Individual Differences In User Performance on Command Line and Direct Manipulation Computer Interfaces.

Konrad Morgan

Ph.D.
University of Edinburgh
1989.
Abstract.

This dissertation describes 4 experiments which formed part of an investigation into human-machine anomalies, attitudes towards technology, and a comparison of two computer interfaces (direct manipulation, and command line). A computer database system with these two different interfaces was developed, along with a series of questionnaires. The parapsychological investigation used random number generators of various types and complexities, the behaviour of which determined the performance of the computer system. This methodology was adopted to investigate a series of empirical and philosophical hypotheses.

The HCI interface data showed that over 3 different subject populations (computer naive & keyboard literate, computer naive & keyboard illiterate, and finally experienced computer users), the direct manipulation interface was superior with regard to several measures. It was faster (p < 0.05), had fewer errors (p < 0.01), and was rated on post-session questionnaires as being better (p < 0.01), more satisfying (p < 0.05), more stimulating (p < 0.01), and easier to use (p < 0.01) than the comparable command line system. However, users did not rate the direct manipulation interface as being more powerful than the command line system. The technology attitudes questionnaires showed significant correlations between early parental encouragement of machine exploration, and later technology competence.

The parapsychological aspects attempted to develop a secure computerised methodology for investigating the possibility that humans interact with computer systems using as yet unknown methods. These experiments found little evidence of any parapsychological effects, but concluded that the area needed more empirical investigation.
# Table of Contents

1. **Background to the Study.** - 5

1.1. **Introduction.** - 5


2.1. **Introduction.** - 32

2.2. **The RNG Devices.** - 33

2.3. **Experiments Designed To Demonstrate An Effect Which Appeared To Show Subjects Guessing The Outcome Of A Random Event Generator (REG).** - 40

2.4. **Experiments Designed To Demonstrate An Effect Which Appears To Influence The Outcome Of An REG.** - 49

2.5. **Investigations Into The Effect Of Varying The Behaviour Of The RNG.** - 55

2.6. **Investigations Into PK-RNG Conducive States.** - 68

2.7. **Learning PK.** - 78

2.8. **Animal PK.** - 81

2.9. **Allobiofeedback.** - 86

2.10. **Theories Of PK.** - 87

2.11. **Review Of Discussions Of The Use Of RNGs.** - 96

2.12. **Alternative Hypothesis For These Effects.** - 97

2.13. **Discussion.** - 101


3.1. **Security Of Parapsychological Computer Systems.** - 102
3.2. The Most Fundamental Security Problem. - 103.


3.5. Discussion. - 164.


4.1. Introduction. - 168.

4.2. Trying to know the user. - 179.


4.4. KMDB Questionnaires. - 209.


5.1. Introduction. - 212.


5.5. The Use Of These Signal Enhancement Techniques Within This Dissertation's Study. - 219.


5.7. Applied Uses. - 228.

5.8. Non-Experimental Speculations. - 234.

5.9. Discussion Of The Findings From This Literature Review. - 236.

5.10. Possible Applied Uses Of Smart Noise. - 236.

6.1. Introduction - 239.
6.3. Planned analysis - 244.
6.4. Planned Overall Analysis for combined results from experiments 1 to 3 - 248.
6.5. Discussion on the number of planned analyses - 251.
6.7. Pilot Experiment - 259.
7. Experiment 1 - 271.
   7.1. Introduction - 271.
   7.3. Results - 272.
   7.4. Parapsychological Findings - 283.
   7.5. Discussion - 298.
8. Experiment 2 - 302.
   8.1. Introduction - 302.
   8.4. Procedure - 303.
   8.5. Results - 304.
   8.6. Parapsychological Findings - 313.
8.7. Questionnaires. - 314.


9.5. Procedure. - 332.

9.6. Results. - 333.


9.10. Findings From The Comparison Of The Overall Results From Experiments 1 to 3. - 354.

10. Overall Discussion on Experiments 1 to 3. - 373.

10.1. Conclusions From The Comparison Of The Overall Results Of Experiments 1 To 3: Experiment 4. The TAO Factor Analysis And The Overall Conclusions And Recommendations For Future Research. - 373.

10.2. Conclusions, Biology Questionnaire, and Experiment 4. - 377.

10.3. Overall Conclusions, And Possible Directions Of Future Research. - 397.

11. Appendices


11.4. Appendix 4. Full Colour Plates (used to provide digitised images in Chapter 6). - 424.

11.5. Appendix 5. Sample Error Categories and Types. - 427.


11.10. Appendix 10. Experiment 4's Pre-Session Questionnaire. - 461.

11.11. Appendix 11. Experiment 4's Post-Session Questionnaire. - 463.


12. References.
Dedication.

This thesis, and the research described therein is dedicated to memory and ideals of Cynthia and Arthur Koestler, without whose generous bequest this research could never have taken place.

Acknowledgements

I would like to express my gratitude to those people without whose help, advice, or support the research described in this thesis would never have taken place.

