interpretation of 'experience' is taken as the ability quickly to form and use a motor plan for inputting information, as previously suggested. It would be expected that stable experienced subjects would input information more quickly than inexperienced stable subjects. With neurotic experienced subjects under stress, the plan should break down (motor disorganisation) and over-compensatory action result in a longer input time than that achieved by inexperienced neurotic subjects under the same stress.

The results of Experiment 4 further support the assumption in the foregoing explanation that general experience may be interpreted to mean the ability quickly to develop and use a motor plan and to use an already existing plan. Subjects with typing ability (measured in a simple task) input information faster on the problem solving tasks than subjects with lower typing ability. Secondly, the advantages afforded by general experience were amplified when using error prone input devices compared with less error prone devices.

However, general experience interacted with order of presentation (specific experience) and other factors in a complex way. In general, it was concluded that experienced subjects took less time correcting errors than inexperienced subjects. The degree of effect depended on the subjects' specific experience and whether there was any stress on the subject. Under the stress of using an error-prone input device with an unreliable slow system to solve a difficult problem, subjects' correction times were longer than when not under stress. The difference was greater for inexperienced subjects than experienced subjects. This implies that generally experienced subjects had less need of time to correct errors compared with inexperienced subjects. On gaining specific experience, generally inexperienced subjects were able to perform as well as the experienced subjects on their first trial.

The conclusion that internalised processes are developed through both general and specific experience is also supported by the results of Experiment 2.

The following summarises the effects of individual differences on the time for information input due to differences between input methods, systems and tasks:-

(i) The range of variations in input times due to differences between people and input methods can be of the same order.

(ii) Some input devices allow people more easily to trade-off speed for accuracy according to how they are used. Differences in input times due to differences between people are shown in these circumstances.

(iii) The usefulness of general computer experience depends on the input method and task characteristics. In simple tasks, it is tentatively assumed that general experience enables individuals quickly to develop 'motor' plans (using internalised processes) to use the input device effectively.

(iv) In problem solving tasks, general experience meant that individuals could quickly develop and use 'motor' plans. Under the stress of using error-prone characteristics with sub-optimum distal feedback to solve a difficult problem, the advantages of general experience were amplified.

(v) The effect on input times of specific experience gained using a particular input method and computer characteristics is as large as that of general computer experience. General and specific experience were equivalent in their effects in the experiments. They may be similarly explained as far as input methods are concerned to be due to the effective use of 'motor' plans.

(vi) The personality dimension of neuroticism-stability interacts with general computer experience in its effects on input rates with error-prone input devices. The effects are comparable in size with the range of variation to be expected when comparing different input methods in simple tasks. The effects are tentatively explained in terms of motor response disorganisation in the more 'neurotic' people under stress exacerbated by using error-prone input devices.

3.2 Input Methods in Relation to Task Characteristics

Experiments 1 and 2 examined the differences in information transfer rates due to different input methods used for a menu selection task. These studies were typical of those reported in the literature review in that they used a simple non-problem solving task and gave proximal and distal feedback of the result of each information transfer. This section deals with the effects on the information transfer where these simple tasks are part of a problem solving task.

Experiment 3 investigated the use of four different input methods used in interactive problem solving. Five different problems were used, the first of which was trivial and consisted of a predefined sequence. The times for the four input methods in the first (trivial) problem were 1.32, 1.36, 1.54 and 7.35 seconds per input respectively. In the remainder of the problems the average input times for the same four input methods were 2.08, 2.73, 3.08 and 7.84 seconds respectively. The differences between the two sets of figures are significant ($p < .05$, $t = 6.22$ between means of the smallest difference) except for the last figure 7.84 - 7.35). In the simple task, there was no significant difference between the input times for the fastest two devices, (1.32 and 1.36 seconds). However, in the problem solving task there was a significant difference (2.08 and 2.73 seconds). Thus, extrapolating from the results of simple input tasks to input tasks in problem solving, even where the other conditions remain the same, will generally be misleading. However, in some circumstances, input times may be predicted across tasks using the same input methods. These are when the cognitive, sensory/perceptual and feedback conditions as well as the speed/accuracy emphasis of the task are similar.

For example, in Experiment 2, an item was selected from a menu by typing a digit then confirming entry. In Experiment 5 a similar technique was used for making a command choice. Examination of the data shows that the mean menu entry time (4.5 seconds) is not significantly different from the mean time to enter a command

(4.6 seconds) in an attempted solution. But this was the only time when an explicit command letter needed to be typed. In the case where the command letter was implicit, and the entry not needed, the input times of the first primitives were significantly different.

To summarise this section on the effects of input method and task variables on information input times:

(i) Information input times measured with input methods on simple tasks are, in general, smaller than the times measured with the same input method used in more complex tasks. The extent to which times are increased depends on the input method and the task.

(ii) Exceptionally, in particular circumstances, information input times measured with simple tasks may be similar to those found in a problem solving task with similar input methods. The particular circumstances are when the cognitive, sensory/perceptual and feedback conditions and the speed/accuracy emphasis of the task are similar.

3.3 Input Methods in Relation to Computer Characteristics

Experiments 4 and 5 examined the differences in information input times as a function of a combination of input methods and computer characteristics. Combinations of variables were examined rather than single variables with the consequence that explanations of the effects are both tentative and complex. This discussion centres around two main themes: (i) the effects of 'dialogue' differences (i.e. non-automatic, pragmatic distal feedback with explicit commands versus automatic pragmatic distal feedback with explicit commands), and (ii) the effects of 'system/task' differences (i.e. unreliable systems with slow, inconsistent system response times and a difficult task compared with a fast, reliable system and an easy task).

3.3.1 'Dialogue' Differences

Experiment 5 was concerned with investigating the effects of different dialogue combined with different input methods. An effect was found which showed that the range of variation in command input times due to dialogue differences when using a particular input device was similar to the range of variation found between input methods within the same dialogue type (3.7 - 4.7 seconds).

The light pen was used faster than the keyboard for input of commands with the explicit command dialogue; the reverse was true for implicit command dialogue. A hypothetical explanation for the implicit dialogue results was that the movement time of the light pen was greater than for the keyboard, even though using the keyboard required shifts of attention from the display. The hypothetical explanation for the explicit dialogue results was that movement time was greater for the keyboard than the light pen because it included a time to change attention and search for and locate the required key. Analysis of the results supported this view in that, for the light pen, the average input time of explicit commands was smaller than the average input time of implicit commands, whereas the converse was true for the keyboard.

In Experiment 5, it was found that the explicit command dialogue in combination with an error prone input device, resulted in the longest times for error correction. This is consistent with a hypothetical explanation that the cognitive faculties are involved in determining the long times rather than just the activation of the device.

The results of Experiment 5 also showed that for both input methods, the dialogue requiring explicit requests for distal pragmatic feedback resulted in more frequent input errors than dialogue with continuous distal pragmatic feedback. These errors are assumed to be the result of unreliable human information processing rather than motor skill defects. On this assumption the notion that central cognitive processes are affected by the input method and dialogue is supported.

In Experiment 5 it was shown that there was no significant difference between the mean command entry times of the light pen with implicit dialogue and the mean command entry time of the keyboard with explicit dialogue. Such a finding would not be easily predicted from the results of Experiments 1 and 2 or similar simple task experiments. Thus, the results of Experiment 5 on command information input measures support the notion that dialogue differences interact with input method to alter the results which would be predicted from simple task evaluations.

3.3.2 'System/Task' Differences

The nomenclature of Experiment 4 was that an unreliable, slow computer with a difficult task was called a 'sub-optimum' system. According to the results of Experiment 4, regardless of the input method and dialogue type, the speed with which commands were input was faster with the sub-optimal system than with optimal system (fast, reliable computer with a simple task). The effect was greater with the keyboard than with the light pen.

The proposed explanation was in terms of the readiness of the problem solver to enter a command. The hypothesis was that the problem solver had a number of solution steps to try in sequence: the translation of these into commands was trivial in the experiment so that the problem solver was waiting for the system to be ready for input. This condition was signified on the keyboard by the printing of a prompt which cued the problem solver to enter a command. The condition was not signalled by the display for the light pen, therefore it did not cue the problem solver to react as quickly; but since the problem solver was prepared to input, the measured input time was smaller than in the optimum non-delayed system.

Strong evidence for the view is provided by the observation that subjects often typed a command before receiving a prompt.

3.3.3 Summary

(i) Regardless of variable combinations of dialogue type and input method used in the problem solving experiments the resulting variations in time for input of command information were comparable. The tentative explanation of this is in terms of different dialogues requiring different movement and attention shifts according to the input method.

(ii) Dialogue without automatic distal pragmatic feedback after every message input affected the time taken to correct errors in message input depending on the input method. The more error-prone the input method, the longer the error correction time with the dialogue.

(iii) Regardless of input method, the dialogue without automatic pragmatic distal feedback gave more frequent errors in input commands.

(iv) The speed of command entry was increased in the sub-optimum conditions of slow, inconsistent response time and a difficult task. The degree of increase depended on the input method such that when a prompt was given, faster input ensued. The explanation is that subjects were, in the system delay condition, sharing time with the computer. Thus, in the delay time, subjects searched for and located a key ready for input on receipt of a prompt. The light pen dialogue did not give an explicit prompt whereas the keyboard dialogue did. Hence the smaller input time depended on the input method.

3.4 Summary for General Hypothesis 2

This section has discussed general hypothesis 2 and the experimental results from three points of view. The views were related and overlapping and in addition to resulting in the specific summary of the sub-sections, they revealed a complex but identifiable number of interacting processes in interactive problem solving. The following summary identifies the most important processes relating to input methods.

a) Human Processes:

(i) The development of 'internalised' motor plans in developing motor skill

(ii) The continuation of information processing while waiting for the computer.

b) Computer Processes:

(i) Those determining the response time for distal feedback

(ii) The dialogue in relation to the provision of pragmatic feedback.

c) Interactive Processes:

(i) Interference between parallel human motor (and cognitive) activity and processing by the computer with error prone input methods and devices at all levels of processing. This results in an increase in the time taken by human processes.

4. General Hypothesis 3

"The acceptability of different input methods is based on an individual's judgement of a combination of factors affecting information transfer. The factors are the characteristics of input devices, computers and problems".

Experiments 2 to 5 attempted to explore the subjective preferences for the different combinations of factors. Experiment 2 on simple menu selection tentatively concluded that, if the error rates were equally low, the preferred input method was that which required least division of attention and so enabled faster input. With higher but equal error rates, the input method with the least penalty for error was preferred. With unequal error rates, the input method with the lower error rates was preferred.

Experiment 3 was concerned with a problem solving task and four input methods. Subjects preferences appeared to depend on a discriminative but unconscious judgement based on: first, if the error rates were equal, the time to enter a command, then, if times were equal, the error rates of the devices. The tolerance of comparison was large and the faster and least error-prone was preferred given that the penalties for error were the same. This supports the tentative conclusions of Experiment 2.

Experiment 4 used measures of acceptability which included not just the input method but all the computer and task conditions. It was noted that when the only difference between conditions was the input method, the one enabling the fastest input of data was preferred. This relationship was also constant with respect to individual differences. That is, a person judged acceptability on the basis of his own performance with an input method. Since the speed of input was a variable depending on the subjects general and specific experience as well as his personality, acceptability judgements varied between people depending on the input method.

In Experiments 4 and 5, it was noted that the order of preference of all the conditions (not just with differences in input methods) was based on a judgement of the number of 'plans' (long command input times). The systems needing the largest number of plans were least preferred.

To summarise:

(i) The input methods allowing the fastest input of data at an acceptable error rate are preferred.

(ii) People do not consciously discriminate between the speed and accuracy of different input methods. The association between preference and performance is as observed, not as reported.

(iii) Input methods are used in different ways by different people so that the rate of input varies. The judgement of acceptability varies accordingly.

(iv) Input methods are used in different problems and systems. This also varies the rate at which information is input to the computer and the acceptable error rate. Hence, there is no absolute rank of preference for input methods which can be ascertained from simple experiments.

(v) In making judgements of the acceptability of total systems, (i.e. different combinations of computer and input characteristics) people make judgements as if the basis for decision was the number of planning activities (as measured by long command-input times) required to solve their problems.

5. Conclusion

This chapter has discussed and summarised the experimental results in relation to the general hypotheses presented in chapter 2. Tentative conclusions were drawn and hypothetical explanations were presented based on assumed underlying processes and their interactions between and within man and computer. These form the basis of a more general discussion and recommendations in the following chapters.

CHAPTER 10

GENERAL DISCUSSION

CONTENTS

1. Introduction

Chapter 1 outlined the area of interest and reviewed pertinent literature. The conclusion was reached that research was needed which concentrated on the role of input devices in real time interactive problem solving so as (a) to identify the relative importance of the input method to the total system, and (b) to examine the validity of the assumption that results obtained by comparing input methods using simple tasks can be generalised into problem solving tasks. Chapter 2 then developed a descriptive model of interactive problem solving which emphasised the role of input devices and presented three general hypotheses. Experimental hypotheses were tested and the results discussed in relation to the general hypotheses. This chapter continues by reviewing and discussing the model, and (a) and (b) above, in relation to the experimental results.

2. A Model of Interactive Problem Solving Emphasising the Role of Input Devices

The purpose of this section is to reconsider the model presented in Chapter 2 in relation to the generalised results of this thesis. A brief review is given first of the model, and second, the generalised experimental results. These are then integrated and a revised view is developed.

2.1 A Review of the Model Presented in Chapter 2

The purpose of the model presented in Chapter 2 was to provide a framework for investigation rather than, for example, to provide a deterministic formulation for calculating performance characteristics. The framework was a model of information processing during interactive man-computer problem solving which did not attempt to identify exactly how or when particular processes were carried out in either man or computer.

The language used for describing both information and information processes was that of 'semiotics'. In this language, three levels of processing were identified; these were syntactic, semantic and pragmatic processing. Syntactic and semantic processing is concerned with communication between man and computer (and vice-versa) whereas pragmatic processing is concerned with interpreting information at the other levels in the light of the goals of the receiver.

The model has three building blocks. These describe human problem solving, human input processes and computer processes. Each was elaborated and a composite model formed on the basis of assumptions concerning how the blocks were inter-related.

The operation of this model was briefly described as a sequence of 8 stages as follows:

(i) The man encounters a 'problem'; that is he wants to reach an objective which cannot be obviously achieved.

(ii) The man perceives cues and organises attempts at a solution.

(iii) The man tries out this solution in his head ('internal' testing) and either modifies it or decides to try the solution in the 'real' external world ('external' testing).

(iv) External testing proceeds by taking each logical step in the trial 'internal' solution and translating this into a suitable number of steps for communication to the outside world.

(v) The model of human input processes in inputting to a computer takes each step for communication to the outside world and translates and operates on it until it is a sequence of items for physical transfer across an interface to the computer.

(vi) The items are transferred into the computer by means of human motor activation of the computers input devices. Proximal feedback may be received which aids this process.

(vii) The computer carries out appropriate syntactic and semantic processing of the communicated information providing distal feedback to the human input processes.

(viii) On completion of the input of a message representing part of a problem solution step, the computer carries out pragmatic processing and provides suitable feedback.

Depending on the computer result, the current step is either repeated or the next step is taken in the solution. This process continues until the human problem solver considers his solution is inadequate or complete. If it is inadequate, then he returns to stage (ii) above.

In this description, each stage is considered as being independent of the others and the process is largely sequential. The role of input devices is clearly to enable effective communication of syntactic and semantic information in this process between man and computer. The experimental results are now presented and discussed in relation to this simple descriptive model.

2.2 The Significance of the Results of the Experiments

In general terms, the stages of the descriptive model were observed to be present in the problem solving experiments. That is, the processes of internal and external testing and their sub-processes could be identified. However, it was clear that there were human activities and attributes which were not represented adequately in the descriptive model. These were described in section 3.4 of the previous chapter.

In the particular circumstances of the research of this thesis, the most significant results which affect the descriptive model of human interactive problem solving are as follows.

(i) Parallel processing occurs. Two types of parallel processing were identified; that between man and computer and that within the problem solver. In the former case, the man continues to process information and prepare for motor activity while the computer is processing information for him. In the latter case there is a degree of autonomy between human input processes and human problem solving processes depending on the particular individual, the task and computer characteristics.

The consequence of parallel processing for the descriptive model is that external and internal solution testing cannot be treated as independent stages. They interact with each other at a level depending on the degree of autonomy of the processes. The evidence of this thesis is that there is a human adaptive process which controls this level which should be represented in the descriptive model.

(ii) Limitations of underlying human cognitive faculties must be considered. The particular faculty found to be of importance in this thesis was short term memory. This limited, first the number of steps that were taken in a solution before pragmatic distal feedback from the computer was necessary and second, the amount of information to be communicated between man and computer. The consequence of this to the model is that there must be provision for changes of solution strategy so as to optimise the use of cognitive faculties depending on their loading at any time during the stages of solution testing. This means that the stages of the descriptive model may be interrupted whenever cognitive faculties are overloaded. This is a step towards including dynamic characteristics in the descriptive model and again identifies the need for more complex control stages in the descriptive model.

(iii) Individual differences may be more explicitly described in the model. Both the personality trait and general computing experience of individuals had significant effects in relation to the efficiency of the communication of information to the computer.

These results are guidelines for the development of the model which is reconsidered in the following section.

2.3 A Revised Model

The approach taken in this section is to present a hypothetical and conceptually simple descriptive model whose operation is similar in detail to that of Chapter 2. The revised model is shown in Figure 1 and is directly related to the descriptive model in Chapter 2 in the following way. In figure 1, the central cognitive processes (CCP) include the human problem solving processes of Chapter 2 and the peripheral processes (PP) of figure 1 include the human input processes and motor activity of chapter 2. The operation of the sub-processes is identical in both models. Also the assumed computer processes are exactly similar to those of chapter 2.

The links between the processes are not shown in detail since the purpose of presenting the simplified model is to emphasise the points made in the previous section. Therefore parallel processing and short term memory (STM) and long term memory (LTM) are introduced and emphasised in the simplified model.

In relation to the model of chapter 2, STM is used for the input and output store whereas LTM is used as the connection store and for holding a solution for external testing. However, as a consequence of the results of this thesis an extra, adaptive, process is postulated to be necessary. This controls the balance of the effort put into CCP and PP and acts so as to minimise the mental effort of the problem solver in reaching his goal. (This process is assumed to be a higher level process which monitors and controls the lower interactive processes, shown in chapter 2.)

The postulate is based on Zipf's work (1949) whereby it is assumed that people behave as if they have a general goal of minimising mental effort within their working constraints.

If Zipf's law is assumed to be valid, then the operation of the model is such that the degree of solution development through internal testing by using CCP in parallel with PP depends on an adaptive process. The adaptive process is hypothesised to be dependent upon STM characteristics, since it is the main link (output store) between CCP and PP when information is transferred to the computer.

The main characteristics of STM are that it is of limited capacity (Murdock (1965)) and that information decays through time dependent or interference processes (Newell and Simon (1972) discuss these aspects). These are inter-related in that overloading and decay of STM has similar results; i.e. loss of information from STM.

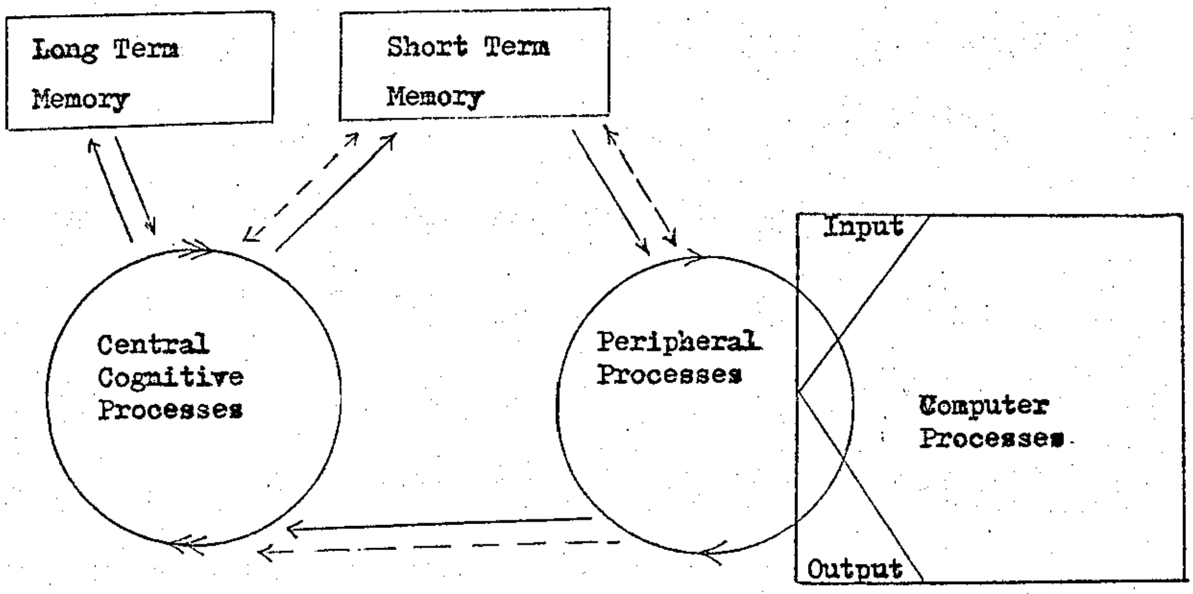


Figure 1

A Conceptual Model of Man-Computer
Problem Solving

Key:

─≫─ Flow of infomration and control in central processes
─>─ " " " " " " peripheral "
– – – – – Interference paths with normal flow

Note: Stability of Peripheral processes depends on personality.
Degree of independence of peripheral processes depends on experience.
Incompatable input characteristics causes interfence between central and peripheral processes.

(i) Limited Capacity: This refers to the number of items of information that can be held in short term memory. The information can be encoded into 'chunks' to enhance the capacity (Miller (1956)). However, this must be decoded before it is used by PP. Therefore, it may be expected that redundancy in the information to be input has the effect of using STM capacity and so leaves less capacity available for other processes.

(ii) Time dependent processes: There are two conditions wherein STM is required to be sustained. These are when there are long computer response times associated with the different levels of processing and when long times are needed for device activation by the problem solver. With respect to the use of keyboards, Conrad (1966) has shown that with low compatability keyboards, errors were made through memory rather than aiming errors. These arose because of the long times associated with locating a particular key.

(iii) Interference processes: This refers to the situation where STM is required to be used for some activity other than for PP. In this thesis an example was when an unexpected event occurred during PP such as accidentally hitting a screen item with the light pen. This caused the item to blink and so gain the problem solvers attention. This used his STM and so some of the previous information in STM was lost. In general, the error proneness of the input method is a particular source of any interference because of the need of the problem solver to divert attention to correct errors.

Given this variety in sources of loss of information from STM, the adaptive process which controls the balance of effort between CCP and PP had a range of flexibility via compensatory procedures which can be adapted. The following list gives examples according to the particular sources (i) - (iii).

(i) Limited Capacity: encoding information and using, when possible, memory aids such as written records.

(ii) Time dependent processes: rehearsal of items in STM by using CCP.

(iii) Interference processes: these effects are minimised by (a) learning to inhibit the response of switching attention given an acceptable error rate of input, and (b) trying to avoid interference by, for example, avoiding frequent use of error prone input methods.

The need for adaptation depends on the individual problem solver. This thesis suggests that people develop integrated motor plans through experience with particular input methods. These plans effectively reduce the effort needed by PP for inputting information to the computer. However, they may be difficult to form for error prone input methods and break down under stress, particularly for people who are neurotic compared with more stable people.

The development of the descriptive model as a framework for discussion and investigation is, for this thesis, complete. Recommendations are provided in the following chapter which are based on the revised model.

3. The Relevance of the Input Method to Real Time Systems

The question of relevance is addressed to both the users and designers of real time systems. This section has the purpose of discussing these questions in relation to the foregoing discussion and experimental results.