Firstly my parents for providing me with such unfailing encouragement, and support. The Koestler Chair of Parapsychology for funding this research for the full three years. My two supervisors Professor Robert Morris, and Dr Hamish Macleod, who have been the best supervisors, and friends I could hope to have had during a PhD. Mrs Shirley Gibbs of Napier Polytechnic of Edinburgh, my HCI co-experimenter, and friend. Dave Nelson, Psychology Department, for his help with regard to the construction of the various questionnaires used in the study. Dr Chris Adie, Dr Sam Wilson, Toby Morris, Christine Rees, Mike Holmes, and Gillian Kirkwood, of Edinburgh University’s Computing Service (EUCS). Damien Murphy - Edinburgh University Chemistry Department for his friendship and support. Napier Polytechnic of Edinburgh for the use of their resources. Colonel and Mrs Bewsher for the kindness they have shown me during my period of study. Dolores Ashcroft-No-wicki for her kindness in providing anecdotal material regarding machine anomalies. Ben Shneiderman - University of Maryland, USA, for his kindness and support in the initial stages of the HCI element of the thesis. Dr Alain Filhol - Institute Laue Langevin, France, for his help and advice regarding random number generators (RNGS). Professor Brian Ripley, Department of Statistics, Strathclyde University (also for help and advice regarding RNGs). Francis Provan - EUCS Statistics Consultant, for her statistical advice. The Royal Statistical Society (for their help with RNGs), The British Computer Society (for the use of their HCI library). I would also like to acknowledge the following companies or institutions for their help in providing information and in many cases either source code, or access to trade restricted information: Anglia PC Users Group, International Business Machines UK plc., Digital Research Inc., Prospero Software Plc., Amstrad (computers) Plc., NASA, Apple Computer Inc., Intel Corp., Motorola Inc., AT&T, Digital Equipment Corp., and
Microsoft Corp. Finally, I must acknowledge the numerous workers at the Koestler Chair of Parapsychology, who have had such a major affect on this work by their critical feedback and support: John Beloff, Helen Simms, Deborah Delanoy, Shari Cohn, Julie Milton, my friend Reimar Gissurarson, Caroline Watt, Robin Taylor, and last but not least Richard Wiseman.

Declaration

I hereby certify that dissertation is my own composition, and that all the work described therein is my own.

Konrad Morgan
1. Background to the Study.

Quote:
Evidence! Hah! You can prove ANYTHING with evidence!
Dr Sam Wilson\(^1\), Edinburgh University. 1988.

1.1. Introduction.

There are many problems involved with starting a project which attempts to investigate alleged paranormal events. One of these problems is that many eminent scientists view, and have viewed, the paranormal with suspicion (McClenon, 1982). Since these scientists are supposed to have an understanding of the current limits of human knowledge, the study of the phenomena could be seen as being hard to justify. Perhaps this is a good point to investigate what it is about paranormal phenomena that these scientists find so hard to accept. These skeptical views can be summarized by a quote from Jacob Bronowski, whose work I view with admiration, both for its technical excellence, and its insight into the human condition. Bronowski uses the concept of Extrasensory Perception (ESP) in his essays on 'The Identity of Man', when he discusses how human knowledge advances by man’s ability to see how the explanations for phenomena blend into a coherent model of how the universe works. He states his objections to paranormal phenomena\(^2\) upon these constructionist grounds:

The suspicion with which all scientists treat the published evidence for extrasensory perception shows this. A set of results is odd and unexpected, in the end it is unbelievable, because it outrages the intricate network of connections that has been established between known phenomena (Bronowski, 1967, p43).

Perhaps the point that Bronowski and other critics (Kurtz, 1978 & 1985; Alcock, 1981; O’Connor, 1986) miss, is that if such paranormal phenomena do exist then it will merely

---

1 Dr Wilson intended this comment humorously.

2 For this discussion the term 'paranormal phenomena' will refer to any supposed event, or events which cannot be explained by current scientific knowledge.
mean that the connections that science has so far established between phenomena was incomplete; or there were phenomena in existence not accounted for in that model. Such changes to the set of accepted scientific knowledge have happened before, and it is likely they will happen again. Indeed skeptics and parapsychologists have both admitted that this conflict with accepted scientific knowledge is not, in of itself, a justification to dismiss parapsychology (Edge, 1977; LeShan, 1978; Pratt, 1979; Truzzi, 1980; Stanford, 1982; Hovelmann, 1983; Palmer, 1986).

However I feel that the rational objection raised by Bronowski, may not be the motivation behind many in the skeptical community. Perhaps the most vocal representative of the skeptical community is a Canadian psychologist, James Alcock. He has stated (Alcock, 1981) that it is the inherently dualistic perspective of parapsychological research that is one of the primary motivations for his opposition to parapsychology. This inherent dualistic interpretation of parapsychological phenomena has been frequently noted by scientists. Nearly 400 years ago, Francis Bacon noted that those who believed in psi accounted for it by postulating that the mind in some way withdrew from the body to allow such phenomena to take place. To justify this Bacon noted that dualists proposed psi events appear most in sleep, in ecstasies, and near death (Bacon, 1605, p119).

The astuteness of Bacon’s observation can only be marvelled upon because it closely reflects the views that many parapsychologists still hold today. Indeed if we changed 'ecstasies' to altered states of consciousness (ASCs), and 'withdrew' to expanded beyond, the phrase could have been quoted from last years Parapsychological Association convention.

Even if one accepts Alcock’s accusation of a dualistic agenda to all parapsychological research (which is plainly not the case), that would not vindicate the rejection of the findings generated by such research. To reject work purely upon the basis of the belief system held by the person conducting that research is highly unscientific. Yet this is the very accusation which Alcock makes of parapsychology. Besides, there are good reasons for adopting the attitude that psi may have something to do with the mind body problem (Beloff, 1977, p94).

It is hard to know if Bacon was reflecting his own views upon this subject, since popular belief attributed such occurrences to the work of discarnate entities (personal communication Morris, 1989).
There are other reasons why the removal of parapsychology would involve a loss to academia. The study of such phenomena encourages investigators to be simultaneously open to the bizarre, and yet to remain discriminating of every claim (Beloff, 1975). Most academics would agree that these are attributes which are to be encouraged in all scientific inquiry.

However it was not these problems that the author found most daunting when he started the inquiry addressed by this thesis. Instead it was being faced with a mass of information, some of which was scientific, some anecdotal, but all incredible, at least to current science. The writer Michael Shallis described this problem beautifully in his book ‘On Time’. In which he said

The difficulty for the scientist faced with the paranormal is that he really has to take it all (Shallis, 1982, p1).

The problem is that scientific training molds the mind into liking definite answers, and the field (except in a few exceptional circumstances) provides nothing but ambiguity. The researcher is left, either to adopt Shallis’s approach, or to set their own criteria for what is possible, and accepting only those findings which conform to these implicit parameters. The human mind being what it is most people will adopt the second method (whether they are consciously aware of it or not), accepting facts that match their belief system, and rejecting those that do not. Within this dissertation I have taken what some may regard as a ‘step back’ from the usual definitions used in parapsychology to show, and describe the process of anomalous interaction. This has to be done since, as I hope to show in later sections, it is a mistake to think that one can define the exact process involved in a human machine anomaly.

1.1.1. Background And Aims Of This Study.

One of the aims of the work described in this dissertation was to try to investigate the potential use of experimentally demonstrated anomalous effects with random number generators (RNGs). This was done with the goal of enhancing the performance of computer systems, and to devise a methodology that could lead to the volitional, or mental control of noise driven components of a computer system. The use of parapsychological phenomena to do something useful (the so called ‘applied stage’ of a discipline), is becoming increas-
ingly more frequent, as the review of applied parapsychology in a later chapter of this dissertation shows. This has increased as pressure has mounted from outside parapsychology to show that it can help to solve problems in other disciplines (Alcock, 1982; Morris, 1985).

My idea of investigating the potential of such a 'psi' interface was derived from reading the doctoral dissertation of Julian Isaacs (1984). In this work Isaacs detailed his investigation into the deformations of piezo semi-conductors. It occurred to me that such a device might prove to be an effective method of allowing some kind of interface between mental volition, and computer actions. My original intention was to have five such detectors, one for each cardinal direction, and one to show that an action was required. The signals produced by the deformations of the crystals would determine the position of a pointing device on a computer display. It was my intention that this system be used in much the same way as the standard input device called a 'mouse' (Lu, 1984). My original plan was to develop such a system as a possible aid for totally disabled people. I felt that not only would such a device be useful to this disadvantaged group, but that they would also probably be highly motivated in learning to use and control such a system. However the precautions required to exclude all known non-paranormal artifacts in such crystal deformation work made such a system construction problematic (Isaacs, 1982 & 1984; Hubbard, Bentley, & Pasturael 1987).

Fortunately I became familiar with the work of Robert Morris, and his idea of using the outputs from sources of randomness in just such useful, and applied ways (Morris, 1983). I was lucky that he had recently been appointed to the position of Koestler Professor, and that when I approached him, he shared my enthusiasm for such an investigation. My previous research work had been in improving the design of the interfaces used in computer aided teaching packages. This gave me a background in the problems found in human computer interaction. Due to the amount of overlap in these two domains, it was a natural step to try to create a study which investigated both problem areas.

However before these areas can be discussed the background to such machine interactions and anomalies must be detailed. Mankind's longing for using unknown or 'magical' forces in applied ways has a long history. Some of the earliest examples of human writing held at the British Museum are directions for usefully predicting the future by an almost mechanical method (Babylonian divination tablets). It is reasonable to suppose that right from man's
earliest moments he has tried to use every means at his disposal to gain a supernatural control over his life and environment. The tools which controlled the physical elements of man’s environment made him assume that via some sympathetic action the same types of tools would be able to control the non-physical world (Frazer, 1917). The assumption that some supernatural power or influence is inherent in inanimate objects has a strong appeal to it. Often this leads people to assume that their interactions with inanimate objects are subject to the capricious nature of the object itself. Some anthropologists, and psychologists have suggested that such ‘magical thinking’ maybe due either to a general propensity for humans to draw false correlations from events, a predilection to search for symbolic and meaningful connections in their environment, or a manifestation of the cognitive processing limitations of the human mind. If these explanations are correct it would have implications for the whole of psychical research, not just for the study of computer based anomalies. In order for us to judge this we must review research that has been conducted in this area.

1.1.2. The Effects Of Belief Upon Judgment, Health, Life Style, And Well Being.

This section will overview the literature that is relevant to the role of belief in human randomization, causality judgment, and health. Since this is a large area to overview, and it is not the major topic of this thesis only limited space will be devoted to it. Readers with a more detailed interest in this subject should refer to the bibliography.