3.1 The Importance of the Input Method to the User

There are two related aspects of importance to the user. These are the performance characteristics and acceptability of the input method. The foregoing section considered a model of interactive problem solving which described performance aspects. The model emphasised that input methods and their error proneness affect problem solving processes at least in the conditions of the experiments. In this sense, therefore, the input method is important to the user since it affects the distribution of mental effort needed to solve the problem.

The users' judgement of acceptability was observed to be related to the number of plans made (inferred to be due to 'internal' solution testing without parallel processing) during interactive problem solving. Some input methods resulted in more plans and these were judged by subjects to be less acceptable than others which resulted in fewer plans. In terms of the revised hypothetical model, this may be expected if it is assumed that the most acceptable circumstances are when parallel processes can take place so as to reduce mental effort in solving the problem. This is not easily achieved with error prone or hard to use input methods as those which are error free and easy to use. With parallel processing, there is continuous solution development during peripheral processing and so there is less frequent need for dedicated use of CCP to develop solutions before they can be tested.

3.2 The Importance of the Input Method to the System Designer

Input methods are part of the dialogue that must be designed when developing a man-computer system. In his book on the design of man-computer dialogue, Martin (1973) suggests a design methodology consisting of a possible 21 inter-related steps. It is very much a 'top-down' design view point starting with user requirements and working towards a computer program specification. Many of the steps are explored in detail in the book, such as assessing the capabilities of the operators of terminals, and relating the dialogue and response time requirements to the computer configuration and control program. This is done by descriptive case studies and by making reference to fragmented literature on experimental work.

The pattern is reinforced by Rouse (1975) who attempted to integrate pertinent literature into a conceptual framework for the design of man-computer interfaces. There is an evident gap between the way in which such information is presented (piecemeal and not related) and the way in which systems designers require it (related so that trade-offs can be seen). This gap is very large and requires sustained efforts by human factors and computer scientists before it may be filled. An indication of the trend towards filling the gap in relation to dialogue design is the recent research carried out by Carlisle (1974). This thesis is another example. The further analysis of the data for real time systems development (presented in Chapter 8) showed how the data in this thesis may be useful. It also shows the relative importance of the input sub-system to total system dynamic behaviour.

Two further possible mechanisms are proposed in this thesis whereby the communication and knowledge gap may be filled. The first is the development of the descriptive model and the second is the use of the language of semiotics for describing dialogue in a formal way.

The descriptive model developed in this thesis is intended to be a convenient and common framework for talking about issues and may be used, in a general way, to discuss user behaviour given a particular set of conditions. Thus, in designing that part of the dialogue concerned with the input method, reference must be made, first to the model, and then to appropriate literature such as that quoted in this thesis. By developing a range of models which have the same basic form, but vary in their emphasis, the designer may form a balanced view of the system. This can be a sound basis for design decisions. This thesis has shown that there is a need to use such a model because of the lack of generalisation of results of comparative experiments using simple tasks with those using problem solving tasks.

The language of semiotics is suitable for such a general view as described in this thesis. It has also been used in a more formal way in each chapter which describes an experiment to specify the input method. The main aim of this was to provide a terse description of the input method with neither redundancies nor omissions.

This can then be used in many ways; for example, to prepare quickly and compare two alternatives and yet provide a reasonable specification to a programmer who must implement it. It would normally be accompanied by a similar description of the output characteristics for the latter purpose. More research is needed into this aspect from the designers point of view.

4. Conclusion

This chapter has briefly discussed general issues in relation to the results obtained in this thesis. The issues concerned the review of the descriptive model presented in chapter 2 and the identification of the importance of the role of input devices to the user and system designer. The thesis continues by presenting recommendations for researchers and designers based on the discussion chapters.

CHAPTER 11

RECOMMENDATIONS

CONTENTS

1. Introduction

The purpose of this chapter is to provide recommendations. They are given within the range of generality implied by the research. That is, within a particular range of types of problem solving and real time command systems.

The usefulness of the research depends on the way the results are used. In the introduction to the thesis, two main groups of users were identified. These were (i) researchers who wish to continue to explore the area of research and (ii) systems designers (and their like) who require to know how the results presented here and elsewhere may be used. Therefore, recommendations are provided for each group.

2. Recommendations for Human Factors Researchers in Man-Computer Interaction

The following recommendations are made for future research:

In general,

(i) Testable research hypotheses should, if possible, be developed from general hypotheses in a systematic way which takes into account a balanced view of man-computer problem-solving. While desirable, the procedure was not developed far in the research reported here which was largely an empirical approach. Further research should aim to test the generality of the results found here both by 'laboratory' and field experimentation. The purpose of this would be to clarify further the picture of the main processes and effects in a descriptive model such as that given in this thesis

(ii) Other views which emphasise the characteristics of computers, people or problems should also be taken and lead to a similar research objective, i.e. providing a clear view of the main processes and effects in a descriptive model.

(iii) Where possible, models should focus on invariant aspects; i.e. they should not be dependent on technological change. This implies a more systematic development which allows the explicit role of particular cognitive faculties like, for example, short term memory, to be examined.

Specifically, during discussion in this thesis a number of points were raised for further research. These were:-

(a) Is the result that input methods affect the number of steps per solution trial an artifact or is it a general finding?

(b) Was the absence of the effects of individual differences on certain measures because of a small sample size or because the effect is not significant in the test conditions?

(c) How does a different construction of individual differences change the view of man-computer problem-solving? (If 'level of aspiration' rather than 'neuroticism' had been postulated would this lead to greater insight into the underlying processes?)

(d) How does a penalty-for-error affect the time to input information with different input methods? Can this be used to predict the time to input in problem solving?

(e) How does the break up of semantic information into groups for input affect the conclusions on information transfer times?

(f) Are differences in input times with different input methods always mainly shown in measures using the first prescriptive primitive of a message?

(g) Are differences between people in information input times always or mainly shown in other than the first entry of a message?

3. Recommendations for System Designers

3.1 General Recommendations

The following recommendations apply to the design of many classes of real time command systems.

(i) The cognitive strain on inexperienced operators should be minimised by designing and testing dialogues for real time problem solving systems. This may be achieved by various means as outlined in, for example, Martin (1973) or Spence (1976). A major way is to develop compatible input methods (as defined in this thesis) and evaluate them according to the recommendations given for human factors researchers.

(ii) In developing compatible input methods allow for two main aspects which, though interactive, may initially be optimised independently. These are the device and the way it is used. Some devices have hardware characteristics which make them more error prone than others, but there may be some optimisation possible through variations in associated variables (for example, light pens and the display variables as outlined in this thesis). All input devices have performance characteristics which may be optimised by using a suitable coding scheme which minimises the operator's cognitive strain. The coding scheme should be based on how the language is used during problem solving so that frequently found sequences of information to be input are coded into one device operation. Thus, it is essential to evaluate input methods in the context of their use, since this determines the overall cognitive strain.

(iii) Inadequate provision (slow and incomprehensible) of distal feedback at the syntactic, semantic and pragmatic levels progressively increases cognitive strain. In conjunction with poor input methods, inadequate feedback may strain the problem solver's cognitive ability so much that serious impairment of performance results. This may be as large as inability to solve the problem or simply a change of the problem solver's strategy in solution so as to minimise his cognitive effort. The change in user behaviour affects both communication and uses of the computer resources. Hence this is one way that may be used to optimise the cost of solving problems with man-computer systems. In general, the greater the cognitive strain imposed by the computer system the less it is used. Thus, in principle, the cost of a man may be balanced with that of the system by altering system characteristics to achieve an optimal problem solving man-computer system.

(iv) People judge the relative acceptability of different input methods according to how fast they can input information in the <u>context of their use</u>, not in different situations. The rates are usually significantly different. Therefore make sure any evaluation of alternatives is done in as realistic a context as possible.

(v) The rate at which people input information depends on the person and not just the input method and the context. Hence their judgement of acceptability may vary as much due to personality as to the differences in input methods. This is true mainly where input methods allow users to trade-off speed for accuracy. Make sure any evaluations use a sample of people who in their characteristics are representative of the eventual operators.

3.2 Recommendations for the Designers of Time-Sharing Multi-User Systems

So far, what has been said has been general to many classes of man-computer systems designed for real-time problem-solving.

The general recommendations for system designers of multi-user time-sharing systems are as follows.

(i) Take into account the fact that input methods affect the rate at which messages are received by the processing computer and therefore both the rate at which resources are used and the rate at which users can be serviced.

(ii) If possible, design the scheduling of time and resources to take into account any patterns in the input message distribution time. Such patterns are specific to tasks and users and terminals. Therefore, the cost of recognising patterns should be borne in mind. However, the benefits are in improving the acceptability of the system for users by reducing their cognitive strain, and increasing the amount of work done by the computer in a given time.

(iii) The rate at which messages can be input is the basis for a user's judgement of the acceptability of a real time system. This partially depends on the input method. Therefore optimise the input method.

(iv) Central computer system performance can be enhanced by optimising the input method for particular circumstances.

4. Conclusion

This chapter has summarised a number of recommendations grouped according to the intended audience and based on the discussion chapters. The thesis concludes by summarising the main points and reviewing the research as a whole.

CHAPTER 12

SUMMARY AND CONCLUSIONS

CONTENTS

1. Introduction

This final chapter summarises the logical approach of the thesis, lists the main findings and then concludes the thesis with a brief appraisal.

2. Summary

2.1 The Logical Development of the Thesis

At the beginning of this thesis it was stated that the objective was to provide information on the constraints imposed by different input devices and how they are used. This was to be done so as to enable the relative importance of the input sub-system of a computer system to be assessed. The means of reaching this objective was through four stages: a review of literature; the development of a descriptive model; deriving and testing experimental hypotheses, and finally, providing guidelines for present and future researchers and system designers.

The literature review showed that, despite a large variety in the number of input devices and how they are used, there was no clear framework within which to carry on investigations; nor were there any evaluations of different alternatives for interactive problem solving.

The thesis continued by proposing a descriptive model which was the framework for developing general hypotheses. The bases of the model were: the language and concepts of semiotics (the theory of signs); a model of how people use a keyboard; a model of human problem solving and a model of computer processes. These were put together in such a way that the role of input devices was emphasised. It was accomplished by borrowing from previous work on modelling man-computer interaction.

The general hypotheses were related to testable experimental hypotheses. Practicalities and the exploratory nature of the research were reflected in the non-systematic relationships between the testable research hypotheses and the general hypotheses.

Five experiments were then described which tested the experimental hypotheses and developed operational definitions of the variables. The results were briefly discussed and tentatively explained in terms of the descriptive model.

The descriptive model was revised and hypothetical processes generated to account for the experimental results. Some analysis and discussion was presented which examined the relative importance of input method to the user and the system designer.

Finally, recommendations were made for human factors researchers and systems designers. The next section lists the general results found in this research.

2.2 Generalised Research Conclusions

Within the limits of generality set by the scope and the experimental contexts of the research the following general conclusions are drawn.

(i) Interactive man-computer problem solving processes are affected by the input method of the computer. The effect is such that the number of steps made in an attempted solution decreases as the time needed to input information to take that step increases. The degree of effect is context dependent; it has most effect in systems where all other characteristics are optimal.

(ii) The variations to be expected in the time to input information due to differences between input methods are large and similar in size to the differences due to variations due to differences between people. The differences due to people are generally shown with input devices which are error prone.

(iii) By being familiar with the input procedures and concepts, experienced people, in problem solving, have less cognitive strain than inexperienced people. The differences between people (due to experience) in input times is amplified by poor system characteristics.

(iv) The advantages of general computing experience and specific experience of particular input methods are similar and of the same order of effect.

(v) The personality trait of neuroticism interacts with general experience in its effects on input time for input methods which are error prone. The sizes of the effects are less than those due to using different input methods. The effect is mainly that neuroticism degrades the advantage of general experience by slowing down input time.

(vi) In general, the time taken to input pragmatic information depends on the task so that input time is increased on a difficult problem solving task compared with a simple problem solving task. The degree of increase partially depends on the particular input methods. Only exceptionally are results directly comparable in different tasks.

(vii) The time needed to input information depends on the interaction between the input method and the computer characteristics. The time needed to input information with poor input methods and adequate computer feedback of information was comparable with the time taken using good input methods with inadequate feedback. The effect is explained in terms of user movement time and attention shifts differentially affected by the cognitive loading according to the input method.

(viii) Computer characteristics with inadequate feedback causes cognitive strain which increases the number of input errors. This effect is increased by poor input methods.

(ix) If people are locked out by the computer, they use the lock-out time to prepare for input when the lock-out time is comparable with the time needed to input without a lock-out.

(x) Those input methods allowing the fastest input time are preferred by a particular individual working in a particular context. If times are comparable for different input methods, the least error prone is preferred. People do not consciously report the basis of these judgements.

(xi) There is no absolute measure of acceptability. In making judgements of acceptability people take into account all the characteristics of tasks and computers and input methods. A comparison made between two conditions of one aspect is only valid if all other conditions are exactly similar. In making a judgement of combinations of characteristics, people base decisions on the number of long times needed to organise and plan solutions using these characteristics. The more of these times, the less acceptable the conditions.

(xii) Despite a wide variety in possible views, the observed behaviour can be readily described using a small number of convenient concepts. The language of description and the general relationships between the concepts have commonalities which should enable system designers, human factor researchers and others to communicate and so develop a model of interactive problem solving. The use and development of a model in this thesis indicates that models may be possible at an adequate level for use by systems designers who wish to compare different designs on the basis of consequences on human and system behaviour. But a great deal of research is needed which focusses on various aspects of man-computer problem solving.

3. Conclusion

In the introduction to this thesis, the need for research in man-computer interaction was briefly discussed and the topic of research for this thesis was presented. The topic is the study of the role of input methods in interactive problem solving with real time command systems. The literature review concluded that there was a need for studies of input methods in interactive problem solving.

The thesis then developed an approach to the study of this topic which attempted to integrate relevant concepts into a descriptive model. This was used to derive a range of hypotheses. These were discussed in relation to the results of five experiments, and the descriptive model was revised. A number of recommendations are made for human factors researchers who wish to develop further this research and for systems designers who wish to use the results for evaluating or designing systems.

The general findings of this research are as follows.

(i) The generalisation of the results of investigations of input methods using simple input tasks (as reviewed in the literature) into situations where the input method is used in problem solving is not valid. The results of time measurements are particularly affected. However, such investigations are useful for examining the error proneness of different input methods and identifying the reasons for their difference.

(ii) A descriptive model of the information processing in interactive man-computer problem solving is a convenient framework for discussing the processes affected by alternative input methods. This can be used for developing investigations and for understanding the implications of the results of other investigations, which focus on similar and related aspects (e.g. system response time).

(iii) The input method used for real time interactive problem solving does affect how the problem is solved and, in multi-user systems, the dynamic characteristics of the system. The user acceptability of a real time system partially depends on the input method. For these reasons, the investigation of the role of input methods is considered to be important for the design of real time systems.

It should be mentioned that the research is multi-disciplinary and was intended to be practically and theoretically relevant to the appropriate group. Consequently, it may be that from the particular viewpoint of a single discipline (e.g. psychology or computer science), this thesis has scope for both development and criticism. The view presented in the thesis is a personal one which has been formed with two aims. These were; (i) the need to make a general contribution to the study of human factors in man-computer interaction, and (ii) the need to provide a practically sound basis for the development of research and systems development. It is intended to continue with these aims in the context of research and development of a large scale CAD system. In particular, field studies will be planned and undertaken on the basis of this research. It is hoped that it may be equally useful for the general reader and that other workers will continue the research.

In 1969, Professor Shackel talked of the contribution of the human sciences and this work grew out of that approach. Compared with the rate of growth of computing and developments within it over the last eight years, the rate at which work like this is produced is very small. The recognition by systems designers of the need for such work has started to generate more research into this area (e.g. Spence, R. and Goodman, T. (1977)). Therefore, it may be appropriate to be optimistic about more intensive development of this and similar research in the immediate future.

13. GLOSSARY

The following terms are provided for reference in alphabetical order. While most terms are not used in any special way, the reader who is unfamiliar with them may find it more convenient to refer to this section than to where the term is first used.

Term	Meaning
Acceptability	In this thesis, it may generally be interpreted as the preference for one set of conditions over another.
ANOVA	An abbreviation for analysis of variance. The technique most frequently used in this thesis for analysing the results of factorial designs.
Command Input	To do with.the input of the first primitives of a message.
Computer Characteristics	In this thesis are details of distal feedback and input error recovery.
Data Input	To do with input other than the first primitive of a message.
Distal Feedback	This refers to feedback from the computer to the man which is specific and indirect (see Proximal).
Evaluation Time	This refers to the time when the user delays his input of a message because of error correction at a syntactic level.
Information	Signals carried by primitives, words or messages.

Input Method	The combination of an input device and how it is used for conveying messages to the computer.
Input Time	The time for the user to enter information from the time when the computer can accept it.
Menu	A list of words or symbols displayed on a computer output device.
Message	Strings of words.
Plans	The activity of planning is assumed to be responsible for a delay of the user input while he works out his next sequence of moves.
Pragmatic Processing	Wherein messages are interpreted in terms of the internal state and goals of the receiver.
Primitives	That which is transmitted between man and computer as a minimum unit (see syntax).
Proximal Feedback	Feedback from the input device which is non-specific but direct and fast.
QWERTY	The abbreviation for the standard typewriter keyboard layout.
R1	To do with the input of the first primitives of a message.
R2	To do with primitives other than the first primitives of a message.
Semantic Processing	Wherein messages are checked against some grammar rules.
Semiotics	The theory of signs which leads the ideas of syntax, semantic and pragmatic processing.

Symbol	That which may represent a primitive.
Syntax Processing	That which decodes primitives.
System Response Time	The time in which the system appears 'dead' to the user through processing or lock-out of input.
Task Characteristics	This refers to task difficulty as defined by the information content and the number of steps to solve a problem.
TOTE	Abbreviation for Test-Operator-Test-Exit - a unit of behaviour.
Transaction	An exchange of pragmatic information between man and computer and back again.
User Characteristics	These are personality, cognitive faculties, motor skill and knowledge of the computer language.
URT	Abbreviation for User Response Time. In this thesis it is used as input time.
Words	Groups of primitives which are for semantic processing.
Work Production	The amount of work done in a particular time. In this thesis, input time is one measure of work production; overall times to solve a problem and number of errors are others.

14. ACKNOWLEDGEMENTS

The ideas and work contained in this thesis have been developed over a period of four years in two environments. The basic philosophy and the first problem solving experiment were developed while the author was a member of the group known as 'HUSAT' (Human Sciences in Advanced Technology) under the directorship of Professor B. Shackel in the Department of Human Sciences of Loughborough University. I would therefore like to record my appreciation to Professor Shackel for both creating such an environment and for encouraging me to develop within it. Without his work, this thesis would not exist.

I would also like to thank Dr. Christopher Evans of the National Physical Laboratory whose encouragement, discussion and management of my work while at HUSAT was a major factor in the production of concrete research results.

Thanks are also due to all those members of the Department of Human Sciences and HUSAT who, at various times, have given their time and ideas freely during my stay there. Particularly long suffering were T.J. Folkard, R.F. Penn, T.F.M. Stewart, L. Damodaran and, of course, my supervisor, K.D. Eason.

No less a debt is due to the second environment wherein four experiments were carried out. This was at the British Ship Research Association, Wallsend, Tyne & Wear. The experiments were carried out as a part of the development of a large scale computer aided ship design system managed by Mr. G.R. Snaith (now Director of the Research

Station) and Dr. M.J. Todd. Both were instrumental in actively encouraging the research and providing an appropriate environment. Each gave both useful advice and stimulation for discussion on the relevance and usefulness of the research. It is a measure of their interest in this research that continuation of it is planned in 'field' situations with the working system. Such people and environments are rare and I would like to record my appreciation of both. In addition, I would like to thank Dr. H.O. Dickinson for commenting on and checking statistics, and BSRA for permission to use the results and for the use of tracing, library and typing facilities in the production of this thesis.

There are a number of long suffering people who deserve a medal for persevering over the years during which this thesis was developed and written. Principal among these are Ken Eason, my supervisor, who read endless drafts and commented both politely and with enthusiasm; secondly, my wife and family who tolerated idiosyncractic behaviour, and finally all those typists who translated the handwritten text into a presentable form, particularly Carol Dawson.

Finally, my thanks go to the seventy or so subjects who gave both time and information in the experiments for no reward.

15. REFERENCES.

ALDEN, D.G. et al. (1972) Keyboard Design and Operation.
Human Factors, 14, 275 - 293.

BAKER, (1960, 1961) Abstracts and Sources in Sperandio and Bisseret (1968)

BAKER, J.D. (1976) Modelling the User.
Paper presented at NATO Advanced Study Institute on Human Factors in Man-Computer Interaction, held in Mati, Greece.

BAKER, J.D. and JOHNSON, E.M. (1974) Field Testing.
Special Issue of Human Factors, 16. 199 -325.

BOEHM, W.M. SEVEN and WATSON. (1971) Interactive problem solving - an experimental study of lockout effects.
AFIPS Conference Proceedings, 38, 205 - 210.

BOIES, S.J. (1972) User Behaviour on an Interactive Computer System.
IBM Research Report RC 4169, Yorktown Heights, New York.

BRADLEY, J.V. (1975) The Optimal - Pessimal Paradox.
Human Factors, 17, 321 - 327.

BRAUNSTEN, M. and ANDERSON, W.F. (1959) A Comparison of Reading Digits Aloud and Keypunching.
IBM Research Memo RC - 185.
Yorktown Heights, New York.

CARBONELL, J.R. (1967) On Man-Computer Interaction; A mode and some related issues.
IEE Conference on System Science and Cybernetics, Oct. 11-13, 34.

CARBONELL, J.R. ELKIND and NICKERSON, R.S. (1968) On the Psychological Importance of Time in a Time-Sharing System.
Human Factors, <u>10</u>, 135 - 142.

CARLISLE, J.H. (1974) Man-Computer Interactive Probl
Solving Relationships between User Characteristics and Interface Complexity.
Thesis, Yale University.
Source; Ergonomics, <u>19</u>, 657.

CHAPANIS, A. (1961) Men, Machines and Models.
American Psychologist, <u>3</u>,113-1

CHAPANIS, A. (1967) The Relevance of Laboratory Studies to Practical Situation
Ergonomics, <u>10</u>, 557 - 577.

CHERRY, C. (1957) On Human Communication
London, Chapman and Hall.

CONRAD, R. (1958) Accuracy of Recall using keyou
and telephone dial.
Journal of Appd. Psychology, <u>42</u>, 285 - 288.

CONRAD, R. (1962) Practice, familiarity and reading rate for words and nonsense syllables.
Quarterly Journal of Experimental Psychology, <u>14</u>, 71-

CONRAD, R. (1966) Short Term Memory factor in th
design of Data Entry Keyboards
Journal of Appd.Psychology, <u>50</u>, 353 - 365.

CONRAD, R. and HULL, A.J. (1968) The preferred layout for numerical data entry layouts.
Ergonomics, <u>11</u>, 165 - 175.

CONRAD, R. (1967) Performance with Different Push Butto
Arrangements.
HET-PTT-BEDRIJF DEEL, 15, 110 -113.
Source; Seibel (1972) .

CORLEY, M.R. and ALLAN, J.J. (1976) Pragmatic Information Processing Aspects of Graphically Accessed Computer Aided Design.
IEEE Trans. on Systems, Man and Cybernetics, SMC-8, 434 - 439.

COX, R.B. and MARTIN, T.B. (1975) Speak and the Machines Obey.
Industrial Research, Nov. 15, 1 - 6.

COX, D.R. (1972) Regression Models and Life - Tables.
J. Royal Statistical Society, B.34 , 187 - 220.

DEININGER, R.L. (1967) Rotary Dial and Thumbwheel Devices fo
Manually Entering Sequential Data.
IEEE Trans. of Human Factors in
Electronics, <u>HFE-8</u>, 227 - 231.

DVORAK, A. (1943) There is a better Typewriter Keyboard
National Business Education Quarterly
<u>12</u>, 51 - 58.