Exactly how much an individual’s belief affects their life is underestimated. It not only affects the health of the person but also the attributions that the individual makes about the environment. As such it is a factor which should not be ignored by any investigation into psi phenomena, and people’s attributions towards machine anomalies. Accounts of such machine anomalies are more frequent now than they have ever been. This is due to the increasing number of people who come into contact with computer systems. As this trend will only increase we must look at the human judgment of events, and how the attribution of the causality of those events is affected by the person’s belief.
1.1.2.1. Health

If there is one area where an individual’s beliefs have a strong effect, it is health. These effects are not always in the direction one would expect. One of the earliest psychologists to realize the impact that belief has upon well being was Sigmund Freud. In his accounts of ‘Some character types met with in Psychoanalytical work’, Freud mentions the ’fear of success’ syndrome. This syndrome is where

People fall ill precisely when a deeply-rooted and long cherished wish has come to fulfillment (Freud, 1953).

They seem unable to tolerate their own success, and fall prey to mental or psychosomatic illness⁴.

1.1.2.2. Randomness And Attributed Causality.

Perhaps of more interest is the ability of humans to understand the concept of chance. Several studies have been conducted into the human ability to generate random sequences, or to recognise chance baselines. They have all revealed that humans are very poor at such estimates. This is especially true when that judgment involves some phenomena which is central to the individual’s beliefs. These findings have often led skeptics to attribute some apparently paranormal events to faulty human causality judgments (Marks, 1986), and poor understanding of probability (Blackmore & Troscianko, 1985). For instance some skeptical anthropologists have proposed

That magical thinking is an expression of a universal disinclination of normal adults to draw correlational lessons from their experience, coupled with a universal inclination to seek symbolic and meaningful connections among objects and events (Sheweder, 1977 p637).

A review of the literature shows that these accusations are only valid for some anecdotal evidence, and cannot be generalized to the bulk of experimental results produced by

---

⁴ This 'fear of success syndrome' may play a role in some negative psi results in parapsychological experimental research.
parapsychology. It would only be a valid objection if the order of targets to be used in an experiment had been assigned arbitrarily by the experimenter. Although this might have happened in the early days of parapsychology, modern standards are much higher.

1.1.2.3. Randomization.

One of the earliest investigations into the randomization ability of humans was conducted by Ross (1955). He reported that when Ss were asked to randomize a series of binary symbols, the series they produced had minor deviations from classic randomness. There was not very much notice taken of this finding, and so it was not until five years later that Ross's findings were followed up. This follow up study (Bakan, 1960) used a similar binary randomization task. Bakan reported finding no first order preferences for either of the two symbols. However he did report that his Ss produced a higher incidence of symbol alteration than expected by chance. He felt that this was due to Ss having a false concept of what randomness looked like, this explanation became known as the 'cognitive hypothesis' (Zahn, 1982). The alternative interpretation on these findings is that put forward by Weiss (1964). He proposed that the simulation of randomness was not an attempt to realize a defective idea, but an effort to overcome an interfering response bias. Weiss proposed that the creation of random sequences required judgment and correction. This has been called the 'motivational hypothesis' (Zahn, 1982). To gather more data on this problem a detailed study was conducted by Bath (1966) two years after Weiss put forward his motivational hypothesis. Bath got his Ss to try and randomize binary, decimal, and alphabetical symbols. He reported finding significant first order deviations from chance in most Ss with the binary symbols, and all Ss with the remaining symbols. The rate of character production decreased as the number of alternatives to be randomized increased.

1.1.2.3.1. Psi Skeptics And Their Interpretations Of These Findings.

Skeptical research in this area has concentrated upon the cognitive hypothesis for explaining the inability of Ss to create random sequences. This is possibly because proposing a cognitive

---

5 An example of this was the pioneer of mechanical REGs (Tyrrell, 1937). He frequently turned off the electric REG, and mentally generated his Ss ESP target sequences, in the presence of his subject. He did this because she complained about the noise the REG made, and scored better when it was not used (West, 1962 p160).
deficit as the basis of 'psi' beliefs is both derogatory to believers, and flattering for the skeptic. However there is experimental evidence that both cognitive and motivational explanations play an equal part in Ss inability to create random sequences. Zahn (1982) reported an experiment which systematically varied the cognitive (Bakan, 1960) and motivational (Weiss, 1964) factors involved in a computer controlled binary randomization task. Zahn reported finding evidence for both explanations, and concluded that both played a role in Ss inability to generate such sequences. Some research has even shown that the factors that bias inference need not enhance the likelihood of an error, since in real world situations there are no secure criteria to the validity of an inference (Kruglanski & Ajzen, 1983).

1.1.2.3.2.Randomness And Spontaneity.

Slak and Shaffer (1984) tried to investigate if there was any correlation between the sequences produced in spontaneous sequencing and those produced when Ss were asked to randomize. In a split condition study 106 Ss were given a spontaneous ordering task, and a randomization task\(^6\). The order of presentation was split to avoid any bias. The authors reported finding better random sequences were produced by Ss who had just completed the spontaneous sequencing task. The authors interpreted this as evidence that the biases normally found when Ss are asked to produce random sequences are due to a cognitive effect. This evidence would seem to confirm Bakan's (1960) hypothesis, although there are alternative explanations (see footnote).

1.1.2.3.3.Intelligence And Randomization.

Skeptics have often reported that belief in psi could be attributed to the lower Intelligence Quotients (IQs) levels which they claim are associated with an inability to randomize (Singer & Benassi, 1981). However these claims are not substantiated by the literature. Jones, Russell and Nickel (1977) reported finding that paranormal belief was correlated with high levels of intelligence. This latter result contradicts the claims made by skeptics, but it may be that both extreme findings are due to some bias in the samples used. In another study

---

\(^6\) It is likely that less notice was taken of feedback in the spontaneous ordering task, and that the use of reduced feedback would tend to be carried over into the next task. This could be a biasing factor, since research has shown the role of feedback to be of importance in randomization tasks (Neuringer, 1986).
the similar theory that dogmatism was correlated with low intelligence was falsified (Steininger & Colsher, 1979). The authors looked at the dogmatism and Machiavellianism scales of 109 Ss (64 female, and 45 male) and their results from a digit span recall test. Steininger and Colsher reported that dogmatism was significantly correlated with recall, but only for female subjects. The authors felt that the modern world required that intelligent females assumed belief in some cause, or authoritarianism. Kreitler and Kreitler (1986) investigated the effects of the judgment of probability and IQ in three groups of children aged 5-6, 8-9 and 11-12 years respectively. 240 Ss were given the Stanford-Binet Intelligence Scale, and four probability based tasks. These included randomization, and odds estimation. The authors reported that age showed some clustering effects (increases in ability were larger in the first two age groups), a positive correlation between ability in all the tasks, and that IQ correlated with the more general ability measures, but not specifically with the ability to randomize. They also reported that the sex advantage differential for males increased with age. Other researchers have tried to steer away from a cognitive explanation. Tresiman and Faulkner (1987) proposed that rather than attribute the human inability to create random sequences to faulty cognitive operations (or concepts of randomness), it should be explained by a basic psychophysical selection mechanism. Using a model they first presented in 1984 they proposed that humans have an internal random generator, the effects of which can be distinguished from those of a cognitive bias.

1.1.2.4. Causality Judgments.

A separate, but closely linked claim often made by skeptics is that psi-believers have poorer causality judging abilities than skeptics. In one of the earliest studies by psi skeptics (Tversky & Kahneman, 1974) believers in psi were reported to have been found to have a greater misunderstanding of causality and probability than the general population. Their research did not point to any specific factors involved in such mis-attributions, and so was not helpful in isolating any situations in which these misunderstandings might occur. In contrast more objective research has tried to pinpoint these relevant factors. An example of this was Langer’s (1975) investigation into causality judgments when Ss were involved in activities which included factors associated with skill situations. These factors include

---

7 Some of the findings generated by skeptics have failed to be replicated by other experimenters (Nanko, 1986).
competition, choice, familiarity, and involvement. Langer found that when any or all of these factors were introduced into a chance situation Ss felt inappropriately confident of success. Langer was the first to term this misplaced confidence 'the illusion of control'. Obviously this is of importance to parapsychological experimentation. Just such factors are introduced into all psi experiments, for both the experimenter and the Ss. The experimenter's responsibility is to be sure that such illusory factors do not influence the assessment of the results.

A follow up study to this looked at the effects of trends in results upon Ss attributions to a task (Langer & Roth, 1975). In this experiment Ss were asked to guess the results of a coin toss. The Ss received ascending, descending or random results from their guesses. Langer and Roth predicted that early success would induce a skill orientation towards the task. The authors reported that Ss given such descending results felt they could influence the outcome of the coin (or predict it). This group also over remembered past success and expected more successes in the future than the other two groups. Langer and Roth also noted that involvement had the effect of increasing Ss expectations and increased evaluations of past performance. Such findings are of direct relevance to the reports of spontaneous cases, and of especial relevance to machine breakdowns. They show that only a few meaningful coincidences are enough to generate misleading causality judgments.

In another skeptical study Ayeroff and Abelson (1976) looked at what they termed ExtraSensible Beliefs (ESB), and ESP. They got Ss to complete 100 fixed choice ESP trials, which had covertly manipulated conditions. After each trial Ss said if they thought a hit or a miss had taken place. Ayeroff and Abelson split the Ss conditions so they compared several separate effects. The first of these was allowing the Ss to choose the card symbols and letting the sender hand shuffle the deck of cards. The second was allowing Ss 'warm up periods'.

---

8 Hand shuffling of ESP decks by Ss is so open to the possibility of fraud (and is such a poor randomizing technique) that it is not used in experimental work. Even allowing the experimenter to hand shuffle target decks was stopped after the discovery of the so called 'psychic shuffle' (Rhine, Smith & Woodruff, 1938).
Ayeroff & Abelson reported that both of these factors increased ESB independently of ESP

One of the skeptics most embarrassing findings came from an experiment into PK dice throwing (Benassi, Sweeney, & Dreven, 1979). In these series of four experiments the authors tried to show that Ss estimates of success at a PK task were affected independently of actual success. In the first experiment those Ss who were given positive introductory results, or no introductory session reported more illusory control than those given negative introductory results. The second experiment investigated the effect of belief and the number of practice trials Ss were allowed before each session. The most illusory control was shown by believers in psi (known as 'Sheep') with 10 practice trials, and the lowest by skeptics (known as 'Goats') with one practice trial. The third study found that Ss who were actively engaged judged their performance as being better than those who had been passively engaged. The final experiment looked at the effect of locus of control (LOC). The authors reported that when Ss became actively involved, those with an internal LOC gave higher estimates of their success than similarly involved Ss with an external LOC. The authors reported there was no difference between LOC groups when they were passively involved. Benassi, Sweeney, and Dreven felt that these findings supported Langer's "illusion of control theory".