EARL, W.K. and
GOFF, J.D. (1965) Comparison of Two Data Entry Methods.
Perceptual and Motor Skills, <u>20</u>,369-3

EASON, K.D. (1976) Understanding the Naive Computer User
Computer Journal, <u>19</u>, 3 - 8.

ELLIS, T.O. and
SIBLEY, E. (1969) Grail - An Experiment in Man-Machine
Communications.
Rand Corp. Report RM-5999 - BRPA.
(see Newman and Sproull (1973)).

ENGLISH, W.K.
ENGLEBART, D.C. and
BERMAN, M.L. (1967) Display Selection Technique for Text
Manipulation.
IEEE Trans. of Human Factors in
Electronics, <u>HFE-8</u>, 5 - 15.

EYSENCK, H.J. (1952) The Scientific Study Personality.
Routledge and Kegan-Paul Ltd., London

GOODWIN, N.C. (1975) Cursor Positioning Using Light Pen,
Light Gun and Keyboard.
Human Factors, 17, 289-295.

GREEN. et al. (1970) Problem Solving: Research Method and
Theory.
J. Wiley, New York.

GUETZKOW, H. (1951) An Analysis of the operation of Set i
Problem Solving Behaviour.
Journal of General Psychology,
45, 219 - 244.

GRIFFITH, R.T. (1949) The Mini-Motion Typewriter Keyboard.
Journal of the Franklin Institute,
Report No. 246.5, 339 - 425.

GROSSBERG, M. et al. (1976) An Experiment on Problem Solving with
Delayed Computer Responses.
IEEE Trans. on HFE, F176, 219 - 222.

HAYES, J.R. (1966) Memory, Goals, and Problem Solving.
Chapter 6 in "Proceedings of
Conference on Problem Solving;
Research Method and Theory", reported
in Green et. el. (1970).
Editor : Kleinmuntz. J.Wiley & Sons,
New York.

HILLIX, W.A. and
COBURN, R. (1961) Human Factors in Keyset Design.
US Navy Electronics Lab. Report No.
AD 622 368.

INNOCENT, P.R. and
CLAYTON. (1975) Methods of Using the Weibull
Distribution.
British Ship Research Association
(Wallsend) Internal Report.

JENKINS, W.L. and KARR, A.C. (1954) The use of a joy-stick in making settings on a simulated scope face. Journal of Appd. Psychology, 38, 457-461.

JOHNSON, E.A. (1967) Touch display: a programmed man-machine interface. Ergonomics, 10, 271 - 277.

LEONARD, J.A. and NEWMAN, R.C. (1965) On the acquisition and maintenance of high speed and high accuracy in a keyboard task. Ergonomics, 8, 281 - 304.

LICKLIDER, J.C.R. (1960) Man-Computer Symbiosis. IRE Trans. Human Factors in Electronics, HFE-1, 4-11.

LONG, J. (1976) Effects of Delayed Irregular Feedba on Keying Performance. Ergonomics, 19, 183 - 202.

MAIER, N.R.F. (1931) Reasoning in Humans. Journal of Comparative Psychology, 12, 181 - 194.

MARTIN, J. (1973) Design of Man-Computer Dialogues. Prentice-Hall.

MEISTER, D. (1973) Systems Development. Ergonomics, 16, 267 - 280.

MILLER, G.A. (1956) Information and Memory. Scientific American, 195, 42 - 46.

MILLER, R.B. (1968) Response Time in Man-Computer Conversational transactions. AFIPS Conference Proceedings, 33, 267 - 277.

MILLER, R.B. (1969) Archetypes in Man-Computer Problem Solving. Ergonomics, 12, 559 - 581.

MILLER, G.A., GALLANTER
and PRIBRAM. (1960) Plans and the Structure of Behaviour.
Holt, Reinhart and Winston.

MILLS, R.G. and
HATFIELD, S.A. (1974) Sequential Task Performance: Task
Module Relationships, Reliabilities
and Times.
Human Factors, 16, 117 - 128.

MINOR, F.J. and
REVESMAN, S.L. (1962) Evaluation of Input Devices for a Dat
Setting Task.
Journal of Appd. Psychology, 46,332-33

MUMFORD, E.and
SACKMAN, H. (1975) Human Choice and Computers.
North-Holland Publishing Co.

MURDOCK, B.B. (1965) A Test of the 'Limited Capacity'
Hypothesis.
Journal of Experimental Psychology,
69, 237 - 240.

NEWELL, A. and
SIMON, H.A. (1972) Human Problem Solving.
Prentice-Hall.

NEWMAN and SPROULL (1973) Principles of Interactive Graphics.
McGraw-Hill.

NICKERSON, R.S.
ELKIND, J.E.and
CARBONELL, J.R. (1968) Human Factors and the Design of Time-
Sharing Computer Systems.
Human Factors, 10, 127 - 134.

NICKERSON, R.S. (1969) Man-Computer Interaction: a Challenge
for Human Factors Research.
Ergonomics, 12, 501 - 507.

PLATH, D.W. and
KOLESNICK, P.E. (1967) The Thumbwheel Switch as a Data Entry
Device.
Human Factors, 9, 479 - 482.

RAY, W.S. (1955) Complex tasks for use in Human Problem
Solving Research.
Psychological Bulletin, 52, 134 - 149.

ROUSE, W.B. (1975) Design of Man-Computer Interfaces for On-Line Interactive Systems.
Proceedings of IEEE, 63, 847 - 854.

SACKMAN, H. (1970) Experimental Analysis of Man-Computer Problem Solving.
Human Factors, 12, 187 - 201.

SEIBEL, R. (1962) Performance on a 5 finger keyboard.
Journal of Appd. Psychology, 46, 165 - 169.

SEIBEL, R. (1964) Data Entry through Chord parallel entry devices.
Human Factors, 6, 189 - 192.

SEIBEL, R. (1972) Data Entry Devices and Procedures.
Chapter 7 in "Human Engineering Guide to Equipment Design".
Editors: Van Cott/Kinkade, published b
US Armed Forces Joint Steering Committ
American Institutes for Research, Washington, D.C

SHACKEL, B. (1969) Man-Computer Interaction - the Contribution of the Human Sciences.
Ergonomics, 12, 485 - 499.

SHACKEL, B. (1976) Man-Computer Interaction(MCI) - Realit and Problems.
Opening Address at NATO Advanced Study Institute on Man-Computer Interaction, Mati, Greece.

SIEGEL, S. (1970) Non-Parametric Statistics.
McGraw-Hill Series in Psychology, 1970 edition.

SIMON, H.A. (1966) Reflections on Time-Sharing from a User's Point of View.
Computer Science Research Review, 43-52
Carnegie Institute of Technology.

SPENCE, R. and
GOODMAN, T. (1977) Response Delay Effects on Problem
Solving.
SRC Report of Work, No.1, carried out
at Department of Electrical Engineerin
Imperial College, London University.

SPENCE, R. (1976) Human Factors in Interactive Graphics.
Computer Aided Design, 8, 49 - 53.

SPERANDIO, J.C. and
BISSERET, A. (1968) Human Factors in the Study of
Information Input Devices.
Bulletin de CERP, 17, 269 - 294.
Library Translation No. 1728 from
RAE, Farnborough.

STAMPER, R. (1973) Information.
Wiley and Sons, 1st edition.

STEWART, T.F.M. (1974) Computer Terminal Ergonomics.
Statskontoret Rapport - Source:-
Department of Human Sciences,
Loughborough University.

THORNTON, G.B. (1954) Comparison of an experimental rolling-
ball control and a conventional
joystick.
Defence Research Board of Canada, DRML
Report 107-1

VAN GEFFEN, L.M.H.J.
(1966) A Review of Keyboarding Skills.
Advances in Computer Typesetting, 2 - 1
Proceedings of International Typesettin
Conference, IOP, London.

VARTABEDIAN, A.G. (1970) Human Factors Evaluation of several
Cursor Forms for use on Alphanumeric
CRT displays.
IEEE Trans. on Man-Machine Systems.
MMS-11,132-137.

WICKELGREN. (1972) How to Solve Problems.
McGraw-Hill.

WILKES, M. (1970) Time-sharing Computer Systems.
2nd Edition. London; MacDonald.

YULE, A. (1972) Human Response Times in a Graphic Environment.
Computer Bulletin, 16, 304 - 306.

ZIPF, G. (1949) Human Behaviour and The Principle of Least Effort.
Hapfner, New York/London.

16. BIBLIOGRAPHY.

ALLPORT, G.W. (1961) Expressive Behaviour.
From Pattern and Growth in Personalit
461 -490,
Holt, Reinhardt and Winston.

AMBROZY, D. (1971) On Man-Computer Dialogue.
International Journal of Man-Machine
Studies, 3, 375 - 383.

ANNETT, J. (1969) Feedback and Human Behaviour.
Penguin Science of Behaviour Series.

ANNON Introduction to Programming - Part 2.
An algorithm for the Tower of Hanoi,
11 - 64. Published by Digital Equipme
Corporation.

ANTHONY, W.S. (1969) Working Backward and Forward in Proble
Solving. British Journal of Psycholog
57, 53 - 62.

BLOOM, B.S. (1964) Stability and Change in Human
Characteristics.
J. Wiley and Sons.

BRIGHAM, F. (1974) Analysis of Variance in Ergonomics.
Ergonomics, 17, 259 - 268.

CARNFORTH, C. (1976) A Model of Architectural Design
Activity.
M.Sc. Thesis, Department of Design
Research, Royal College of Art, Londo

CHAPANIS, A. (1973) Interactive Human Communications.
Paper presented at NATO Advanced Stud
Institute on Man-Computer
Interaction, Mati, Greece.

CHERRY, E.C. (1957) On the Validity of Applying
Information Theory to Experimental
Psychology.
British Journal of General Psychology,
48, 176 - 188.

CLARIDGE, G.S. (1961) Arousal and Inhibition as Determinants of the Performance of Neurotics.
British Journal of Psychology, 52, 53 - 60.

CONRAD, R. and HULL, A.J. (1968) Input modality and the serial position curve in short-term memory.
Psychonomic Science, 10, 210 - 215.

CONRAD, R. and HULL, A.J. (1967) Copying alpha and numeric codes by hand an experimental study.
Journal of Appd. Paychology, 51, 444 - 448.

CORCORAN, D.W.J. (1965) Personality and the Inverted U Relation
British Journal of Psychology, 56, 267 - 273.

CROPPER and EVANS. (1968) Ergonomics and Computing.
Computer Bulletin, 12, 94 - 98.

DEVOE, D.B. (1967) Alternatives to Handprinting in the Manual Entry of Data.
IEEE Human Factors in Electronics, HFE-8, 21-32.

DE GREENE, K.B. (1974) Models of Man in Systems in Retrospect and Prospect.
Ergonomics, 17, 437 - 446.

EASON, K.D., DAMODARAN, L. and STEWART, T.F.M. (1976) MICA Survey Report.
S.S.R.C. Publication under Grant No. HR 3844/1. Work carried out in Department of Human Sciences, Loughborough University of Technology.

EYSENCK, N.J. and RACHMAN, S. (1965) 'Dimensions of Personality'
From 'The Causes and Cure of Neurosis'.
Routledge and Kegan-Paul.

FELL, J. and
LAUGHERY, K. (1969) Short term Memory; Mode of Presentatio
for Alphanumeric Information.
Human Factors, 11, 401 - 415.

GAINES and FACEY (1975) Some Experience in Interactive System
Development and Application.
Proceedings of the IEEE, 63, 894 - 911

GANDELSONAS, M. (1974) Linguistic and Semiotic Models in
Architecture.
in 'Basic Questions of Design Theory'.
39 - 54. Editor: William Spillers.
McGraw-Hill.

GEISSER, S. and
GREENHOUSE (1958) An Extension of Box's results on the u
of the F distribution in Multivariate
Analysis.
Anals of Mathematical Statistics, 29,
885 - 891.

GRIGNETTE, M.C. and
MILLER, D.C. (1968) Modifying Computer Response
Characteristics to Influence Command
Choice.
IEEE Conference Publication on Man-
Computer Interaction, 68, 201 - 205.

HERSHMAN, R.L. and
HILLIX, W.A. (1965) Data Processing and Typing.
Human Factors, 7, 483 - 492.

HOPGOOD, F.R.A. (1969) Compiling Techniques.
Computer Monographs, Macdonald/Elsevie

HUNTER, F.M.L. (1964) Memory.
Penguin Books.

HORNBUCKLE, G. (1967) Computer Graphics/User Interactions.
IEEE Transactions on Human Factors in
Electronics, HFE-8, 17 - 20.

HOWDEN, W.F. (1975) Solution Plans and Interactive Problem
Solving.
Computers and Graphics, 1, 21 - 26.

INNOCENT, P.R. (1973) Calling the Tune in Time Sharing System
Digital Equipment Company Users Society,
(DECUS), Conference Proceedings, 9,
191 - 196.

INNOCENT, P.R. (1973) Human Time Sharing with a Dedicated
Interactive Computer System.
Datafair Conference Proceedings, 2,
392 - 398. British Computer Society
Publication.

INNOCENT, P.R. (1973) Studies of the Ease of Use and Learning
of On-Line Computer Systems.
NPL Contract Report No. SN/1170/015.

INNOCENT, P.R. (1976) Concepts of Computers.
British Ship Research Association
Report - Memorandum No. 11 of Ship
Structural Design System (SSDS)Project.

INNOCENT, P.R. (1976) Individual Differences - A Model.
British Ship Research Association
Internal Note.

KAMLETT and MUSA (1974) Human Factors Testing of Computer Syste
Bell Systems House Journal, September,
245 - 251.

KAY, H. (1957) Information Theory in the Understanding
of Skills.
Occup. Psychology, 31, 219 - 224.

KAY, H. (1954) The effects of position in a display
upon problem solving.
Quarterly Journal of Experimental
Psychology, 6, 155 - 159.

KIRK, R.E. (1968) Experimental Design for Behavioural
Scientists.
Brooks/Cole,

KLEMMER, E.T. (1962) Communication and Human Performance.
Human Factors, 4, 75 - 79.

KOLESNICK, P. and TEEL, K.S. (1965) A Comparison of 3 Manual Methods of Inputting Navigational Data. Human Factors, 7, 451 - 456.

KOSTER, W.G. (1969) Attention and Performance. North-Holland Pub. Co.Ltd.

LEWIS, D. (1960) Quantitative Methods in Psychology. McGraw-Hill.

LONDON, K. (1976) The People Side of Systems. McGraw-Hill.

MARKS, M. (1951) Problem Solving as a Function of the Situation. Journal of Experimental Psychology, 55, 74 - 86.

MEISTER, D. (1967) Development of Human Reliability Indices. Proceedings of the Symposium on Human Performance and Quantification in Systems Effectiveness. Naval Material Command, Washington D.C.

MILLER, P.K. (1970) Man-Machine Symbiosis. IEEE 17th Annual Conference on Petroleum & Chemicals. Paper 70-4, 14-2

MILLER, G.A. (1967) The Psychology of Communication. Pelican Books.

MILLS, R.G. (1967) Man-Machine Communication and Problem Solving. Annual Review of Information Science and Technology, 2, 223 - 253.

NAWROCKI, L.H. (1972) Alphanumeric versus Graphic Displays i a Problem Solving Task. NTIS US. Army Report, Behaviour and Systems Res. Lab. Technical Research Note 227.

NEISSER et al. (1963) Search for alphanumeric targets on displays.
Perceptual and Motor Skills, 17, 955 - 961.

NICKERSON, R.S. and PEW (1971) Oblique steps towards the Human Factors Engineering of interactive computer systems.
ARPA, DOD, AFOSRC F44620-71-C-0065.

PALME, J. (1975) Interactive Software for Humans.
FOA-1 Report, Swedish National Research Defense Institute.

POSNER, M.I. and BOIES, S.J. (1971) Components of Attention.
Psychological Review, 78, 391 - 408.

POULTON, E.C. (1973) Unwanted Range Effects from using within subject experimental designs.
Psychological Bulletin, 30, 113 - 121.

POULTON, E.C. and FREEMAN, P.F. (1966) Unwanted Asymmetrical Transfer Effects with balanced Experimental Designs.
Psychological Bulletin, 66, 1 - 8.

RATZ, H.C. and RITCHIE, D.K. (1961) Operator Performance on a Chord Keyboard.
Journal of Appd. Psychology, ,45, 303-3

ROOT and SADAACA. (1967) Man-Computer Communication Techniques.
Human Factors, 9, 521 - 528.

ROSENTHAL ,R. (1967) Covert Communication in the Psychological Experiment.
Psychological Bulletin, 67, 356 - 367.

SACKMAN, H. and CITRENBALM, R.L. (1972) On-line Planning towards Creative Problem Solving.
Prentice-Hall Inc., Englewood Cliffs, New Jersey.

SACKMAN, H. (1970) Man-Computer Problem Solving.
Princeton: Auerbach Publishers.

SACKMAN, H. (1967) Computers, Systems Science and Evolving Society.
Wiley and Sons Ltd., New York.

SACKMAN, H. (1976) Towards a General Theory of Man-Computer Problem Solving.
Proceedings of NATO Advanced Study Institute Conference on Man-Computer Interaction, Mati, Greece.

SCHOONARD, J.W. and BOIES, S.J. (1975) Short type - A Behavioural Analysis of Typing and Text Entry.
Human Factors, 17, 203 - 214.

SHEPHERD, R.D. (1965) Search for Alphanumeric targets on displays.
Occupational Psychology, 39, 191 - 203.

SIMANIS, A. (1975) Human Factors in Interactive Graphics.
Proceedings of the 4th Conference on Man-Computer Communications, Ottawa.
8 - 3 to 8 - 12.

SMITH, W.A. (1967) Accuracy of Manual Entries in data collection devices.
Journal of Appd. Psychology, 51, 362-36

SOLLEY, C.M. and SNYDER, F.W. (1958) Information processing and problem solving.
Journal of Experimental Psychology, 55, 384 - 87.

STERN, G. (1970) People in Context.
J Wiley and Sons, New York.

TAYLOR, R.W. (1975) On the Relation of Interactive Computin to Computer Science.
Proceedings of the IEEE, 63, 843 - 846.

TODD, M.J. (1974) Simulation of a User Terminal Interactive Situation.
British Ship Research Association, Internal Report.

TULLY, C.J. (1974) The Unsolved Problem - A New Look at Computer Science.
Computer Bulletin, Series 2, 2, 26 - 28

WASON and JOHNSON-LAIRD(Editors) (1968) Thinking and Reasoning.
Penguin Modern Psychology Readings.

WATSON, R.W. (1976) User Interface design issues for a larg
interactive system.
AFIPS Conference Proceedings, 45, 357 - 364.

WEINBERG, G.H. (1971) The Psychology of Computer Programming.
Van Nostrand Reinhold Co., New York.

WELFORD, A.T. (1963) Stress and Performance.
Ergonomics, 16, 567 - 572.

WELFORD, A.T. (1962) Arousal, Channel Capacity and Decision.
Nature, 194, 365 - 366.

WRIGHT, D.S. and TAYLOR, A. (Editors.) (1970) Introducing Psychology:
An Experimental Approach.
Penguin Books.

17. Appendices

APPENDIX 1

DATA AND INFORMATION TABLES OF CHAPTER 5

FIGURES

TABLES

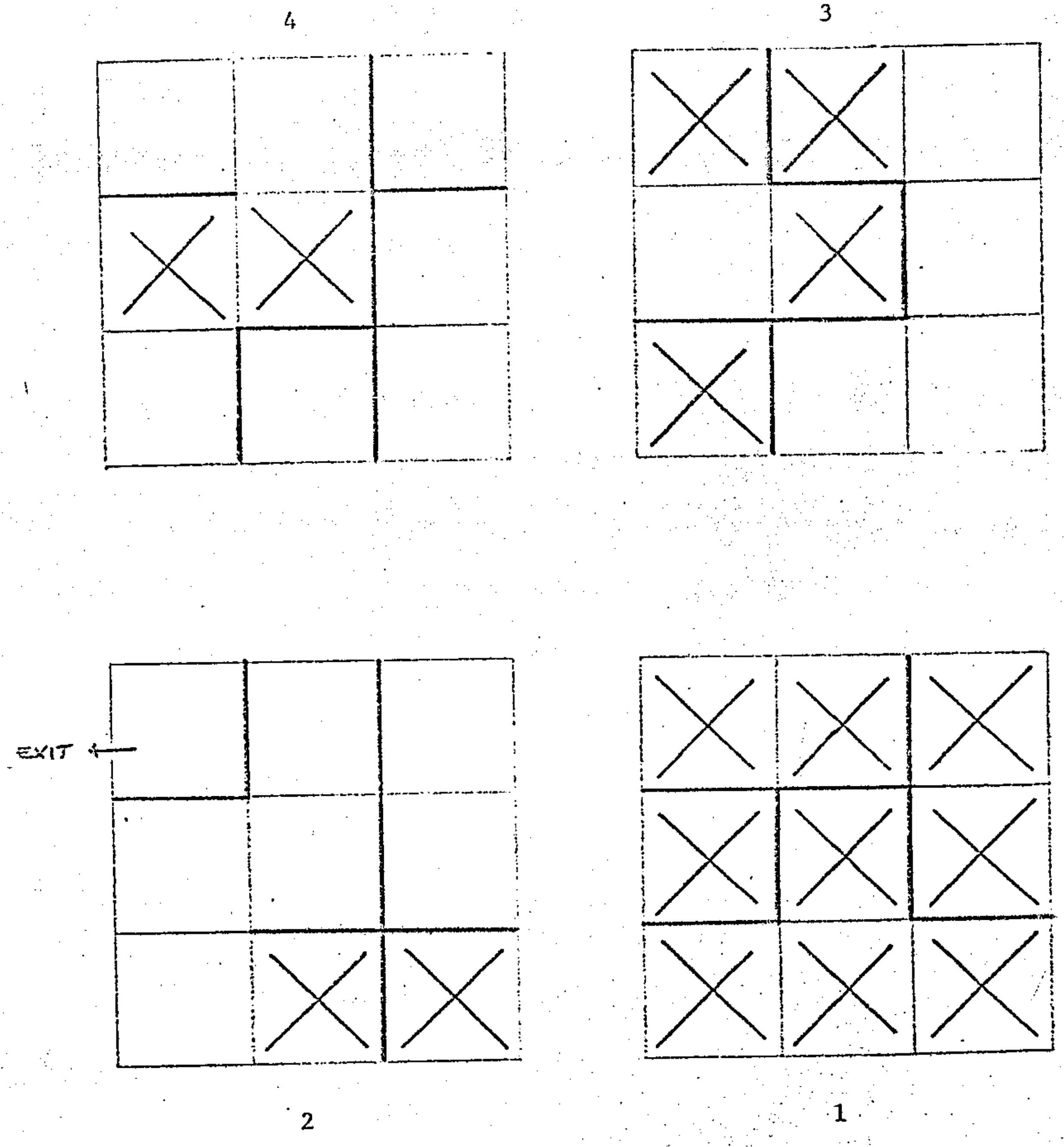
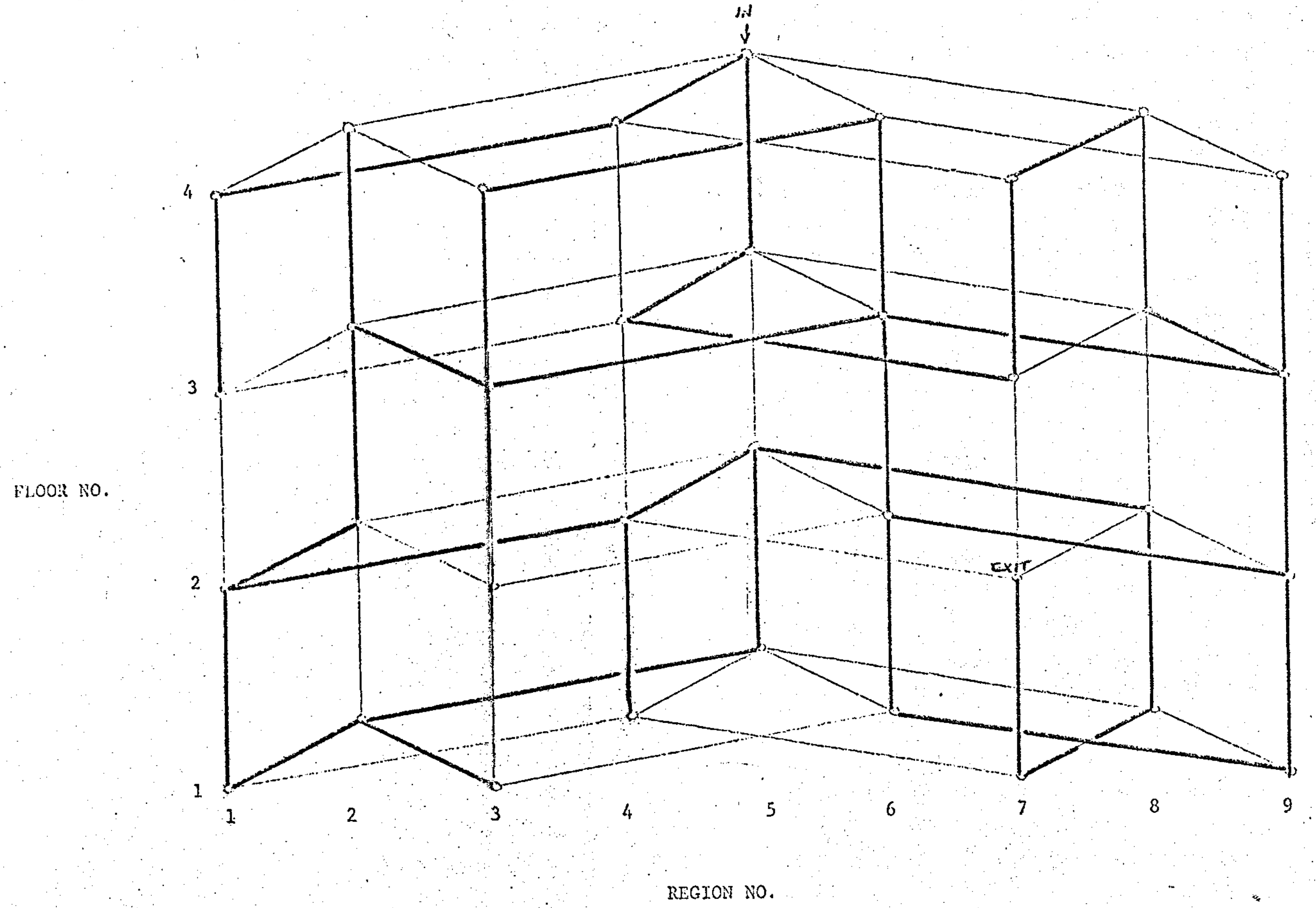


Figure 1. MAZE 1


IN
4
3
2
1
FLOOR NO.
EXIT
1
2
3
4
5
6
7
8
9
REGION NO.

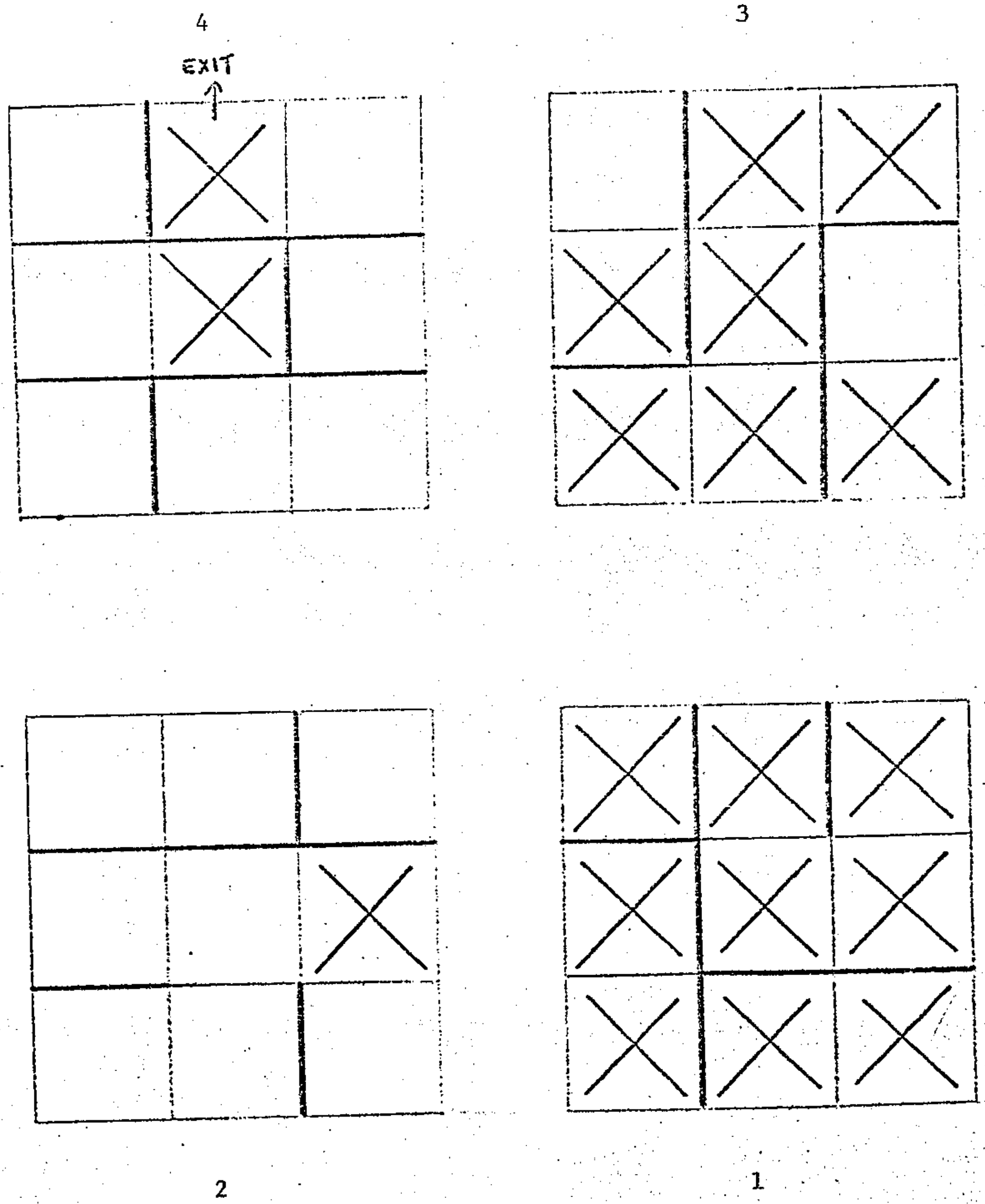
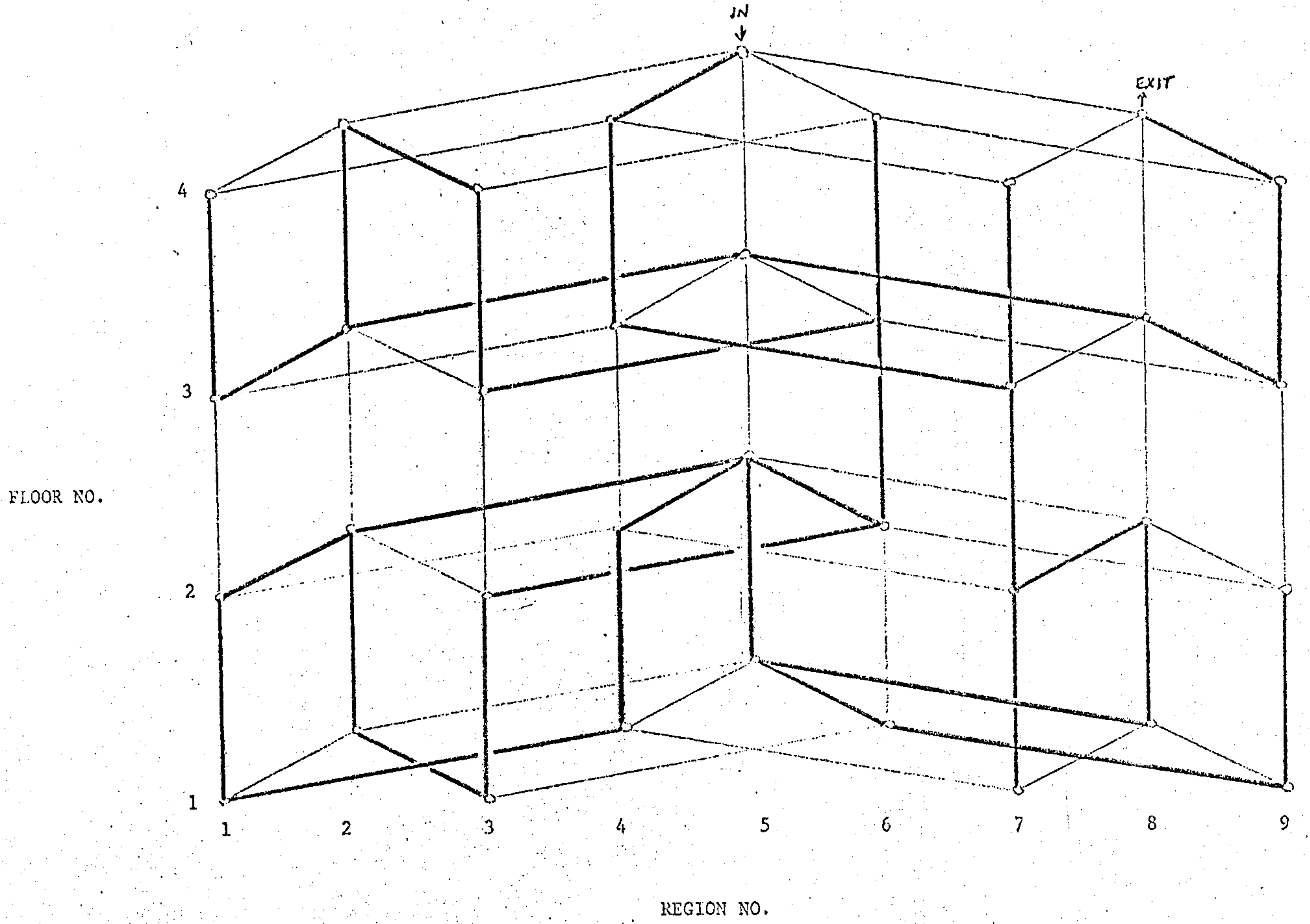


Figure 3. MAZE 2


IN
EXIT
4
3
2
1
FLOOR NO.
1
2
3
4
5
6
7
8
9
REGION NO.

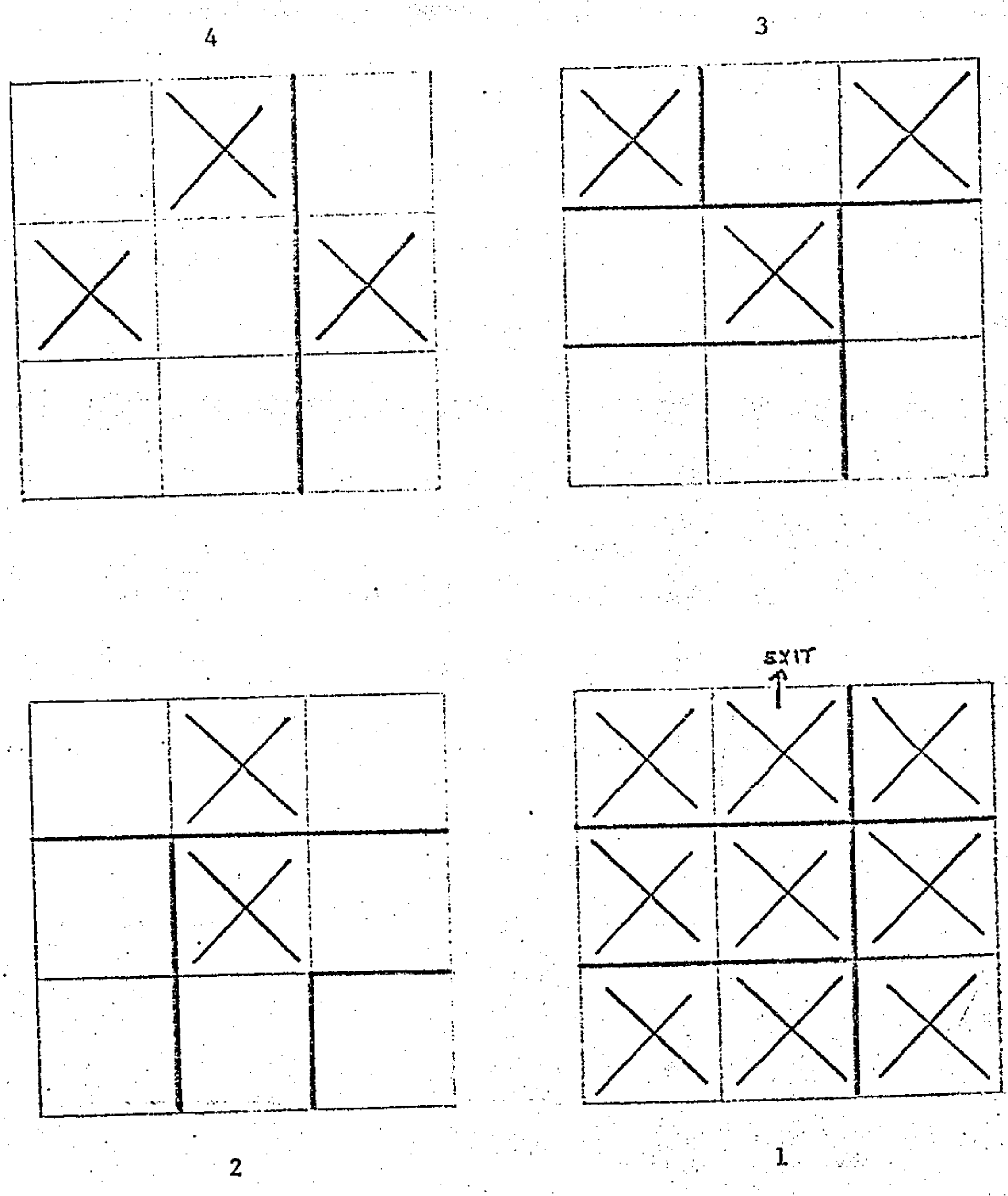
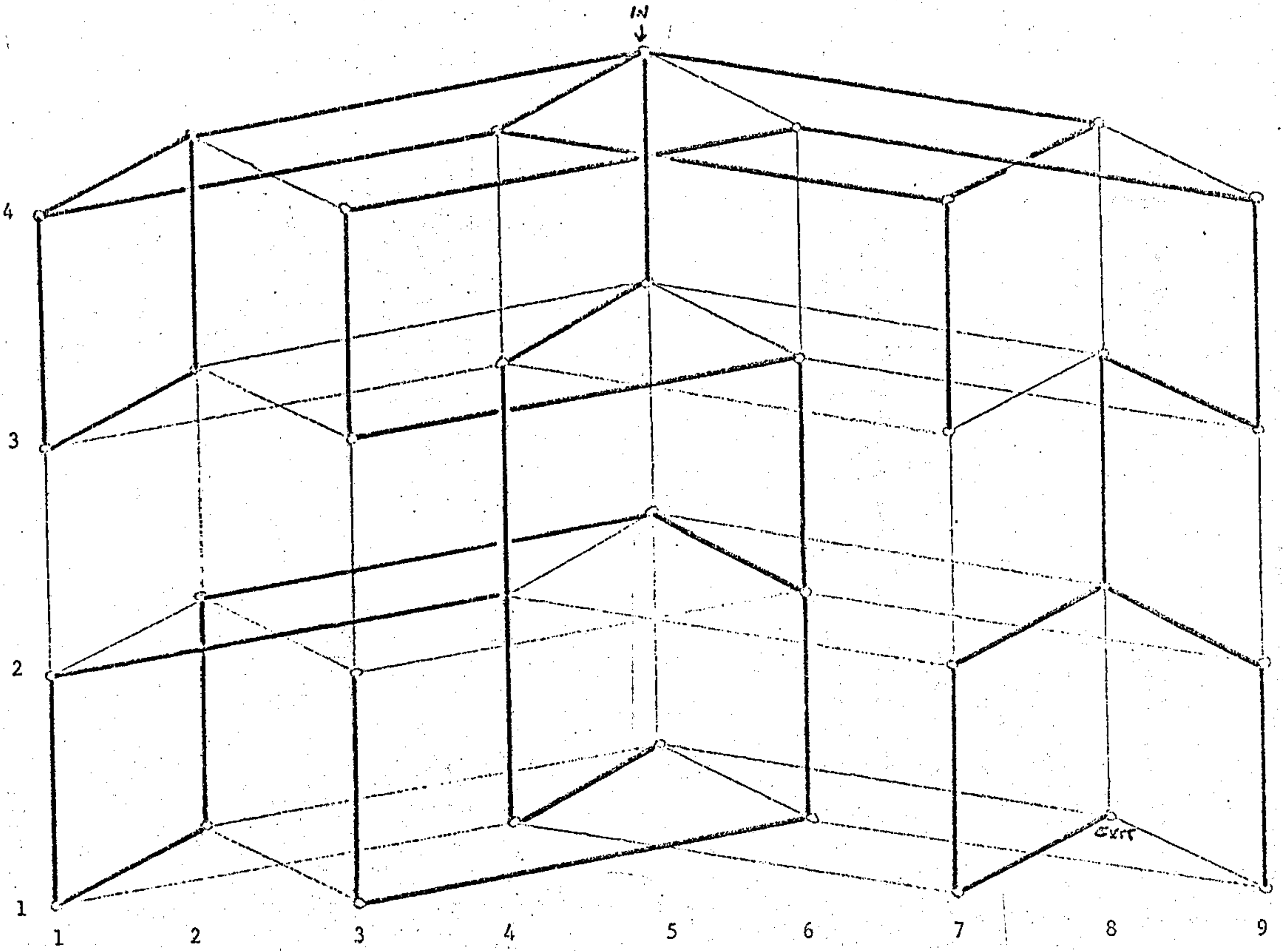


Figure 5. MAZE 3



FLOOR NO.

REGION NO.

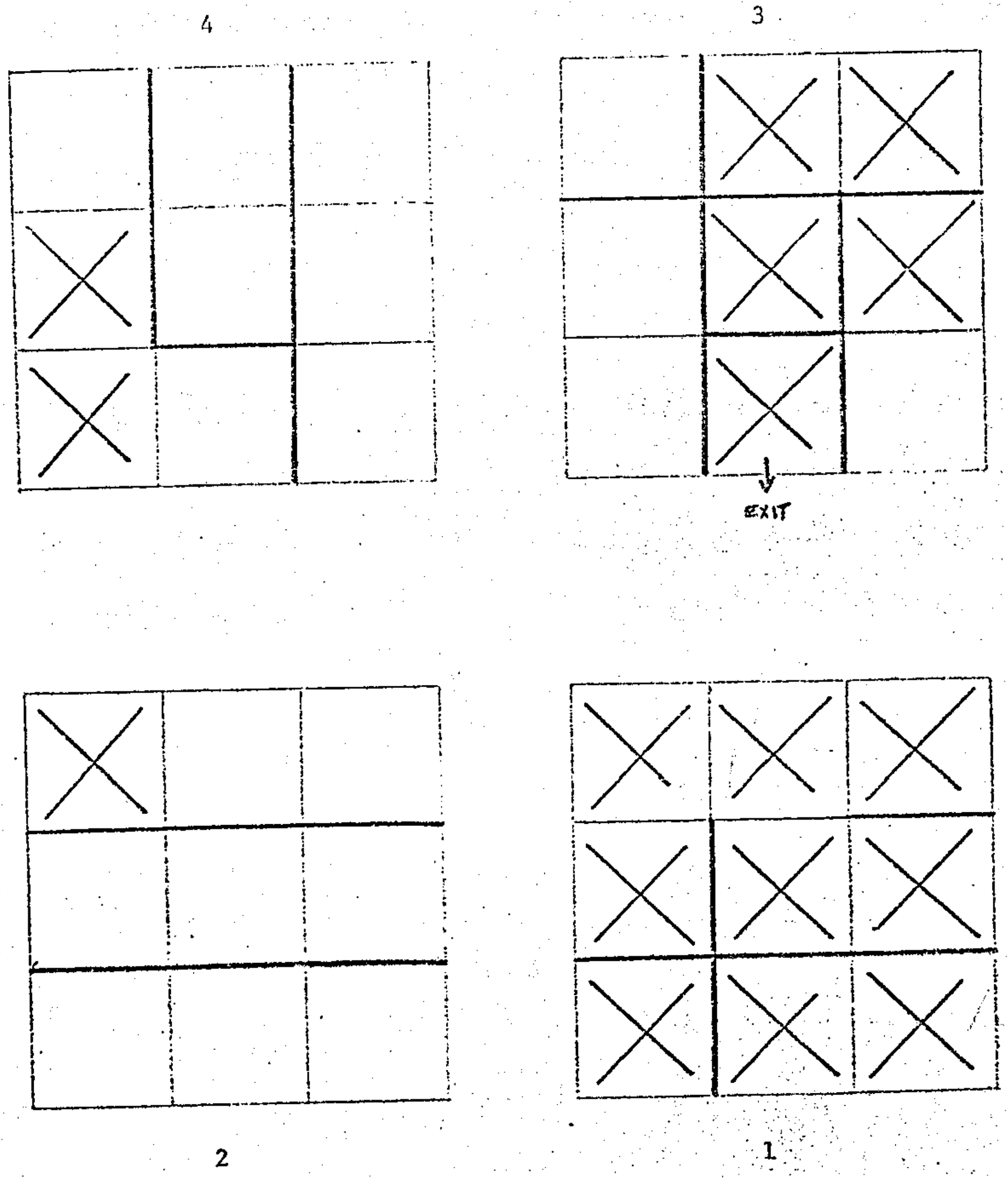
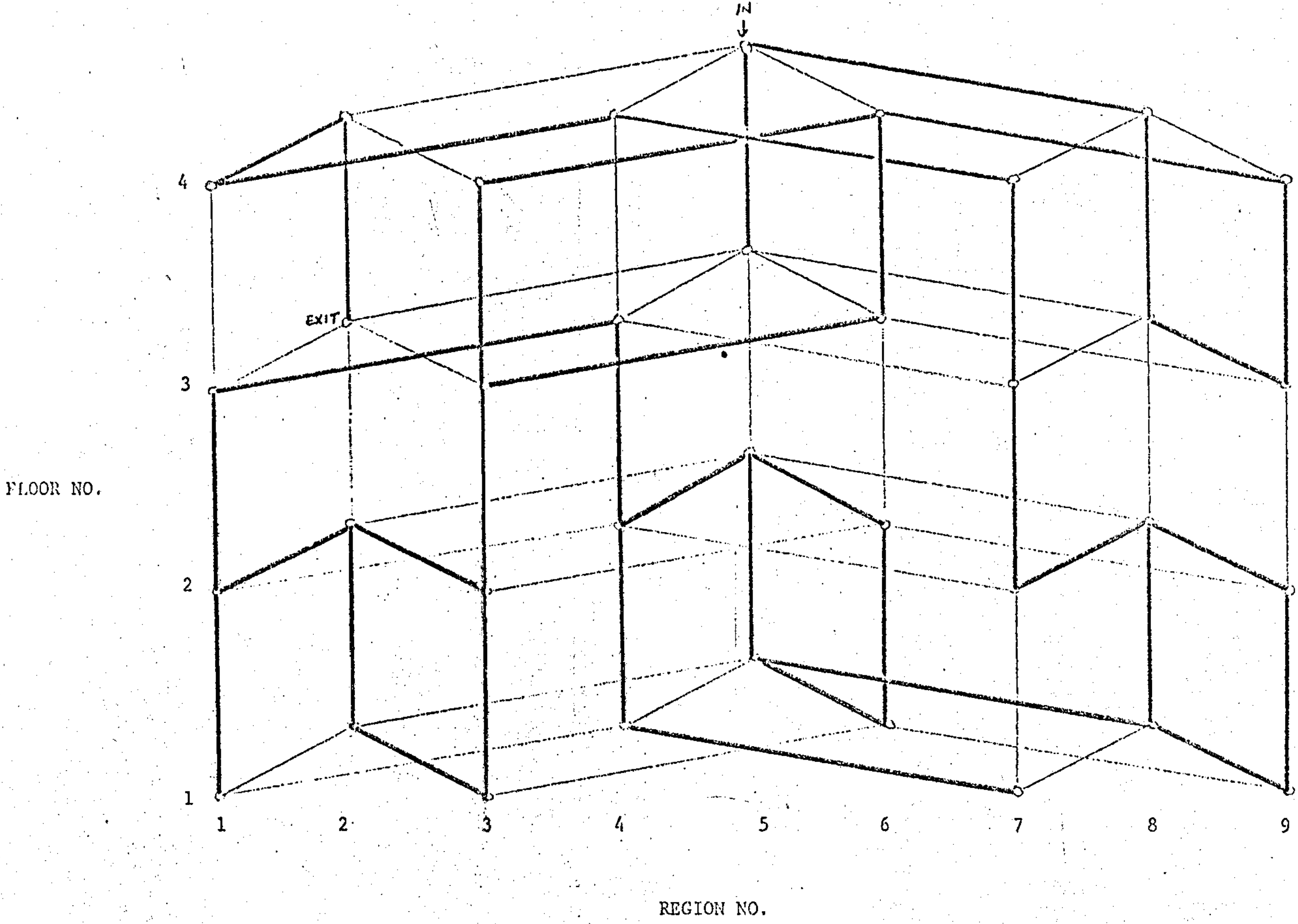