1.1.2.4.1. Faulty Causality And Cognitive Deficit.

Singer and Benassi published another paper two years after their previous one on illusory control. In this later paper they addressed what they felt were the foundations and social processes involved in all so called 'magical' thinking (Singer & Benassi, 1981). Using some of the findings from cognitive science they proposed that flaws in human inference, and increasing environmental uncertainty explained the increasing number of believers in psi phenomena. To back these claims Singer and Benassi reported the results of a series of

---

9 The rate of success was within MCE.
10 Combined PK scores for three of the experiments which used a mechanical dice thrower = + 55.5% (three standard deviations from MCE).
11 This dimension of psi belief was first proposed by Schmeidler & McConnell (1958 & 1973). The terms 'Sheep' and 'Goats' they introduced are taken from Matthew 5:32, where those who believed were 'Sheep', while those who disbelieved were 'Goats'.

15
experiments with a very small selected sample of Ss (6). In these experiments they reported finding that the ’psi believers’ showed a lower understanding of causality than skeptics. These findings have been called into question by the work of Nanko (1986) which used a much larger sample, and failed to replicate this finding.

Two skeptics Blackmore and Troscianko (1985) reported experimental evidence that believers in psi were less able to judge probabilities than skeptics. Using a computer controlled probability task Blackmore and Troscianko conducted two experiments. The first involved 50 female Ss, and the second involved 100 university students. In both these experiments the authors reported that sheep performed worse than goats on most tasks requiring a probability judgement. In a third experiment the authors looked at the hypothesis that sheep were worse at making causality judgments in chance events, the so called ‘illusion of control’ (Langer, 1975). In a computer controlled coin tossing task, the authors reported that sheep felt they were exercising more control than did goats. Sheep were also reported to underestimate the number of hits they had made, and to underestimate the MCE. Blackmore and Troscianko interpreted these findings as evidence that belief in psi was attributable to what they termed a ‘chance baseline shift’.

### 1.1.2.5. Perceived Control And Belief In Psi

Davies and Kirkby (1985) investigated the relationship between perceived control and belief in psi. They issued a battery of perceived control questionnaires to 95 college students. They reported finding two major orthogonal variates. Expectancies of external control were linked with belief in religion, and interpersonal control was linked to belief in witchcraft and parapsychological phenomena. The authors interpreted these results as supporting the theory that a belief in the paranormal was not uni-dimensionally related to perceived control. In the same year Martin Davies (1985) also published the results of an investigation into the links between self-consciousness and belief in psi. He reported finding that high self consciousness and low social anxiety were associated with psi beliefs.

---

12 As with most of these experiments a possible objection is that university students are not a representative sample of the population.
Wicker, Richardson, and Lambert (1985) looked at psychological well being and paranormal belief. They tested 106 Ss with measures that included Jones’ Irrational Beliefs Test (IBT), and other general psychological tests. They reported finding a general relationship between irrational beliefs and emotional difficulty. The specific patterns differed for different subgroups of the IBT. Correlations confirmed the construct validity of the IBT, but also suggested that they reflected motives, and personality constructs. The authors suggested that subscore composites of the IBT may be of more use than the overall score. This reflects that some selected items on the IBT paranormal belief scales are more valid as indicators of psychological disharmony than others.

In his doctoral dissertation Nanko (1986) reported on a study in which he gave 125 Ss an extended paranormal belief scale questionnaire (EPBS). He found no difference in the critical thinking ability of the three main groups, believers in the whole set of parapsychological phenomena, those who just believed in the phenomena reported in academic parapsychology\textsuperscript{13}, and skeptics. The only difference he reported was that traditional religious beliefs correlated with dogmatism, and were negatively correlated with psi belief.

From the research which has been reported after the dissertation’s review finished (1986) only one was judged to be so important that it should be summarized. This study reported finding significant correlations between psi experiences, particularly out of the body experiences (OBE), and narcissism (Tobacyk & Mitchell, 1987). This contradicted the earlier findings of Locke and Shontz (1983), who could find no psychological differences between control Ss and Ss who had OBE like near death experiences (NDEs)\textsuperscript{14}.

1.1.2.6. Computer Anomaly Attributions.

These research findings give evidence that mis-attributions of causality are not directly connected with intelligence. Instead they may be connected with two separate parts of any event. For example in a computer anomaly the user has been found to have two separate

\textsuperscript{13} Note that this category was not present in the other previous studies.

\textsuperscript{14} The NDE-OBE may be in a different category than the ‘normal’ OBE, although Tobacyk and Mitchell made no such distinction in their study.
models of the actions being performed. One of these covers the series of actions needed to complete a task (semantic task concept), and the other is the knowledge of what the task actually involves the system in performing (semantic computer concept). For example, Van-Eekhout and Rouse (1981) reported on a study of 36 professional engineers, and their judgement of the cause of failures in a simulated supertanker control room. Using verbal transcripts, computer logs, observers, questionnaires and interviews, of each simulation they analysed the actions, and attributions of the 36 Ss. They reported that most misinterpretations of the cause of faults were correlated with a lack of knowledge of the system, and the functions of the controlling equipment. In contrast errors in procedure were almost entirely due to the poor design of the control displays. From these findings we can postulate that it is not low intelligence, or faulty cognition that is responsible for many reported computer anomalies. It is more likely to be a mismatch in the users models, so the user thinks a command involves the system one set of actions, when in fact it involves another. This could be due to a poor design of the system, not to any inherent limitation of the users mind (Norman, 1988).