Figure 7. MAZE 4



2-D REPRESENTATION OF ALL POSSIBLE MAZES

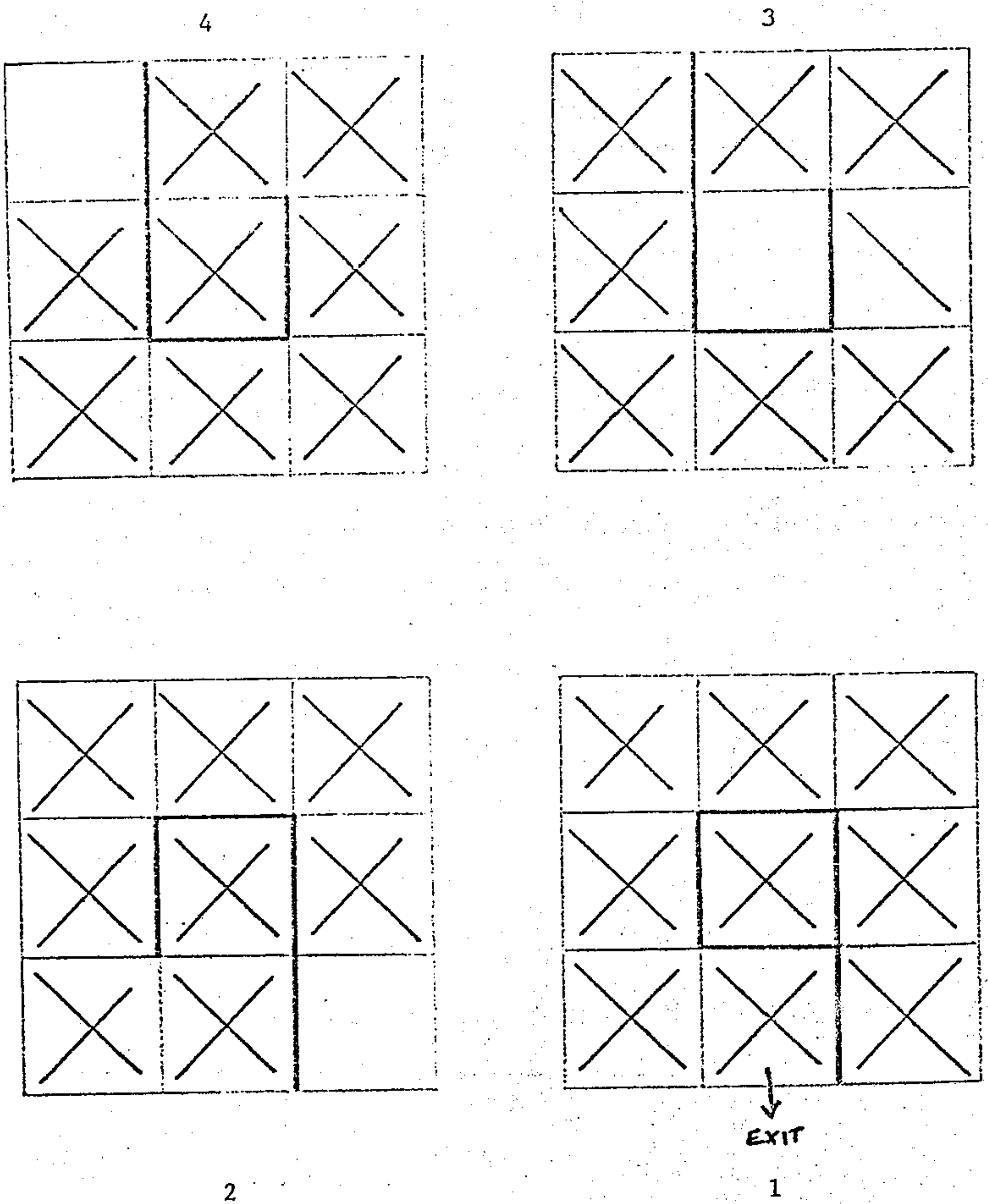
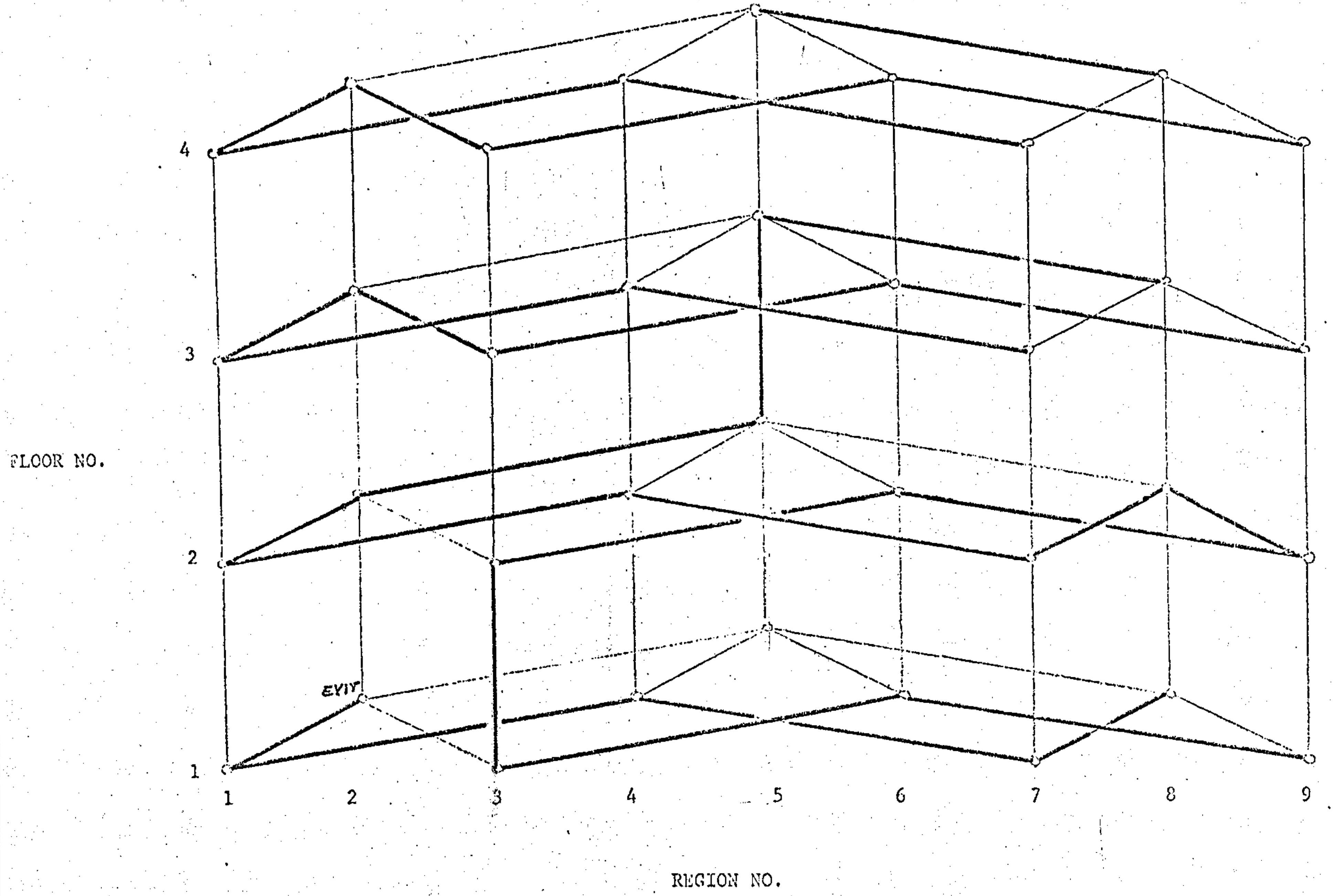


Figure 9. MAZE 5

(Training Maze)



2-D REPRESENTATION OF ALL POSSIBLE MAZES

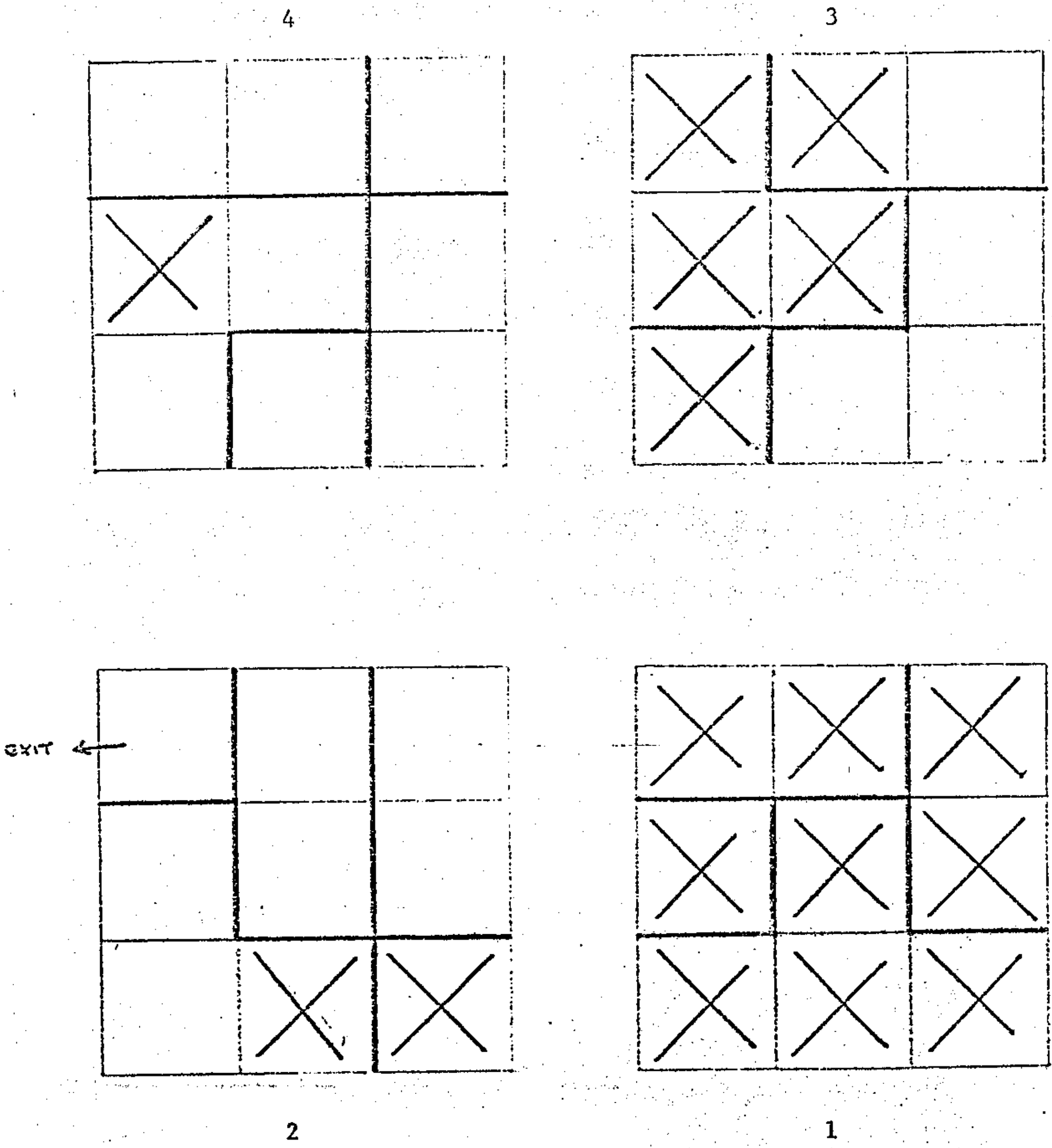
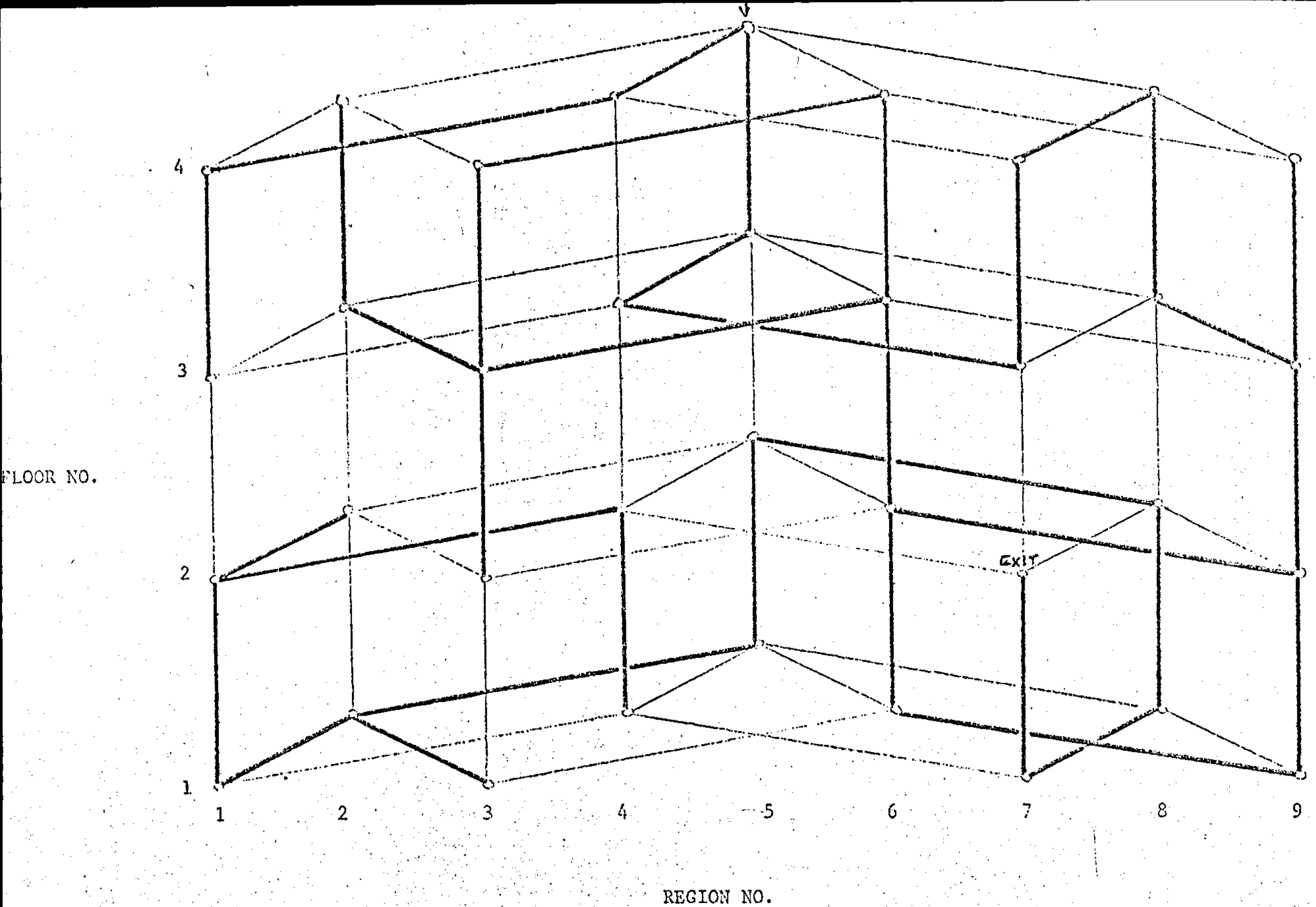


Figure 11. PILOT MAZE (MAZE 6)



2-D REPRESENTATION OF ALL POSSIBLE MAZES Figure 12. PILOT MAZE (MAZE 6)

Table 1 - Spiral Maze Results

	JOYSTICK				KEYBOARD											
					→ ← ↑ ↓				U,D,N,S,E,W,M,A,R				MOVE EAST, GO DOWN			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
S1	37	0.79	0.36	1	56	1.33	0.88	1	35	1.59	1.16	0	435	6.14	2.40	0/0
S2	35	2.16	1.18	0	35	1.87	1.51	0	35	1.84	1.34	0	435	6.47	2.61	0/1
S3	27	1.02	0.82	1	56	0.97	0.66	2	35	1.42	1.57	0	435	5.41	2.26	01
S4	35	1.21	2.36	0	35	1.44	1.23	0	35	1.47	0.96	0	50	5.67	1.17	1/1
S5	35	0.74	0.92	0	37	1.09	0.66	0	35	1.04	0.89	0	59	5.52	1.81	1/0
S6	35	1.24	1.26	0	69	1.39	1.28	1	35	1.54	0.93	0	435	11.69	4.57	0/5
S7	127	1.32	1.09	5	35	1.45	0.93	0	69	1.82	1.07	1	435	9.99	5.27	0/0
S8	35	1.51	0.85	0	35	1.09	0.74	0	35	2.00	1.47	0	104	7.31	2.27	6/8
S9	35	1.17	1.23	0	61	1.17	1.11	1	35	1.48	1.09	0	41	7.93	3.55	1/3
S10	69	0.75	0.47	1	35	0.99	0.72	0	35	0.84	0.51	0	35	3.48	0.66	0/0
S11	70	0.83	0.51	2	60	1.12	0.80	1	35	1.42	0.82	0	35	9.56	3.14	0/0
S12	35	1.45	1.54	0	59	1.06	0.65	1	35	1.38	1.26	0	54	5.53	1.61	1/1
S13	216	1.65	2.35	16	35	2.07	2.38	0	35	2.38	3.69	0	104	3.85	1.49	2/26
S14	209	2.50	2.16	11	35	1.82	1.42	0	35	1.62	1.66	0	142	9.68	5.86	3/41
S15	146	0.66	0.35	4	35	1.21	0.62	0	35	1.46	1.11	0	35	12.2	22.4	0/6
S16	94	2.08	2.95	4	35	1.75	1.54	0	35	1.40	0.89	0	35	7.11	5.05	0/3
				3.24%				2.61%				0.25%				1.8%/ 11%

KEY:

Column	Contents
1	No. of Moves through spiral to get out.
2	Mean time per move through spiral (Maze 5) in seconds for successful trial.
3	S.D. associated with 2.
4	No. of errors in trials/No. of typing errors detected before input entered.

NOTE: N = 35 for all samples of mean time

	JOYSTICK				KEYBOARD											
					→ ← ↑ ↓				U,D,N,S,E,W,M,A,R				MOVE EAST, GO DOWN			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
S1	3					12					10					41
S2		9					5					4	1			
S3			4					1	1					1		
S4				8	1					1					1	
S5				8			2			7			1			
S6	1							2			2			7		
S7		4			2							1			1	
S8			3			7			16							19
S9				4			11			23			2			
S10	1							6			4			6		
S11		2			2							3			1	
S12			5			1			1							8
S13	3					2					3					3
S14		8					20					23	5			
S15			2					5	1					2		
S16				4	2					1					1	

Table 2 NO. OF TRIALS

	JOYSTICK				KEYBOARD											
					→ ← ↑ ↓				U,D,N,S,E,W,M,A,R				MOVE EAST, GO DOWN			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
S1	1.42					24.75					4.04					42.27
S2		16.79					13.1					19.91	2.94			
S3			3.2					2.31	1.209					6.13		
S4				12.63	1.32					3.80					4.72	
S5				8.22			2.59			4.36			2.08			
S6	5.38							15.06			9.70			38.9		
S7		15.31			2.56							16.59			12.06	
S8			3.48			10.59			15.63							58.51
S9				9.45			9.65			36.12			3.96			
S10	1.19							6.65			4.23			5.50		
S11		11.13			3.30							12.40			9.67	
S12			5.78			4.10			1.62							16.65
S13	4.91					18.32					10.97					22.15
S14		10.58					23.44					44.24	7.79			
S15			10.8					23.64	2.84					12.73		
S16				9.15	5.63					2.98					7.76	

Table : 3 OVERALL TIMES
(Minutes) to get out of the maze.