1.1.2.7. Psi And Belief.

The reasons why the skeptical community seems to have spent so much effort investigating the role of expectancy and belief could be to gain an understanding of why certain people believe in psi. It could also be because the effect of belief is one of the more repeatable parapsychological findings. There is experimental evidence that genuine psi works best for those that believe in it (Schmeidler & McConnell, 1958 & 1973; Mischo, 1982; Schmeidler, 1985). For example Judith Taddonio (1975) reported on a study that investigated the effects of induced expectancy about psi performance and ESP scores. She reported finding that Ss who expected to perform well at an ESP task did significantly better than those who did not. Obviously these expectancy effects will be exactly the same effects as found in Langey’s experiments showing ‘illusory control’. Yet the experimental (Isaacs, 1984, Giesler 1984) and anecdotal evidence (Owen & Sparrow, 1976, Batcheldor 1979, 1982, Playfair 1985)

15 These models are outlined in chapter 4.
points towards a seemingly anomalous effect often being preset with this induced feeling of control. The problems belief impose upon objective experimentation could be seen to be worse if one considers another important parapsychological aspect of belief, the experimenter effect (Rosenthal, 1969; Thouless, 1976). Some evidence suggests that to achieve success parapsychologists might have to induce mental states in which both the S and E’s causal judgments could become distorted. If this is the case then it could be dangerous since the experimenter’s judgement of some critical aspects of experimental protocol could be biased. However this need not be a fatal problem, so long as either second parties, or automated controls are in place during experiments, and only objective statistical measures are used in evaluation. In fact strong perceptions of ‘illusory’ control are not in of themselves bad. On the contrary, a body of research shows (Perlmuter & Monty, 1975; Hoekstra, 1983; Staub, 1984; Krause, 1985 & 1986 & 1987; Lorimor et al., 1985; Norris, 1986; Gist, 1987) that high levels of perceived control correlate with improved performance on a variety of tasks, and insulated the person from environmental stress. Other work (Studt, 1977; Pestonjee et al., 1980; Satler, 1980) has shown that people with low belief in their ability to control their environment have a poor quality of life due to their ‘accident proneness’.

1.1.2.8. Discussion On The Role Of Belief.

The areas that we have just discussed are obviously still highly controversial. Some of the research may have been affected by the belief system of the researchers involved, and what they wanted to find. Although this review has found the skeptics explanations of psi belief unconvincing, there are some other correlates with belief in psi that are more problematic. There are indications that belief in psi is correlated with fewer cases of neurosis, but that it may be correlated with more serious mental illness. Eckblad and Chapman (1983) reported finding that Ss who responded more than 1.91 SD above the mean on a magical ideation and psi belief scale had significantly more psychotic experiences than a control group. These 28 Ss from a pool of 1512 Ss had more affective symptoms, and difficulties in concentrating.

16 A potentially complicating factor with the problem of belief is the highly publicized ‘fear of psi’ hypothesis (Tart & Labore, 1986), which postulates that a deep unconscious fear of psi might be responsible for psi’s elusive nature.
Eckblad and Chapman concluded that strong magical thinking and belief in psi correlated with schizotypy. Parapsychologists have tried to experimentally investigate these findings with mixed success (Thalbourne, 1984). Although a pilot study with Ss from a parapsychology class failed to replicate Eckblad & Chapman’s findings, Thalbourne’s major experiment showed a similar trend to Eckblad & Chapman’s reported findings. This raises ethical concerns when parapsychologists start encouraging positive psi belief in their Ss, and in psi training regimes. Such practices may have long term effects on the Ss mental health and well being. Neppe (1984) discussed a recognised psychosis known as subjective paranormal experience (SPE). Its symptoms include subjective feelings, or impressions about the present, past, or future. Although these experiences are shared by ostensibly normal people, Neppe argued that too much reliance upon these intuitions were directly correlated with acute psychosis in some cases. Parapsychologists have therefore to take a responsible attitude to psi training regimes, which actively encourage increased awareness of, and acting upon subjective feelings. Potential Ss have to be briefed about these dangers, and the risks involved in such work, both at the beginning, and throughout such training procedures.

1.1.3. Machine Anomalies.

Humanity still seems to accept that inanimate objects malfunction at the most critical moments. Indeed even the most skeptical have acknowledged that inanimate objects seem to break down at the worst time. In their book 'The Psychology of the Psychic', arch-skeptics Marks and Kammann use the example

Why do I get a flat tyre on the very day I left my jack at home (Marks & Kammann, 1980 p167).

They even went so far as to name it the 'Gremlin Illusion'. They feel that such events are artifacts of humanities selective memory, which has a tendency to recall events that seem anomalous, and in this way exaggerate the recalled consistency. Cognitive psychology and Human Computer Interaction (HCI) specialist Donald Norman also feels that so called 'computer based anomalies' are due to faulty human memory and causality judgments.

---

17 Russel Targ and Arthur Hastings (1987) have warned psychiatrists that claims of psychic abilities and experiences should not be taken as sole indicators of mental illness.
Touch a computer terminal just when it fails, and you are apt to believe that you caused the failure, even when the failure and your action were related only by coincidence. Many of the peculiar behaviors of people using computer systems or complex household appliances result from such false coincidences. (Norman, 1988 p11)

Norman even goes so far as giving an example mis-attribution of a computer user who felt his terminal's breakdown was attributed to a remote library system (Norman, 1988 p39). He later states that when computer users remember computer anomalies only the exceptional events are easily recalled and linked, the rest are forgotten (Norman, 1988 p123). Other work has shown that users are more likely to attribute animate descriptors to the computer when they describe an unsuccessful experience (Wise, Robinson-Stavely, & Nelson, 1987).