	JOYSTICK				KEYBOARD											
					→ ← ↑ ↓				U,D,N,S,E,W,M,A,R				MOVE EAST, GO DOWN			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
S1	58					345					30					178
S2		68					24					44	10			
S3			50					24	14					27		
S4				228	10					21					16	
S5				159			35			57			13			
S6	10							24			33			63		
S7		29			10							24			16	
S8			29			52			99							188
S9				113			137			251			11			
S10	12							149			43			57		
S11		21			20							58			18	
S12			73			19			30							88
S13	12					28					46					24
S14		193					271				[illegible]	527	20			
S15			196					175	34				20	62		
S16				49	26					19					18	

Table 4 TOTAL MOVES

	JOYSTICK				KEYBOARD											
					→ ← ↑ ↓				U,D,N,S,E,W,M,A,R				MOVE EAST, GO DOWN			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
S1	0					4					2					28
S2		9					4					3	1			
S3			2					1	1					1		
S4				2	1					1					1	
S5				3			1			4			1			
S6	1							2			1			5		
S7		3			1							1			1	
S8			2			6			9							15
S9				1			2			11			1			
S10	1							3			2			4		
S11		1			1							2			1	
S12			1			1			0							6
S13	3					2					1					3
S14		7					17					21	4			
S15			1					2	0					0		
S16				2	1					0					1	

Table 5 NO. OF PLANS/AIDS USED

	JOYSTICK				KEYBOARD											
					→ ← ↑ ↓				U,D,N,S,E,W,M,A,R				MOVE EAST, GO DOWN			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
S1	18					31					11					55
S2		4					3					6	2			
S3			4					3	1					6		
S4				9	3					3					1	
S5				7			5			7			0			
S6	1							1			4			3		
S7		5			2							0			1	
S8			0			7			10							42
S9				4			1			24			2			
S10	3							7			0			7		
S11		7			3							5			3	
S12			4			2			7							15
S13	2					12					9					6
S14		12					20					40	6			
S15			8					30	5					10		
S16				9	8					5					3	

Table 6 NO. OF RECALLS/EVALUATIONS

	JOYSTICK				KEYBOARD											
					→ ← ↑ ↓				U,D,N,S,E,W,M,A,R				MOVE EAST, GO DOWN			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
S1	4.21					25.8					27.32					31.7
S2		121.1					218.1					842.5	89.3			
S3			72.65					75.45	26.71					151.0		
S4				93.0	30.5					155.8					157.2	
S5				100.9			57.39			43.05			52.6			
S6	98.6							381.6			413.6			191.2		
S7		268.24			122.0							897			504	
S8			82.6			73.3			102.3							83.9
S9				334.8			150.8			81.99			121.5			
S10	38.1							43.27			67.25			24.56		
S11		516.5			143.6							501.3			368.1	
S12			371.9			204.8			0							62.29
S13	47.2					448.7					338.9					275
S14		28.7					58.2					107.2	40.71			
S15			28.3					145.1	0					0		
S16				282.4	103.2					35.3					268.9	

Table 7 MEAN PLANNING TIME (SECONDS)

	JOYSTICK				KEYBOARD											
					→ ← ↑ ↓				U,D,N,S,E,W,M,A,R				MOVE EAST, GO DOWN			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
S1	10					75.1					6.6					8.3
S2		6.4					5.2					12.3	18			
S3			23					21	13					21		
S4				82.5	7					18					15	
S5				48.6			31			11			10			
S6	10							12			29			11		
S7		8			8							24			15	
S8			9.66			7.6			10.4							9.4
S9				55.0			78.3			19.7			8			
S10	11							45.5			23			12.7		
S11		7			17							26			15	
S12			14			17			0							16.4
S13	33					8					19.8					6
S14		25.6					14.5					24.2	4.5			
S15			28					38	0					0		
S16				17.5	10					0					15	

Table 8 MEAN NO. OF MOVES BETWEEN AID AND AID

	JOYSTICK				KEYBOARD											
					→ ← ↑ ↓				U,D,N,S,E,W,M,A,R				MOVE EAST, GO DOWN			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
S1	1/3					70/12					0/ 3					19/26
S2		1/9					1/5					2/4	0/1			
S3			2/4					0/1	0/1					0/1		
S4				26/8	0/1					0/1					0/1	
S5				14/8			0/2			4/7			0/1			
S6	0/1							0/2			1/2			4/7		
S7		1/4			0/2							0/1			0/1	
S8			3/3			5/6			15/14							19/21
S9				12/3			10/20			39/20			0/1			
S10	0/1							15/6			0/23			4/6		
S11		0/2			1/2							2/3			0/1	
S12			19/5			0/1			0/1							4/8
S13	0/3					1/2					0/3					0/3
S14		30/8					71/20					97/23	0/5			
S15			3/2					5/5	0/1					2/2		
S16				1/4	1/2					0/1					0/1	

Table 9 MEAN ERRORS PER TRIAL

	JOYSTICK				KEYBOARD											
					→ ← ↑ ↓				U,D,N,S,E,W,M,A,R				MOVE EAST, GO DOWN			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
S1	1.44					4.42					5.52					6.667
S2		2.56					2.96					3.61	5.88			
S3			2.04					2.41	2.04					7.87		
S4				2.37	3.21					2.93					7.38	
S5				1.77			2.57			2.09			6.72			
S6	2.64							4.54			3.11			10.97		
S7		1.98			2.05							3.77			10.72	
S8			2.20			2.24			2.25							6.05
S9				1.73			1.89			3.23			8.27			
S10	1.17							1.52			3.05			3.43		
S11		1.51			1.69							3.05			9.44	
S12			0.79			2.63			1.85							5.19
S13	3.58					3.69					4.02					13.01
S14		2.07					2.55					3.33	7.46			
S15			2.69					3.01	3.32					8.52		
S16				3.9	2.47					2.16					7.75	

Table 10 MEAN TIME PER MOVE (seconds)

	JOYSTICK				KEYBOARD											
					→ ← ↑ ↓				U,D,N,S,E,W,M,A,R				MOVE EAST, GO DOWN			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
S1	2.61					8.71					11.27					15.86
S2		14.94					12.66					29.13	13.99			
S3			9.2					6.11	17.35					13.07		
S4				12.55	9.84					8.04					11.22	
S5				20.16			6.88			5.35			0			
S6	0							18.00			17.28			35.11		
S7		10.79			5.95							0			36.11	
S8			0			5.91			6.62							16.98
S9				8.91			20.9			14.73			16.89			
S10	6.67							4.98			0			6.95		
S11		13.68			6.94							11.95			17.33	
S12			6.25			4.15			7.51							13.82
S13	6.16					10.66					16.2					33.5
S14		7.86					9.26					16.43	19.07			
S15			13.03					10.94	14.11					30.34		
S16				15.86	8.97					9.5					21.65	

Table 11. MEAN TIME PER RECALL/EVALUATION (SECONDS)

Subject No.	Comments
1	Used the mark facility to mark path and to mark the room use. Also used the aid and forgot errors on verbose in
2	Used aid to locate exit then worked back about 3 moves; developed a vivid visual model of a building; used idea lifts rather than staircases.
3	Used aid to plan then relied on visual-spatial memory.
4	No marks because of difficulty of remembering the meanin Worked back from exit then forwards from entrance.
5	No internal model; ignored floors of little choice, but remembered floors of many; if error, immediate restart; saw a digital space.
6	Remembered mnemonics from the aid; no 3-D model.
7	Memorised each floor; cues from spatial recognition; staircases confusing - used a plan mode, no 3-D.
8	Memorised first five or six moves then forgot first step no 3-D model. Found down-up-down idea difficult.
9	Mark could be ambiguously used; visual model; forward and back strategy; didn't restart after errors to get feedback on current path.
10	Conceptually difficult; goal oriented strategy; no mar or 3-D model.
11	Used relative position in the maze to remember path by learning only short steps within floors.
12	Tried to remember only the entrance and exit points of floors; developed cube model.
13	Had problems with up-down-up idea.
14	Mazes all spiral variants; used trial and error on maze goal strategy on others; 3-D model; joystick incompatibility with motion on display.
15	No use of aid, look through floors by going down through all; trial and error learning; no mental model.
16	3-D model developed; used aid for multiple choice situations.

Table 12

SUBJECT COMMENTS

S. No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
N-Score	8	16	11	13	5	3	1	8	0	6	12	9	2	2	5	10
E-Score	6	16	7	12	7	16	15	18	13	9	18	8	12	14	8	7
Lies	3	4	0	0	0	2	1	3	2	1	1	2	5	4	5	1

Table 13. Esyenck Personality Inventory (EPI) Measures for all Subjects.

Category	Upper Bound (secs)	MAZES				INPUTS				MAZE 5			
		1	2	3	4	A	B	C	D	A	B	C	
1	0.100	16	28	48	29	77	17	14	13	115	12	10	
2	0.147	6	0	10	0	10	4	0	2	7	1	2	
3	0.215	1	3	13	17	15	11	7	1	14	5	11	
4	0.316	1	6	7	26	17	13	9	1	25	17	27	
5	0.464	11	4	26	51	48	26	18	0	79	76	58	
6	0.681	6	21	26	84	63	51	23	0	196	113	90	
7	1.000	19	41	54	120	129	62	41	2	279	70	64	
8	1.468	46	156	119	222	231	166	141	5	258	179	101	
9	2.154	89	252	166	281	235	280	255	18	199	127	160	
10	3.162	73	261	210	285	218	289	291	31	112	56	102	
11	4.642	50	194	168	217	136	229	252	112	50	23	29	1
12	6.813	47	181	102	331	92	134	181	254	18	5	8	3
13	10.000	35	105	77	215	51	74	115	192	19	11	7	2
14	14.68	25	76	45	136	24	41	50	167	7	4	2	1
15	21.54	14	32	23	73	16	15	23	88	7		3	
16	31.62	6	17	6	35	6	6	15	37	3		1	
17	46.42	2	19	2	16	5	6	9	19			1	
18	68.13	11	6	3	13	0	4	6	13				
19	100.00	3	6	3	8	5	5	5	5				
20	146.8	4	4	2	11	3	4	2	12				
21	215.4	3	8	2	3	5	4	3	4				
22	316.2	22	2	4	2	3	3	2	2				
23	464.2		2	4	2	1	3	3	1				
24	631.3		4	1	4	2	2	1	4				

Table 14. Frequency Distribution of Input Information

ASURE	SOURCE	DF	SSQ	MSQ	F	M	MEASURE	SOURCE	DF	SSQ	MSQ	F
PRT	S	15	86.67	5.77	10.62*	9	ERRS	S	15	9900	660	2.96
	I	3	423.5	141.2	260*			I	3	603.125	441.2	1.98
	ERROR	45	24.43	0.54				E	45	1002.8	222.8	
	TOTAL	63	534.6					T	63	20531		
NOTS	S	15	1335.7	89.05	2.45	10	TRM	S	15	537.5	35.84	–
	I	3	58.2	19.4	–			I	3	960.8	320.26	5.38
	E	45	1637	36.39				E	45	2679	59.5	
	T	63	3031.6					T	63	4177		
OVTM	S	15	2268.7	151.24	1.22	11	TMM	S	15	182.35	12.16	9.06
	I	3	511.3	170.4	1.38			I	3	332.0	110.6	79.3
	E	45	5575	123.9				E	45	62.75	1.39	
	T	63	8355					T	63	577.1		
OTMOV	S	15	232592	15506	2.23	5	NOPLANS	S	15	653.2	43.5	2.12
	I	3	12655	4218	–			O	3	41.55	13.84	–
	E	45	312759	6950				E	45	921.1	20.46	
	T	63	558006					T	63	1615.8		
PLANS×	S	15	653.25	43.5	2.55	8	MMAA	S	15	7071	471	2.3
	I	3	167.63	55.8	3.16*			O	3	1036.7	345.2	1.68
	E	45	794.9	17.66				E	45	9210.3	204.67	
	T	63	1615.8					T	63	17318		
ORCLS	S	15	3195.9	213.06	2.54	10	TRM	S	15	537.5	35.84	–
	I	3	136.0	45.3	–			O	3	908.25	302.75	5.05
	E	45	3766	83.69				E	45	2693.6	59.85	
	T	63	7098.4					T	63	4139.4		
MPT	S	15	1049495	69966	2.96	11	TMM	S	15	182.35	12.15	1.4
	I	3	76460	25486	1.08			M	3	5.18	1.73	–
	E	45	1061629	23591				E	45	389.5	8.6	
	T	63	2187584					T	63	577.1		
MMAA	S	15	7071	471	2.5	8	MMAA	S	15	7071	471	2.2
	I	3	1756	585.33	3.1*			M	3	624.27	208.9	–
	E	45	8491	188.7				E	45	9622.7	213.84	
	T	63	17318					T	63	17318		
						11	TMM	S	15	182.35	12.15	–
								O	3	3.24	1.08	–
								E	45	391.5	8.700	
								T	63	577.1		

15 - 2-WAY ANOVA RESULTS

Input Method E = Error
Subjects T = Total
Maze * = 5% level significance
Order × Non-normal data.

t values between Mean No. of Moves Between Use of Aid (MMAA)

	A	B	C	D
A		0.04	1.22	2.15*
B			1.54	2.52*
C				1.45
D				

t values between Mean times of recalling (TRM)

	A	B	C	D
A		0.48	0.98	3.35*
B			0.64	3.17*
C				2.76*
D				

t values between Mean time per move in problem solving (TMM)

	A	B	C	D
A		2.34*	4.32*	9.57*
B			1.09	8.31*
C				7.74*
D				

t values between Mean time per move in spiral maze (PRT)

df = 30

	A	B	C	D
A		0.32	1.39	9.32*
B			1.50	9.36*
C				9.19*
D				

TABLE 16

t-values of Significant Measures (5%)

APPENDIX 2

Appendices of Chapter 6

APPENDIX 2.1

Procedure for calculating the Reliability time (TL)

TL may be estimated in the following way:

TL = time to make all moves (TM) + time to set up all subgoals (TS)

i.e. $TL = TM + TS$ — ①

$TM = (\text{number of moves}) \times (\text{time per move})$

Assuming N discs and 2 mistaken moves and corrective moves, the number of moves $= 2^N + 3$

$\therefore TM = (2^N + 3) \times (\text{time per move})$.

The time to move a disc depends on the mean S.R.T., the number of subject actions, and the time per action. If these are assumed to be 2.5 seconds, a minimum of 2 actions and a maximum B1 seconds per action, then the time per move $\doteq 2B1 + 2.5$.

Then $TM \doteq (2^N + 3) \times (2B1 + 2.5)$ seconds —— ②

Now $TS = (\text{number of sub-goals}) \times (\text{time per sub-goal})$

$= (N-2) \times (\text{time per sub-goal})$.

If B2 is an estimate of the maximum time for a move including setting up a sub-goal then TS is estimated arbitrarily as follows

$$TS = (N-2) \times \frac{(B1 + B2)}{2} \quad \text{——— ③}$$

If the keyboard task was used, the time of solution must include another term to allow for (i) printing "illegal' and (ii) drawing the current status. "Illegal" is printed, $(2^{N-2} + 2)$ times in the course of an optimum solution (refer to section 5.2.6) and each print is assumed to take 0.7 seconds.

Drawing the current status is assumed to occur every third move and involve a time of 3 seconds. Therefore the extra term for keyboard task is

$$(2^{N-2} + 2) \times 0.7 + 36 \text{ seconds.}$$

The estimated times of solution are, therefore,

(1) for Light pen (LP) = $(2^{N} + 3) \times (2B1 + 2.5) + (N-2) \times (B1+B2)$ seconds

(2) for Keyboard (Ky) = $LP + (2^{N-2} + 2) \times 0.7 + 36$

The values of B1 and B2 are further discussed in Section 6.1 (chapter 6)

APPENDIX 2.2

Techniques for Collecting Personality Data

The data described in this report was collected by using two mediums; paper questionnaires and by using a computer video display (VDU) with keyboard. The actual techniques used were the State Anxiety Questionnaire (STAI) developed by Spielberger and Associates and the Eysenck Personality Inventory (EPI).

The STAI consists of twenty questions about emotional disposition. Each question has a choice of one of four possible answers which relate (along a continuum) directly to emotional disposition. The questions are semantically balanced in an attempt to avoid response preferences unrelated to question content. The difference in the mediums (VDU and paper) is mainly that the VDU allows the presentation of one question at a time and disallows the scanning of previous or following questions compared with the paper questionnaire. The advantage of the VDU method is that data is collected and processed quickly, automatically and therefore reliably (with suitable error correction allowances for the respondent) whereas paper questionnaires allow of both respondent error and processing error and is time consuming. The disadvantage of VDU use is that programmes must be written and computer resources used; both of these are time consuming and expensive. Also, if there is a long Computer response time the respondent may be caused to change his emotional state thereby affecting the results.

In order to test whether differences would be reflected in the measure being taken, a simple experiment was carried out. Twenty four subjects were asked to fill in the STAI on both mediums within one hour. The computer version deliberately utilised a slow device (VDU operating at 30 characters/sec) so that it would present a long response time.

The order of presentation of each medium was alternated for the subjects in the sample and spaced by 1 hours. The results are as follows:

Mode	VDU		Paper (15/24 were > than for paper)		
Order	Mean	(S.D.)	Mean	(S.D.)	
VDU 1st	39.58	(6.67)	40.17	(5.78)	NSD
Paper 1st	40.58	(8.34)	40.92	(7.4)	NSD
Total	40.54 NSD	(6.51)	40.08 NSD	(7.24)	

STAI scores (assumed on equal interval base)

*NSD means no significant difference (@ 5%)

As this Table shows, there are no significant differences (at 5% level) between order or mode as shown in the STAI scores. Since the effect of computer systems utilsing fast VDU's and binary choice personality tests measures would be less than for the conditions used here, the EPI was presented one question at a time on a fast VDU rather than on paper.

In both cases, the subjects were self paced and left in isolation to complete the tests.

APPENDIX 2.3

Raw Data

TABLE:

NOTE: For meaning of R1, R2, P1-3 and t_c, refer to treatment of data section of Chapter 6.

TABLE A1

Lp				KYB			
OPT		SUB-OPT		OPT		SUB-OPT	
1st	2nd	1st		1st	2nd	1st	2nd
24		39*			23		31*
11		19*			24		37*
	17		67	46		47	
53		58*			28		42*
28		32			25		31*
	83		50	26		29*	
27		59			15		48*
	38		42	50		52*	
	15		30	28		62*	
	17		28	17		42*	
	15		28	67		16*	
13		58*			24		32
13		30			15		32*
	15		43	54		16*	
	30		30	18		43	
	35		24	22		32*	
	24		33	27		43*	
	93		38	28		37*	
19		56			19		33*
18		29			18		39
	14		30	21		86*	
13		31			16		45*
32		62*			27		29*
25		41			21		35

No. of logical moves to solve problem
(* means not solved - Nos. replaced by 100 in analysis)

R1
MEAN
PEAK 1

	LP				KYB			
	OPT		SUB		OPT		SUB	
Subject No.	ORDER 1st	ORDER 2nd	1st	2nd	1st	2nd	1st	2nd
1	4.25		4.52			4.31		1.82
2	4.38		5.26			6.17		3.76
3		5.09		6.28	4.58		1.98	
4	5.83		7.77			5.40		3.68
5	2.96		2.71			4.70		2.33
6		4.81		4.93	6.34		3.61	
7	2.91		2.83			3.68		2.14
8		7.71		5.94	6.12		4.32	
9		3.17		2.87	3.95		2.56	
10		3.91		3.50	4.11		1.91	
11		4.10		4.01	5.53		5.37	
12	8.04		6.06			5.26		2.19
13	3.50		2.98			3.50		0.70
14		8.97		6.18	4.72		3.98	
15		5.97		4.37	7.88		2.78	
16		2.87		2.87	5.13		4.66	
17		5.42		5.19	4.08		2.06	
18		5.36		3.66	6.94		2.96	
19	4.08		5.44			4.04		1.80
20	3.61		4.03			2.64		1.84
21		3.00		2.30	3.09		0.61	
22	3.58		3.78			5.22		2.22
23	7.28		5.55			4.08		2.76
24	4.58		3.79			3.00		3.38
Σx	55.0	60.38	54.72	52.1	62.47	51	36.78	28.63
Σx^2	281.83	342.08	274.72	246.74	346.75	230.46	133.24	76.92

=19

R1 P1 $\bar{t}$ (SECONDS)

2	LP				KYB			
	OPT		SUB		OPT		SUB	
ubject No.	1st	2nd	1st	2nd	1st	2nd	1st	2nd
1	3		10			1		3
2		4		8	3		5	
3	0		5			3		6
4		3		6	4		8	
5	1		2			1		0
6		8		2	6		5	
7	1		6			1		0
8		7		4	18		10	
9	0		1			6		4
10		1		3	1		6	
11	1		2			21		0
12		4		3	1		3	
13	2		4			3		6
14		1		8	12		2	
15	4		0			2		3
16		5		7	6		1	
17	1		0			7		10
18		3		0	5		5	
19	0		3			1		6
20		0		2	0		1	
21	1		2			1		2
22		2		6	1		5	
23	10		9			3		4
24		3		1	2		4	
Sx	24	41	44	50	59	50	55	44
Sx^2	134	203	280	292	597	562	331	262

R1 P2 N

TABLE A3

R1 MEAN PEAK 2 - R1 MEAN PEAK 1

	LP				KYB			
	OPT		SUB		OPT		SUB	
Subject No.	1st	2nd	1st	2nd	1st	2nd	1st	2nd
1	9.25		8.59			6.19		13.45
2		8.87		7.51	19.33		16.99	
3	0		13.06			13.59		19.1
4		16.34		19.13	10.85		15.72	
5	7.54		11.39			9.8		0
6		17.19		17.35	22.33		29.75	
7	16.59		10.97			19.82		0
8		12.22		18.24	17.66		18.18	
9	11.07 0		8.13			9.05		20.74
10		10.59		7.72	25.39		11.36	
11	10.4		6.68			16.59		14.8_07
12		9.96		11.61	12.24		19.08	
13	8.5		7.43			6.33		10.88
14		20.53		15.59	21.03		12.29	
15	12.53		9.93_0			18.19		18.42
16		7.03		8.67	19.87		15.5	
17	13.08		9.93_0			22.09		11.04
18		12.81		13.65_0	3.96		19.01	
19	11.07_0		8.66			20.06		13.65
20		6.89		10.2	13.65_0		13.5	
21	6.5		8.6			3.41		11.59
22		6.92		9.62	11.28		20.15	
23	15.32		15.76			10.75		13.24
24		15.59		24.61	17.0		9.12	
Σx	99.71	144.94	99.27	150.25	180.94	155.87	187.87	132.11
Σx^2	1203.09	1975.25	1057.68	2374.58	3365.34	2472.63	3668.57	2051.36

R1 P2 $\bar{t}_c$ (SECONDS)

TABLE A4

K 3

	LP				KYB			
	OPT		SUB		OPT		SUB	
Subject No.	1st	2nd	1st	2nd	1st	2nd	1st	2nd
1	0		3			1		0
2	0		7			0		4
3		0		4	5		5	
4	0		1			1		4
5	0		0			0		0
6		0		3	0		7	
7	1		0			0		2
8		0		0	0		12	
9		0		0	0		5	
10		0		2	0		5	
11		1		2	0		0	
12	0		0			0		1
13	0		1			1		2
14		0		1	0		0	
15		0		0	0		1	
16		1		0	0		3	
17		0		0	1		2	
18		2		1	2		5	
19	0		1			1		6
20	0		0			0		1
21		0		0	0		0	
22	1		1			0		8
23	0		2			1		3
24	0		2			2		2
Σx	2	4	18	13	8	7	45	33
Σx^2	2	6	70	35	30	9	307	155

R1 P3 N

TABLE A5

R1 P3 - R1 P1

	LP				KYB			
	OPT		SUB		OPT		SUB	
Subject No.	1st	2nd	1st	2nd	1st	2nd	1st	2nd
1	0		49.63			21.19		0
2		0		92.94	0		45.54	
3	0		44.29			24.12		35.04
4		0		74.93	24.8		38.12	
5	0		0			0		0
6		0		151.07	0		99.69	
7	25.4		0			0		44.66
8		0		0	0		51.75	
9	0		0			0		59.44
10		0		25.25	0		43.69	
11	25.4		71.26			0		0
12		0		0	0		30.11	
13	0		22.4			17.0		16.53
14		0		26.12	0		0	
15	0		0			0		29.72
16		26.63		0	0		388.34	
17	0		0			14.42		23.34
18		25.14		45.54	22.56		47.6	
19	0		27.8			26.46		48.2
20		0		0	0		42.62	
21	0		0			0		0
22		0		17.82	0		50.20	
23	0		58.55			25.42		59.05
24		0		45.51	26.5		45.92	
Ex	50.8	51.77	273.87	479.18	47.36	128.61	883.58	315.98
Ex^2	1290.32	1341.17	14205.43	42856.91	1123.99	2585.03	178477.4	14266.85

R1 P3 $\bar{t}_c$ (SECONDS)

TABLE A6

TABLE A7

	LP				KYB			
	OPT		SUB		OPT		SUB	
bject No.	1st	2nd	1st	2nd	1st	2nd	1st	2nd
1	26		47			37		60
2	15		29			46		68
3		17		73	83		89	
4	46		60			46		74
5	29		34			48		62
6		88		51	24		51	
7	27		60			30		96
8		44		46	91		102	
9		16		31	50		117	
10		18		30	34		77	
11		17		29	126		58	
12	16		58			36		63
13	14		32			28		53
14		16		50	93		27	
15		29		31	30		86	
16		34		28	42		58	
17		24		32	46		81	
18		96		39	42		74	
19	19		59			37		45
20	17		30			30		65
21		14		32	30		49	
22	16		35			30		68
23	39		64			37		58
24	28		44			36		70
x	292	413	552	472	691	441	869	782
x^2	8250	23079	27432	20562	51371	16719	69835	52676

R2 P1 N

2
EAN
EAK 1

TABLE A8

	LP				KYB			
	OPT		SUB		OPT		SUB	
Subject No.	1st	2nd	1st	2nd	1st	2nd	1st	2nd
1	2.19		3.16			2.28		2.49
2	2.43		4.42			2.63		3.04
3		2.62		3.53	2.79		2.47	
4	3.91		2.91			3.78		4.00
5	2.78		2.86			2.24		2.42
6		2.11		2.64	4.88		3.71	
7	1.83		3.91			3.3		2.80
8		4.27		3.91	2.37		2.77	
9		2.69		3.43	3.92		4.06	
10		2.22		2.22	1.97		2.65	
11		2.26		2.68	3.79		4.46	
12	5.00		3.46			3.22		2.59
13	2.00		3.09			2.61		2.82
14		5.94		5.88	4.91		4.42	
15		4.71		4.21	4.87		3.52	
16		1.24		2.21	2.67		4.28	
17		3.96		4.21	2.48		2.26	
18		2.94		4.15	2.69		2.28	
19	1.66		3.05			2.07		2.92
20	2.74		3.59			1.70		2.52
21		1.29		2.07	1.80		1.59	
22	2.13		2.21			2.20		2.44
23	3.83		4.07			2.53		3.28
24	2.36		2.87			2.39		1.62
Σx	32.86	36.25	39.6	41.14	39.32	30.95	38.47	32.94
Σx^2	101.1	131.82	134.85	155.01	143.8	83.62	134.01	94.01

= 978.22

R2 P1 t (SECONDS)

TABLE A9

2

	LP				KYB			
	OPT		SUB		OPT		SUB	
Subject No.	1st	2nd	1st	2nd	1st	2nd	1st	2nd
1	0		1			1		2
2		0		2	0		1	
3	0		1			1		1
4		1		2	2		2	
5	0		0			0		0
6		0		1	0		1	
7	1		5			0		0
8		0		0	1		0	
9	0		0			3		2
10		0		3	0		2	
11	0		3			2		0
12		0		1	0		0	
13	1		3			2		3
14		0		2	7		1	
15	2		0			0		0
16		5		3	0		1	
17	0		1			0		0
18		0		0	0		0	
19	0		0			1		2
20		1		1	0		3	
21	1		0			0		1
22		0		2	0		0	
23	1		3			2		0
24		0		0	1		0	
Σx	6	7	17	17	11	12	11	11
Σx^2	8	27	55	37	55	24	21	23

=250

R2 P2 N

TABLE A10

R2 P2 MEAN - R2 P1 MEAN

	LP				KYB			
	OPT		SUB		OPT		SUB	
Subject No.	1st	2nd	1st	2nd	1st	2nd	1st	2nd
1	0		13.06			9.22		12.61
2	0		8.25			0		15.00
3		0		18.27	16.71		13.39	
4	17.59		27.89			12.47		11.6
5	0		0			0		0
6		0		33.76	0		26.89	
7	16.67		10.89			0		0
8		0		0	27.31		0	
9		0		0	8.91		21.34	
10		0		10.45	0		12.25	
11		0		12.33	25.71		0	
12	0		14.44			0		0
13	5.5		6.03			4.89		5.81
14		0		23.02	8.91		16.04	
15		18.29		0	0		0	
16		1.24		6.90	0		20.12	
17		0		21.59	21.16		0	
18		0		0	0		0	
19	0		0			2.03		11.27
20	7.76		6.59			0		9.25
21		1.11		0	0		8.81	
22	0		7.19			0		0
23	2.83		17.72			15.47		0
24	0		0			9.11		0
Σx	50.35	20.64	102.33	126.32	108.71	53.19	118.84	65.54
Σx^2	685.77	337.29	1789.07	5590.55	7417.0	8007.86	10255.39	11020.86 =

45103.79

R2 P2 $\bar{t}_c$ (SECONDS)

TABLE A11

3

	LP				KYB				
	OPT		SUB		OPT		SUB		
Subject No.	1st	2nd	1st	2nd	1st	2nd	1st	2nd	
1	1		3			0		0	
2	0		1			0		1	
3		0		0	0		0		
4	0		0			0		2	
5	0		0			0		0	
6		0		1	0		0		
7	1		0			0		0	
8		0		0	0		0		
9		0		0	0		1		
10		0		0	0		2		
11		0		0	0		0		
12	1		1			0		1	
13	0		0			0		2	
14		0		0	0		2		
15		0		0	0		0		
16		0		0	0		1		
17		0		0	0		1		
18		1		0	0		0		
19	0		0			0		7	
20	0		0			0		0	
21		0		0	0		0		
22	0		1			0		0	
23	0		0			1		0	
24	0		0			1		0	
Σx	3	1	6	1	0	2	7	13	
Σx^2	3	1	12	1	0	2	11	59	=89

R2 P3 N

TABLE A12

R2 P3 MEAN - R2 P1 MEAN

	LP				KYB			
	OPT		SUB		OPT		SUB	
Subject No.	1st	2nd	1st	2nd	1st	2nd	1st	2n
1	20.31		29.94			0		
2	0		36.98			0		44.8
3		0		0	0		0	
4	0		0			0		43.4
5	0		0			0		
6		0		33.76	0		0	
7	27.67		0			0		0
8		0		0	0		0	
9		0		0	0		37.84	
10		0		0	0		41.05	
11		0		0	0		0	
12	21.5		33.24			0		23.7
13	0		0			0		20.1
14		0		0	0		48.13	
15		0		0	0		0	
16		0		0	0		58.5	
17		0		0	0		22.34	
18		26.56		0	0		0	
19	0		0			0		38.2
20	0		0			0		0
21		0		0	0		0	
22	0		37.39			0		0
23	0		0			26.97		0
24	0		0			27.11		0
Σx	69.48	26.56	137.55	33.76	0	54.08	207.86	170.37
Σx^2	1640.38	705.43	4766.83	1139.74	0	1462.33	9354.79	6326.81

25396.

R2 P3 $\overline{t}_c$ (SECONDS).

TABLE A13

OVERALL TIME TO SOLVE (MINUTES)

OPT

	N1				N2				
	E1		E2		E1		E2		
	01	02	01	02	01	02	01	02	
	5.15	4.72	2.38	2.48	2.48	5.38	2.64	8.72	
	11.26	1.82	4.74	11.37	12.36	16.50	2.27	18.65	
	4.66	3.00	3.55	1.70	4.31	2.58	2.11	4.67	Ex^2
Ex	21.07	9.54	10.67	17.25	19.15	24.46	7.02	32.04	1336.2

SUB.

	N1				N2				
	E1		E2		E1		E2		
	01	02	01	02	01	02	01	02	
	22.5*	20.98	11.59	9.79	8.78	8.58	27.8*	9.94	
	29.21*	8.14	13.45	15.57	24.5*	23.7	18.5	11.71	
	18.12	11.65	8.45	6.99	27.18*	29.79*	9.56	10.34	Ex^2 381
Ex	69.83	40.77	33.49	32.35	55.46	62.07	55.86	31.99	7345.0

B OPT

	N1				N2				
	E1		E2		E1		E2		
	01	02	01	02	01	02	01	02	
	5.95	31.56	4.14	4.73	2.75	13.07	8.16	7.18	
	6.65	15.56	7.29	24.50	10.75	13.41	4.94	8.76	
	4.52	33.90	5.29	2.40	4.85	15.16	3.53	6.76	Ex^2
Ex	17.12	81.02	16.72	31.63	18.35	41.64	16.63	22.70	4218.6

B SUB.

	N1				N2				
	E1		E2		E1		E2		
	01	02	01	02	01	02	01	02	
	14.49	10.50*	23.04*	18.40*	12.70	39.07*	22.42*	22.39	
	17.61*	38.15*	10.78	36.27*	27.23*	29.70*	24.34*	20.21*	507
	24.92*	16.80*	10.34*	20.10*	11.67	19.82*	12.97*	20.29*	Ex^1
Ex	57.02	65.45	44.16	74.77	57.60	88.59	59.73	62.89	12226.1

ns not solved

TABLE A14

	Inexp'd	Exp'd
Solved	7	11
Not Solved	5	1

	Order 1st	Order 2nd
Solved	7	11
Not Solved	5	1

$p = 0.071$ (Fisher Exact Probability Int). $p = 0.071$

Significant Frequency tables of Solution using Light-Pen Sub-optimum System

		Light Pen	
		Not Solved	Solved
Keyboard	Solved	3	3
	Not Solved	3	15

McNemar Test.

$X^2 = 5.8$, $d_f = 1$

$p = \frac{1}{2} (0.02) < 0.01$ (1 tailed test)

Between Input tests for the sub-optimum systems

TABLE A15

		LP Sub Opt.		KYB Sub Opt.	
		Solved	Not Solved	Solved	Not Solved
A	N1	10	2	3	9
	N2	8	4	3	9
B	E1	7	5	3	9
	E2	11	1	3	9
C	O1	7	5	5	7
	O2	11	1	1	11
Fisher Exact Probability (1 tail)	PR(A)	0.2427		0.3596	
	PR(B)	0.0706*		0.3596	
	PR(C)	0.0706*		0.0706*	

Frequency of Solution in different groups

TABLE A16

	LP				KYB			
	OPT		SUB		OPT		SUB	
Subject No.	1st	2nd	1st	2nd	1st	2nd	1st	2nd
1	1		2			0		0
2	0		0			0		1
3		1		0	1		0	
4	2		0			0		0
5	0		0			1		0
6		0		0	0		0	
7	0		0			0		0
8		0		0	0		1	
9		0		0	0		0	
10		0		0	0		1	
11		0		0	1		0	
12	0		0			0		0
13	0		0			0		0
14		0		0	0		0	
15		1		0	0		0	
16		1		0	0		0	
17		1		0	0		1	
18		2		0	1		0	
19	0		1			0		0
20	0		0			0		1
21		0		0	0		0	
22	0		1			0		0
23	0		1			1		0
24	0		1			0		0
	3	6	4	0	3	2	3	2

$Ex^2 =$

Frequency of logical errors

TABLE A17

KEYBOARD RESULTS

FREQUENCY OF INPUTS

	1 (OPT)					2 (SUB-OPTIMUM)				
Subject No.	M	D	R	S	TE	M	D	R	S	TE
1	23	10	3	1	2	31	14	0	0	0
2	24	15	2	0	2	37	19	1	0	2
3	46	23	3	1	1	47	12	0	0	0
4	28	18	1	1	6	42	31	0	1	3
5	25	8	3	0	3	31	8	0	0	0
6	26	21	1	2	8	29.	26	3	0	2
7	15	14	0	0	0	48	48	0	0	0
8	50	23	1	2	5	52	24	1	2	1
9	28	27	1	0	0	62	61	0	0	0
10	17	7	0	0	1	42	12	0	0	1
11	67	63	10	0	14	16	15	0	0	0
12	24	19	2	1	2	32	30	0	0	0
13	15	0	0	0	0	32	4	0	0	4
14	54	50	3	1	4	16	15	0	0	0
15	18	9	1	1	2	43	15	0	0	0
16	22	19	1	0	2	32	31	1	0	3
17	27	19	2	1	0	43	26	1	0	3
18	28	13	2	1	3	37	29	0	0	0
19	19	18	0	0	0	33	23	0	0	1
20	18	17	1	0	2	39	14	2	0	0
21	21	5	1	1	1	86	16	1	0	2
22	16	11	1	0	3	45	25	0	0	2
23	27	11	4	1	5	29	12	0	0	0
24	21	6	1	1	2	35	5	0	0	0

KEY : M – MOVES
D – DRAWS

TABLE A18

Acceptability Ranks:

Sign Test Results

a. Within Inputs:

L.P.

Opt vs sub.opt.
x = 3, n = 24

Order 1st vs
Order 2nd
x = 11 n = 24

Keyboard

Opt. vs sub opt.
x = 1, n = 23

Order 1st is
Order 2nd
x = 10 N = 23

b. Between Inputs:

Lp vs KYB
(Optimum)
x = 5N = 24

LP vs KYB (sub-optimum)
x = 11N = 23(NSD)

LP OPT vs KYB sub
opt; x = 4 N = 24

LP Sub vs KYB opt
x = 10, N = 24

Key x = numbers of fewer signs
N = no. of matched points who showed differences.

Subject	LP		KYB	
No.	OPT	SUB-OPT	OPT	SUB-OPT
1	1	4*	3	2
2	1	4*	$2\frac{1}{2}$	$2\frac{1}{2}$*
3	1	2*	3	4
4	3	4*	1	2*
5	1	4	2	3
6	1	$3\frac{1}{2}$	2	$3\frac{1}{2}$*
7	2	4	1	3*
8	1	2	3	4*
9	1	2	3	4*
10	4	3	1	2
11	1	4	2	3
12	1	4*	2	3*
13	1	4	2	3*
14	1	2	3	4*
15	2	1	3	4
16	1	2	3	4*
17	3	4	1	2*
18	2	1	3	4*
19	1	2	3	4*
20	3	4	1	2
21	1	3	2	4
22	1	2	3	4*
23	1	4*	2	3*
24	1	2	3	4
h_j	36	$71\frac{1}{2}$	$54\frac{1}{2}$	78

* Failed to solve

$X_p = 26.56$ $d_f = 3$ (Significant at .1%)

TABLE A19

Subject No.	Acceptability Score	
	LP	KYB
1	5	4
2	5	4
3	7	6
4	7	8
5	7	4
6	9	6
7	4	3
8	7	3
9	3	8
10	6	6
11	5	7
12	6	3
13	6	2
14	7	5
15	7	6
16	7	7
17	6	5
18	5	6
19	6	8
20	7	7
21	6	7
22	7	10
23	5	3
24	5	6

$t = 0.963$

TABLE A20

LIGHT PEN

		Acceptability	Work Production							
		θ	R1 P1 N	R1 P1 MEAN	R1 P2 N	R1 P2 MEAN t	OUTM	R2 P1 N	R2 P1 MEAN	R2 P2 N
θ	RS	–	0.226	0.171	0.123	0.039	0.216	0.197	0.153	0.135
	T		1.090	0.812	0.580	0.182	1.039	0.942	0.728	0.639
R1 P1 N	RS		–	0.197	0.411*	0.451*	0.770	0.217	0.144	0.978
	T			0.944	2.116	2.371	5.658	1.044	0.681	21.84*
R1 P1 MEAN	RS			–	0.404*	0.658*	0.612*	0.256	0.786*	-0.246
	T				2.075	4.098	3.628	1.240	5.965	-1.188
R1 P2 N	RS				–	0.550*	0.723*	0.487*	0.160	0.188
	T					3.089	4.905	2.616	0.761	0.899
R1 P2 MEAN t	RS					–	0.756*	0.493*	0.501*	0.031
	T						5.416	2.658	2.713	0.146
OVTM	RS						–	0.816*	0.471*	0.078
	T							6.653	2.507	0.370
R2 P1	N							–	0.178	0.113
R2 P1	MEAN								–	-0.181
R2 P2	N									–
R2 P2	MEAN									

KEYBOARD

		θ	R1 P1 N	R1 P1 t MEAN	R1 P2 N	R1 P2 t MEAN	OUTM	R2 P1 N	R2 P1 t Mean	R2 P2 N
θ	RS	–	0.026	-0.052	-0.023	-0.081	0.127	0.026	-0.020	.006
	T		0.123	-0.244	-0.108	-0.382	0.600			
R1 P1 N	RS		–	0.248	0.113	0.593*	0.327	1**	0.321	0.473*
	T			1.199	0.533	3.452	1.625			
R1 P1 MEAN t	RS			–	0.453*	0.377	0.304	0.248	0.561*	-0.178
	T				2.382	1.912	1.497			
R1 P2 N	RS				–	0.359	0.429*	0.593*	0.598*	0.444*
	T					1.805	2.227			
R1 P2 MEAN t	RS					–	0.296	0.113	0.388	-.188
	T						1.452			
OVTM	RS						–	0.327	0.682*	0.214
	T									
R2 P1	N								0.321	0.473*
R2 P1	MEAN								–	0.261
R2 P2	N									–
R2 P2	MEAN t									

Key

* signifies that the correlation coefficient (RS) has a t-value (T) that it is significantly different from zero at 5% level.

RS calculated by ranking measures over subjects for each measure then applying Spearman rank correlation procedure (Siegel (1956), p. 20

Intercorrelations between measures within input devices for OPTIMUM SYSTEM

TABLE A21

Inter-Correlations between Measures within Input Device for Sub-optimum system

Sub-optimum system

		Acceptability		Work Production							
		θ	OVTM	R1 P1 N	R1 P1 MEAN	R1 P2 N	R1 P2 MEAN t_c	R1 P1 N	R2 P1 MEAN	R2 P2 N	R2 P2 MEAN t_c
θ	RS T	-	0.141 0.67	0.062 0.29	0.228 1.098	0.026 0.123	0.33 1.64	0.064	-0.187	-0.097	0.165
OVTM	RS T		-	0.069 0.326	-0.194 -0.93	-0.305 -1.501	0.055 0.257	0.032	0.016	-0.090	-0.142
R1 P1 N	RS T			-	0.522* 2.868	0.223 1.074	0.62* 3.703	0.984*	0.237	-0.075	0.397
R1 P1 MEAN	RS T				-	0.253 1.229	0.477* 2.543	0.503*	0.401	-0.161	0.511*
R1 P2 N	RS T					-	0.378 1.913	0.312	0.041	0.605*	0.427*
R1 P2 MEAN t_c	RS T						-	0.625*	-0.023	-0.026	0.317
R2 P1 N	T RS							-	0.179	-0.046	0.390
R2 P1 MEAN	T RS								-	0.064	0.171
R2 P2 N	T RS									-	0.581*
R2 P2 MEAN t_c	T RS										-

θ		-	-0.127 -0.602	0.155 0.738	0.047 0.219	0.003 0.017	0.206 0.986	-0.031	0.116	0.269	0.312
OVTM	RS T		-	-0.044 -0.209	-0.233 -1.122	-0.255 -1.238	0.002 0.010	0.184	-0.315	-0.073	-0.167
R1 P1 N	RS T			-	0.358 1.798	0.015 0.072	0.286 1.402	0.688*	0.239	-0.267	-0.107
R1 P1 MEAN	RS T				-	-0.105 -0.498	0.344 1.717	0.077	0.550*	-0.401	0.061
R1 P2 N	RS T					-	0.370 1.871	0.294	-0.140	0.183	0.097
R1 P2 MEAN t_c	RS T						-	0.257	0.221	-0.028	0.309
R2 P1 N								-	-0.167	-0.240	-0.219
R2 P1 MEAN									-	0.232	0.454*
R2 P2 N										-	0.743*
R2 P2 MEAN t_c											-

Key : As For TABLE A20

TABLE A22

Rank Correlations Between Inputs and Systems for 4 Measures

LPOPT

		R1 P1 N	R1 P1 MEAN	R2 P1 N	R2 P1 MEAN
LPSUB	R1 P1 N	.334	.477*	.344	.258
	R1 P1 MEAN	.125	.784*	.201	.493*
	R2 P1 N	.285	.428*	.297	.176
	R2 P1 MEAN	.059	.584*	.046	.664*

LPOPT

		R1 P1 N	R1 P1 MEAN	R2 P1 N	R2 P1 MEAN
KYB OPT	R1 P1 N	.032	.285	.060	.436*
	R1 P1 MEAN	.278	.446*	.367	.370
	R2 P1 N	.032	.285	.060	.436*
	R2 P1 MEAN	.231	.416*	.246	.378

LPSUB

		R1 P1 N	R1 P1 MEAN	R2 P1 N	R2 P1 MEAN
KYB SUB	R1 P1 N	-.008	.150	-.098	.188
	R1 P1 MEAN	-.080	.176	-.139	.144
	R2 P1 N	.053	.086	-.010	.310
	R2 P1 MEAN	-.089	.160	-.101	.095

KYBOPT

		R1 P1 N	R1 P1 MEAN	R2 P1 N	R2 P1 MEAN
KYB SUB	R1 P1 N	.328	.387	.328	.482*
	R1 P1 MEAN	.504*	.657*	.504*	.636*
	R2 P1 N	.180	.139	.180	.158
	R2 P1 MEAN	.249	.327	.249	.610*

KEY : * AS FOR TABLE A20

Intercorrelations Between Subsidiary Variables and Measures

				Light Pen								KeyBoard							
				Optimum System				Sub-Optimum System				Optimum System				Sub-Optimum System			
				A	E	O	TA	A	E	O	TA	A	E	O	TA	A	E	O	TA
WP	R1	P1	n t	.247 -.201	.178 -.331	.272 .146	.403 .206	-.124 -.045	-.044 -.093	.32 .111	.08 .177	-.110 .308	-.127 -.185	.294 .018	.110 .521*	-.049 .251	-.036 .030	.416* .358	.278 .543
		P2	n t_c	.138 -.246	-.095 -.034	-.179 .235	.566* .223	.143 .308	-.316 -.197	.009 .316	-.047 .064	-.016 .107	.151 -.197	.393 .027	.413* .036	-.010 .230	-.020 -.295	-.208 .156	-.112 .374
	R2	P1	n t	.177 -.417	.487 -.060	.212 .198	.458* .073	.015 -.037	-.150 -.193	.239 .247	.053 -.223	-.110 .047	-.127 .065	.294 .540*	.110 .464*	.176 .107	-.138 .150	.175 .311	-.225 .478*
		P2	n t_c	.251 .165	.077 .434*	.080 .121	.088 .092	.219 .096	.181 .138	.206 .396	-.057 .036	-.255 -.171	-.014 -.135	.222 .262	.195 .190	.274 .363	.071 .026	-.004 .236	.478* -.020
Overall time				.109	-.008	.315	.530*	-.178	-.059	.189	.001	.153	.053	.40	.210	-.247	-.266	-.016	-.136
ERRORS	Failure to Solve +			-	-	-	-	0.24	0.34	0.36	-	-	-	-	-	0.15	0.23	0.14	0.04*
	Logical +			0.14	0.14	0.34	-	0.14	0.14	0.21	-	0.34	0.34	0.29	0.34	0.34	0.14	0.37	0.14
	Typing +			-	-	-	-	-	-	-	-	-.102	.032	.190	.586*	.303	.184	-.308	.018
θ				.089	.322	.074	.047	.089	.322	.074	.047	.156	-.039	.112	.175	.156	-.039	.112	.175

KEY
A - Anxiety
E - Extraversion
O - Occupation
TA - Typing Ability
+ - for these measures, entry of - means not tested.

* Significant at 5% level (RS different from 0).
-. means -0. eg -.016 is -0.016.

TABLE A24

Intercorrelations between Main and Subsidiary Independent Variables

(Rank Correlation Coefficients Across all Subjects)

		EP	ET	N	A	O	T
Ep	Experience		+0.28	+0.32	+0.16	+0.84*	+0.32
Er	Extraversion			-0.17	+0.19	-0.30	-0.17
N	Neuroticism				-0.67	-0.12	-0.07
A	Anxiety					-0.06	-0.04
O	Occupation						-0.63
T	Typing Ability						

Measure	Source	d_f	SSQ	MSS	F
Acceptability Score for Light Pen System	N	1	3.37	←	4.74*
	E	1	0.04	←	-
	O	1	1.04	←	1.46
	NxE	1	9.38	←	13.21*
	ExO	1	0.38	←	-
	NxO	1	1.04	←	1.46
	ExNxO	1	0.37	←	-
	ERROR	16	11.33	0.71	
	TOTAL	23	26.95		

Measure	Source	d_f	SSQ	MSS	F
Light Pen Optimum Number of Logical Errors (for n)+	N	1	4.16	←	208*
	E	1	0.00	←	- -
	O	1	0.67	←	33.5*
	NxE	1	0.00	←	-
	ExO	1	1.50	←	75*
	NxO	1	0.67	←	33.5*
	ExNxO	1	1.50	←	75*
	ERROR	16	0.33	0.02	
	TOTAL	23	8.83		

TABLE B25

2x2x2 ANOVA TABLES FOR MEASURES SHOWING SIGNIFICANT EFFECTS

Note: Choice of error terms for calculating F was based on a fixed effects model, i.e. the single error term was used. This limits the generality of the results to the population used in the experiment.

* 5% level of significance

\+ a number of the results are associated with extreme values. In these cases, the ANOVA technique may lead to too many significant results.

Measure	Source	d_f	SSQ	MSS	F	Measure	Source	d_f	SSQ	MSS	F
Light Pen Optimum R2 P2 $\bar{t}_c$	N	1	34.8	←	-	Light Pen Sub-Optimum R2 P3 $\bar{t}_c$ (small n)	N	1	0.67	←	-
	E	1	66.33	←	1.09		E	1	100.04	←	1.11
	O	1	45.65	←	-		O	1	100.04	←	1.11
	NxE	1	9.50	←	-		NxE	1	165.37	←	1.83
	EXO	1	274.1	←	4.54*		ExO	1	486.00	←	5.38*
	NxO	1	19.94	←	-		NxO	1	165.37	←	1.83
	NxExO	1	2.60		NS		NxExO	1	0.667	←	NS
	ERROR	16	965.2	60.325			ERROR	16	1444.8	90.30	
	TOTAL	23	1418.1				TOTAL	23	2463.0		
Light Pen Optimum P2 P3 n (small values)	N	1	0.00	←	-	Keyboard Sub-Optimum P2 P1 $\bar{t}$	N	1	0.06	←	-
	E	1	0.17	←	1.36		E	1	4.82	←	13.12*
	O	1	0.17	←	1.36		O	1	1.27	←	3.46
	NxE	1	0.17	←	1.36		NxE	1	1.26	←	3.47
	ExO	1	0.67	←	5.36*		ExO	1	1.46	←	3.97
	NxO	1	0.17	←	1.36		NxO	1	0.69	←	1.88
	NxExO	1	0.00	←	NS		NxExO	1	0.10	←	NS
	ERROR	16	2.00	0.125			ERROR	16	5.88	0.3675	
	TOTAL	23	3.33				TOTAL	23	15.53		

TABLE A25 (contd.)

Measure	Source	d_f	SSQ	MSS	F	Measure	Source	d_f	SSQ	MSS	F
Light Pen Sub-Optimum R2 P1 $\bar{t}$	N	1	0.11	←	-	Light Pen Sub-Optimum Logical errors (small n)	N	1	0.00	←	
	E	1	0.20	←	-		E	1	0.00	←	
	O	1	0.09	←	-		O	1	1.50	←	5.17*
	NxE	1	6.39	←	9.83*		NxE	1	0.17	←	-
	ExO	1	0.37	←	-		ExO	1	0.00	←	
	NxO	1	0.002	←	-		NxO	1	0.00	←	
	NxExO	1	0.53	←	NS		NxExO	1	0.17	←	
	ERROR	16	10.44	0.65			ERROR	16	4.67	0.29	
	TOTAL	23	18.15				TOTAL	23	6.50		
Light Pen Sub-Optimum R2 P2 $\bar{t}_c$	N	1	108.42	←	-	Light Pen Sub-Optimum Overall time to solve	N	1	34.89	←	-
	E	1	970.6	←	7.82*		E	1	230.89	←	4.91*
	O	1	1.59	←	-		O	1	93.83	←	1.99
	NxE	1	0.00	←	-		NxE	1	9.48	←	-
	ExO	1	2.92	←	-		ExO	1	0.27	←	-
	NxO	1	18.04	←	-		NxO	1	6.98	←	-
	NxExO	1	15.57	←	NS		NxExO	1	142.11	←	NS
	ERROR	16	1736.2	124.06			ERROR	16	752.17	47.01	
	TOTAL	23	2853.4				TOTAL	23	1270.6		

TABLE A25 (contd.)

Measure			BETWEEN I (LP vs KYB)		WITHIN I (BETWEEN S)		TABLE 8
			OPT	SUB	LP	KYB	
R1	P1	$\bar{t}$	0.10	4.65*	0.75	5.64*	I,S
	P2	n	1.60	-0.24	1.78	0.34	NSD
		$\bar{t}_c$	2.14*	1.44	0.18	0.25	I
	P3	n	1.49	2.82*	2.85*	4.08*	I,S
		$\bar{t}_c$	1.07	1.09	3.45*	2.68*	S
R2	P1	$\bar{t}_c$	0.12	0.15	1.68	0.14	NSD
	P2	n+	1.04	1.45	2.47*	0.02	S
		$\bar{t}_c$+	0.73	0.27	2.03*	0.12	S
	P3	n	0.87	1.76	0.79	2.42*	S
		$\bar{t}_c$	0.73	1.71	0.93	3.07*	S
Overall time to solve			2.14*	2.45*	5.6*	4.7*	I,S
Logical Errors+			1.08	0.05	1.21	0	NSD
Typing Errors+			-	-	-	2.55	S

* significant difference at 5% - 2 tailed test, df = 23

\+ non-normal distribution

TABLE A26

paired t-values for Tables A1-A12

Raw Data of Chapter 7

TABLE:	
1	R1 N P3
2	Overall Time to Solve
3	R1 P1 $\bar{t}$
4	R1 P2 $\bar{t}_c$
5	R1 P3 $\bar{t}_c$
6	R1 P2 N
7	R2 P1 $\bar{t}$
8	R2 P2 N
9	R2 P2 $\bar{t}_c$
10	R2 P3 N
11	P2 P3 $\bar{t}_c$
12	Number of Moves
13	Individual Differences for LP, D, L
14	Individual Differences for I and D
15	2x2 Mixed Model ANOVA Results
16	Paired t Values

NOTE: For meaning of R, P and t_c, refer to section on treatment of data in Chapter 6.

: 1

1

3

LP				KYB			
UIC1		UIC2		UIC1		UIC2	
01	02	01	02	01	02	01	02
-	0	0	-	-	1	1	-
0	-	-	0	0	-	-	0
-	0	0	-	-	1	0	-
0	-	-	1	0	-	-	0
-	2	1	-	-	0	2	-
0	-	-	2	1	-	-	1
-	0	13	-	-	2	0	-
1	-	-	3	2	-	-	0
-	1	5	-	-	0	0	-
0	-	-	1	1	-	-	1
-	1	1	-	-	0	0	-
0	-	-	1	0	-	-	1
-	0	0	-	-	0	0	-
0	-	-	3	1	-	-	0
-	0	0	-	-	0	0	-
0	-	-	0	0	-	-	0

KEY:
LP - LIGHT PEN
KYB - KEYBOARD
UIC 1 - User Input Characteristics 1, (automatic feedback)
UIC 2 - " " " 2, (no " ")
O 1 - ORDER OF PRESENTATION - 1 = 1ST
O2 - " " " - 2 = 2ND

THIS AND UP TO TABLE 12,

EASE NOTE: 1. UIC IS ALSO CALLED "DIALOGUE" IN THE TEXT.
2. N = NUMBER OF OR FREQUENCY
3. P1, P2, P3 = PEAKS 1, 2, 3 OF DISTRIBUTION OF INPUT TIMES
4. R1, R2 - SEE TEXT OF CHAPTER 6 FOR DEFINITION
5. $\bar{t}_c$ = MEAN CORRECTED TIME = $R_i P_j \bar{t} - R_i P_2 \bar{t}$ WHERE j = 2, 3 AND i =

E 2

TM
O
LVE
UTES)

LP				KYB			
UIC1		UIC2		UIC1		UIC2	
01	02	01	02	01	02	01	02
–	4.67	12.66	–	–	3.089	6.74	–
2.48	–	–	3.46	4.022	–	–	2.75
–	1.82	4.228	–	–	2.44	15.56	–
3.55	–	–	2.14	1.71	–	–	5.92
–	18.65	3.35	–	–	1.49	8.76	–
2.27	–	–	6.47	5.06	–	–	4.94
–	11.37	21.80	–	–	3.98	24.5	–
2.38	–	–	4.94	3.09	–	–	4.14
–	5.38	16.89	–	–	2.65	13.07	–
2.11	–	–	3.43	1.79	–	–	3.53
–	3.00	11.13	–	–	2.35	33.9	–
11.26	–	–	8.50	2.38	–	–	6.65
–	8.72	5.86	–	–	2.12	7.18	–
2.64	–	–	10.82	4.78	–	–	8.16
–	2.48	5.45	–	–	1.22	4.73	–
5.15	–	–	3.37	–	5.95	2.23	–

R1 P1 $\overline{t}$

TABLE 3

R1 P1 $\overline{t}$
(SECONDS)

LP				KYB			
UIC1		UIC2		UIC1		UIC2	
01	02	01	02	01	02	01	02
-	5.42	3.49	-	-	4.52	4.08	-
3.61	-	-	3.125	4.114	-	-	2.64
-	3.17	3.93	-	-	4.42	3.95	-
2.96	-	-	2.687	3.553	-	-	4.70
-	5.36	4.225	-	-	3.742	6.94	-
4.08	-	-	3.137	3.500	-	-	4.04
-	7.71	4.08	-	-	5.70	6.12	-
3.58	-	-	2.96	4.27	-	-	5.22
-	2.87	2.72	-	-	2.83	5.13	-
3.50	-	-	3.208	2.913	-	-	3.50
-	4.10	4.48	-	-	5.32	5.50	-
7.28	-	-	4.05	3.92	-	-	4.08
-	5.97	4.78	-	-	4.52	7.88	-
4.38	-	-	4.56	4.22	-	-	6.17
-	3.91	4.53	-	-	2.50	4.11	-
8.04	-	-	3.74	3.79	-	-	5126

LP				KYB			
UIC1		UIC2		UIC1		UIC2	
01	02	01	02	01	02	01	02
-	13.04	6.20	-	-	5.68	22.09	-
6.89	-	-	3.813	8.76	-	-	13.65
-	11.07	9.57	-	-	3.18	9.05	-
7.54	-	-	0	11.05	-	-	9.08
-	12.81	0	-	-	0	3.96	-
11.07	-	-	8.37	8.84	-	-	13.65
-	12.22	8.04	-	-	4.50	17.66	-
6.92	-	-	4.26	5.68	-	-	11.28
-	7.03	9.77	-	-	6.21	19.87	-
8.5	-	~~9.77~~	3.53	4.71	-	-	6.33
-	10.4	8.48	-	-	8.44	16.59	-
15.52	-	-	9.68	6.75	-	-	10.75
-	12.53	10.85	-	-	7.60	18.19	-
8.87	-	-	7.70	8.50	-	-	19.33
-	10.59	7.80	-	-	7.80	25.39	-
9.96	-	-	5.14	7.33	-	-	12.24

TABLE 5

R1

P3$\bar{t}$-P1$\bar{t}$

3 $\bar{t}_c$(SECONDS)

LP				KYB			
UIC1		UIC2		UIC1		UIC2	
01	02	01	02	01	02	01	02
-	0	0	-	-	18.35	14.42	-
0	-	-	0	0	-	-	0
-	0	0	-	-	7.00	24.12	-
0	-	-	12.34	0	-	-	0
-	25.14	10.81	-	-	0	22.56	-
0	-	-	18.87	21.37	-	-	26.46
-	0	16.64	-	-	14.91	0	-
0	-	-	17.83	17.23	-	-	0
-	26.53	20.51	-	-	0	0	-
0	-	-	12.42	21.97	-	-	17.0
-	25.4	20.40	-	-	0	0	-
0	-	-	20.70	0	-	-	25.42
-	0	0	-	-	0	0	-
0	-	-	20.29	20.65	-	-	0
-	0	0	-	-	0	0	-
0	-	-	0	0	-	-	0

ABLE 6

R1 N P2

LP				KYB			
UIC1		UIC2		UIC1		UIC2	
01	02	01	02	01	02	01	02
-	5	11	-	-	8	6	-
2	-	-	5	1	-	-	3
-	1	10	-	-	2	21	-
10	-	-	1	0	-	-	3
-	4	2	-	-	2	2	-
4	-	-	10	5	-	-	3
-	1	5	-	-	1	1	-
4	-	-	3	3	-	-	1
-	1	4	-	-	1	7	-
0	-	-	4	2	-	-	0
-	0	2	-	-	2	6	-
1	-	-	0	1	-	-	1
-	0	0	-	-	0	1	-
0	-	-	2	7	-	-	1
-	7	19	-	-	3	18	-
2	-	-	3	3	-	-	1

R2

P1 $\bar{t}$

TABLE 7

R2

P1 $\bar{t}$

(SECONDS)

LP				KYB			
UIC1		UIC2 (withdraw)		UIC2		UIC1 (withdraw)	
01	02	01	02	01	02	01	02
-	3.96	3.13	-	2.48	-	-	2.69
2.74	-	-	2.003	-	1.70	1.98	-
-	2.69	3.38	-	3.92	-	-	3.67
2.78	-	-	1.56	-	2.24	2.058	-
-	2.94	2.09	-	2.69	-	-	1.73
1.66	-	-	2.18	-	2.07	2.52	-
-	4.27	3.22	-	2.37	-	-	2.58
2.13	-	-	2.11	-	2.20	2.20	-
-	1.24	1.94	-	2.67	-	-	1.73
2.00	-	-	2.41	-	2.61	1.73	-
-	2.26	3.84	-	3.79	-	-	2.48
3.83	-	-	2.67	-	2.53	2.41	-
-	4.71	3.77	-	4.87	-	-	2.71
2.43	-	-	3.72	-	2.63	2.59	-
-	2.22	2.76	-	1.97	-	-	1.31
5.00	-	-	2.38	-	3.22	2.29	-

P2 n

LP				KYB			
UIC1		UIC2		UIC2		UIC1	
01	02	01	02	01	02	01	02
-	0	6	-	0	-	-	0
1	-	-	2	-	0	2	-
-	0	1	-	3	-	-	0
0	-	-	0	-	0	0	-
-	0	0	-	0	-	-	0
0	-	-	1	-	1	2	-
-	0	10	-	1	-	-	3
0	-	-	1	-	0	0	-
-	5	4	-	0	-	-	2
1	-	-	1	-	2	0	-
-	0	4	-	2	-	-	0
1	-	-	0	-	2	0	-
-	2	3	-	0	-	-	0
0	-	-	6	-	0	2	-
-	0	2	-	0	-	-	0
0	-	-	2	-	0	1	-

ABLE 8

2

2 n

P2 $\bar{t}_c$

TABLE 9

R2

P2 $\bar{t}_c$

(SECONDS)

LP				KYB			
UIC1		UIC2		UIC2		UIC1	
01	02	01	02	01	02	01	02
-	0	8.50	-	21.16	-	-	0
7.76	-	-	7.87	-	0	4.75	-
-	0	4.3	-	8.91	-	-	0
0	-	-	0	-	0	0	-
-	0	0	-	0	-	-	0
0	-	-	6.9	-	2.03	5.3	-
-	0	6.9	-	27.31	-	-	6.2
0	-	-	8.7	-	0	0	-
-	1.24	10.98	-	0	-	-	4.02
5.5	-	-	5.6	-	4.89	0	-
-	0	8.8	-	25.71	-	-	0
2.83	-	-	0	-	15.17	0	-
-	18.29	5.6	-	0	-	-	0
0	-	-	7.9	-	0	11.3	-
-	0	6.0	-	0	-	-	0
0	-	-	4.8	-	0	13.6	-

P3 n

LP				KYB			
UIC1		UIC2 (withdraw)		UIC2		UIC1 (withdraw)	
01	02	01	02	01	02	01	02
–	0	1	–	0	–	–	0
0	–	–	0	–	0	0	–
–	0	0	–	0	–	–	0
0	–	–	0	–	0	0	–
–	1	0	–	0	–	–	0
0	–	–	0	–	0	0	–
–	0	4	–	0	–	–	1
0	–	–	1	–	0	1	–
–	10	1	–	0	–	–	0
0	–	–	0	–	0	0	–
–	0	0	–	0	–	–	0
0	–	–	0	–	1	0	–
–	0	0	–	0	–	–	0
o	–	–	0	–	0	0	–
–	0	0	–	0	–	–	0
1	–	–	0	–	0	0	–

BLE 10

2

3 n

P3 $\overline{tc}$

LP				KYB			
UIC1		UIC2		UIC2		UIC1	
01	02	01	02	01	02	01	02
-	0	18.2	-	0	-	-	0
0	-	-	0	-	0	0	-
-	0	0	-	0	-	-	0
0	-	-	0	-	0	0	-
-	26.56	0	-	0	-	-	0
0	-	-	0	-	0	0	-
-	0	16.4	-	0	-	-	16.2
0	-	-	14.7	-	0	16.5	-
-	0	22.9	-	0	-	-	0
0	-	-	0	-	0	0	-
-	0	0	-	0	-	-	0
0	-	-	0	-	26.97	0	-
-	0	0	-	0	-	-	0
0	-	-	0	-	0	0	-
-	0	0	-	0	-	-	0
20.1	-	-	0	-	0	0	-

TABLE 11

R2

P3 $\bar{t}_c$

(SECONDS)

BLE 12

UMBER OF
VES TO
LVE

LP				KYB			
UIC1		UIC2		UIC2		UIC1	
01	02	01	02	01	02	01	02
-	24	32	-	27	-	-	21
18	-	-	15	-	18	26	-
-	15	16	-	28	-	-	15
28	-	-	16	-	25	15	-
-	93	16	-	28	-	-	15
19	-	-	18	-	19	26	-
-	38	54	-	50	-	-	15
16	-	-	19	-	16	17	-
-	35	78	-	22	-	-	20
15	-	-	18	-	15	15	-
-	15	28	-	67	-	-	15
32	-	-	35	-	27	16	-
-	30	19	-	18	-	-	15
16	-	-	28	-	24	21	-
-	17	20	-	17	-	-	15
15	-	-	16	-	24	15	-

TABLE 13

LP. UIC1	N ↓				N ↑			
Results only	E ↑		E ↓		E ↑		E ↓	
Measure	01	02	01	02	01	02	01	02
R2 P1 t	3.223	2.111	3.384	1.558	2.090	2.184	3.125	2.002
	2.76	2.375	3.840	2.672	3.765	3.72	1.94	3.208
R2 P2 $\bar{t}_c$	6.88	8.76	4.28	0	0	6.09	8.55	4.70
	6.00	4.75	8.78	0	12.09	7.90	11.06	3.52
R2 P3 n	4	1	0	0	0	0	1	0
	0	0	0	0	0	0	1	1
R2 P3 $\bar{t}_c$	16.44	14.76	0	0	0	0	18.18	0
	0	0	0	0	0	0	22.97	12.72
No. of Draws/ Moves	30/54	10/19	16/16	10/16	3/18	17/28	28/32	9/15
	8/20	8/16	27/28	18/35	11/19	20/28	17/78	4/18

TABLE 14

R1 P1 $\bar{t}$ (SECONDS)

DIALOGUE TYPE I (UIC 1) DIALOGUE TYPE II (UIC

	LP		KYB		LP		KYB	
	N↓	N↑	N↓	N↑	N↓	N↑	N↓	N↑
	8.04	3.61	3.97	2.64	3.74	3.13	5.26	4.11
	7.28	2.87	3.92	2.83	4.05	2.72	4.08	5.13
	3.17	4.38	4.42	4.22	3.93	4.56	3.95	6.17
	4.1	4.08	5.32	3.5	4.48	3.14	5.50	4.04
	3.58	3.5	4.27	2.91	2.96	3.21	5.22	3.50
	2.96	5.97	3.55	4.92	2.69	4.78	4.70	7.88
	3.91	5.36	2.50	3.72	4.53	4.23	4.11	6.94
	7.71	5.42	5.7	4.52	4.08	3.49	6.12	4.08
Σx	40.75	35.43	33.65	28.95	30.36	29.26	38.47	42.02
Σx^2	240.81	163.09	148.54	108.52	119·07	111·07	193·85	236·74
Mean	5.09	4.42	4.21	3.62	3.79	3.63	4.81	5.25
SD	2·05	0·72	0·92	0·67	0·72	0·71	1·05	1·42

SURE	SOURCE	DF	SSQ	MSS	F	MEASURE	SOURCE	DF	SSQ	MSS	F
1 1 $\overline{t}$	S	15	51.44	3.42	4.09	R2 P1 $\overline{t}$	S	15	22.93	1.53	3.
	I	1	0.87	0.87	–		I	1	1.38	1.38	3.
	D	1	0.007	0.007	–		D	1	0.20	0.20	–
	IxD	1	15.73	15.73	18.8*		IxD	1	1.88	1.88	4.
	E	45	37.59	0.83			E	45	18.30	0.41	
	TOTAL	63	105.64				TOTAL	63	44.61		
1 2 n	S	15	386.18	25.7	1.49	R2 P2 n	S	15	53.30	3.55	1.
	I	1	0.76	0.76	–		I	1	14.10	14.10	5.
	D	1	83.26	83.26	4.95*		D	1	16.00	16.00	6.
	IxD	1	0.391	0.391	–		IxD	1	17.25	17.25	6.
	E	45	773.91	17.19			E	45	118.8	2.64	
	TOTAL	63	1244.5				TOTAL	63	219.8		
1 2 $\overline{t}_c$	S	15	407.87	27.19	2.51	R2 P2 $\overline{t}_c$	S	15	469.0	31.2	–
	I	1	67.81	67.81	–		I	1	7.68	7.68	–
	D	1	60.87	60.87	–		D	1	215.8	215.8	5.6
	IxD	1	540.22	540.22	49.8*		IxD	1	0.148	0.148	–
	E	45	487.81	10.84			E	45	1759.4	39.09	
	TOTAL	63	1564.4				TOTAL	63	2542.2		
1⌿ 3 n	S	15	38.1	2.53	1.02	R2⌿ P3 n	S	15	30.69	2.05	1.5
	I	1	12.3	12.3	–		I	1	4.00	4.00	2.9
	D	1	15.03	15.03	–		D	1	7.81	7.81	5.7
	IxD	1	19.29	19.29	7.77*		IxD	1	12.37	12.37	9.0
	E	45	111.91	2.48			E	45	61.76	1.37	
	TOTAL	63	196.5				TOTAL	63	116.44		
1⌿ 3 $\overline{t}_c$	S	15	1585	105.7	1.05	R2⌿ P3 $\overline{t}_c$	S	15	681.9	45.46	–
	I	1	0.20	0.20	–		I	1	3.9	3.9	1.0
	D	1	163.3	163.3	1.64		D	1	15.01	15.01	–
	IxD	1	113.53	113.53	1.14		IxD	1	5.85	5.85	–
	E	45	4478.9	99.53			E	45	2426	53.9	
	TOTAL	63	6341				TOTAL	63	3183		
rall e to ive	S	15	1610.8	107.4	17.2	R2⌿ P3 $\overline{t}_c$ (Light Pen)	N	1	31.26	–	0.7
	I	1	2.08	2.08	–		E	1	31.26	–	0.7
	D	1	345.3	345.3	55.4*		O	1	57.78	–	1.3
	IxD	1	86.02	86.02	13.7*		NxE	1	449	–	10.7
	E	45	280.7	6.23			ExO	1	45.7	–	1.0
	TOTAL	63	2324.9				OxN	1	45.7	–	1.0
							NxExO	1	57.78	–	1.3
							ERROR	8	332.7	41.58	
							TOTAL	15	1051.3		

: * = 5% level significance
⌿ = non-normal distribution

TABLE 15

2x2 Mixed Model ANOVA - Significant Results

TE: Assumes SxI, SxD, SxIxD interactions are not significant.
Conservative F-test used for 5% level significance (related sample mixed model); df = (1, 15).

Assumes also that order effects are not significant.

Measure	Factor	Between D's		Within D's	
		LP	KYB	D1	D2
R1 P1 $\bar{t}$	I x D	2.67*	-2.92*	NT	NT
R1 P2 n	D	-2.02	-1.31	NT	NT
R1 P2 $\bar{t}_c$	I x D	4.48*	-5.68*	NT	NT
R1 P3 n	I x D	-1.97	0.76	NT	NT
R2 P1 $\bar{t}$	I x D	0.83	2.59*	NT	NT
R2 P2 n	I D I x D	-2.75*	-0.16	-0.33	2.62*
R2 P2 $\bar{t}_c$	D	-2.43*	1.59	NT	NT
Overall Time	D I x D	-1.39	-2.94*	NT	NT
R2 P3 $\bar{t}_c$		N1		N2	
		E1	E2	E1	E2
		↔ 0.84	↔ 0	↔ 1.91	
			0.84		1.91

TABLE 16

Paired t-values for significant results of Table 15

* means significant @ 5% using 2 tailed test

NT " Not Tested