The universality of this experience can be seen by the international recognition given to such rules as "Murphy's Law" which is attributed to Captain Edsel Murphy DEC. Murphy was a development engineer on USAF project MX981 at Wright Field Aircraft Laboratory, Ohio, in 1949 (Dickson, 1978 p136). Examples of these informal 'official rules of life' are given here.

My own view is that these laws have a similar rationale behind them. They explain an unusual event, and leave the human operator the dignity of having an excuse for the incident that occurred. There is plenty of anecdotal evidence of people with bad machine reputations. An often quoted example is the theoretical physicist Wolfgang Pauli, whose reputation for breaking sensitive lab equipment was legendary (Gamow, 1959). However one intriguing thing is the number of these reported equipment anomalies that affect people involved in some form of psychic training, or who are known for some apparently inexplicable ability.
(Shallis, 1982). Such people seem prone to problems of this nature with machines (Greenhouse, 1975; Ashcroft-Nowicki, 1986; personnel communication Ashcroft-Nowicki, 1987; Geller, 1986 & 1987; Richards, 1988). A typical example of these accounts would be the following anecdote from the author Steve Richards, who has written a series of psychic development books. I have chosen to paraphrase this example from the many I could have chosen, since it illustrates a typical account of the apparent effect of emotional tension surrounding sensitive electrical instruments. Richards recounts that on one Friday afternoon he was emotionally upset, and in rapid succession every oscilloscope he used broke down for different and inexplicable reasons. So he could get some work done that day, he decided to try to solve some problems on the company computer. This was a time sharing terminal connected to “an 1108” several miles away. However nothing he tried would 'run' (run is a computer term used to describe the execution of a computer program or instruction). The following Monday, he was told that all 14 tape drives on the computer system he had been using had broken simultaneously, but only so it affected Richards, and one other user. Furthermore it was one of the rare failures that totally eluded the diagnostic programs. Richards stated

Any engineer can tell you that when a project deadline date nears, and tension builds among the design team as they race to produce something workable before the customer shows up to use it, things start going wrong with their instruments. This is always mentioned jokingly as just an example of Murphy’s law in action, but I have suspected for some time that these very delicate devices may be subject to being affected by psychic energy in a way that a less delicate device, such as a mechanical contraption is not (Richards, 1988, p40).

Although Richards does not report the details of what “an 1108” computer was, by a little detective work it is possible to determine the identity of 'an 1108'. This is important since this can help us find out how rare a malfunction was with such a machine. Fortunately there are only two computers that have used the name 1108. One was the Xerox 1108 LISP workstation (known as a 'Dandelion'), and the other was the Univac 1108. This was a 36 bit mainframe, introduced as an upgrade to the Univac 1107 in about 196518. It was phased out by the early 1970s. It had 128 'thin film' (glass backed) registers, a large fixed drum storage, and was a heavy user of magnetic tape drives. From these descriptions it is likely that the machine which Richards mentions was the Univac 1108. Unfortunately Univac no

18 Precise dates for the introduction of the Univac 1108 vary from country to country
longer exists (it merged with Burroughs to become Unisys). However these facts allow us to date the events Richards describes to a period when computers were lucky to have a mean time between system crashes\(^\text{19}\) of longer than a few days. However, Richard’s story is more than a simple crash, since according to his account it involved a system malfunction that only affected two people, and was not detected by the operating system\(^\text{20}\). Such software faults are rare in modern systems, but in the mid sixties Univac’s operating systems were largely handwritten in assembler\(^\text{21}\). Such code is hard to make error free, and many sections of code would seldom be run. The errors in these pieces of code would therefore only appear at irregular intervals, when that code was run. It is likely that Richards and the other user were unfortunate enough to run just such a piece of code, which resulted in the anomaly.

It is always hard to know if individual anecdotal accounts have a genuinely anomalous part. The researcher has to be cautious when examining such anecdotal evidence, since there could be some personal bias involved. This is especially true of the accounts in the popular press. These are often full of supposed ‘supernatural’ events with machines. Many of these may well contain a large element of sensationalist journalism, or attempts to benefit the subject of the story. An example of these sorts of anecdote are shown by the story of a personal computer system owned by a United States based spiritualist movement that claims the telepathic channel can take dictation directly into the wordprocessor (Computing, 1987). I think most people would agree that this is taking Koestler’s ‘Ghost in the Machine’ (Koestler, 1976) too far!

In contrast, the kinds of anecdotes that I have paid attention to are those that have caused some kind of unwelcome affect. An example of such unwelcome effects are the mechanical problems which plagued a home in Minsterworth, Glos. over a five year period. Alternatively there was the case of the frequent wordprocessor breakdowns reported by a professional writer as her deadlines approached, and which disappeared when an engineer was called to examine them (Personal communication Ashcroft-Nowicki, 1987). In the former example,

---

19 A 'crash' is a computing term used to denote any form of system breakdown.
20 Such an isolation of an error to just two people is a byproduct of the use of what are called 'user partitions'. These isolate each user, or group of users, to their own 'virtual machine'.
21 Personnel communication with one time 1108 system manager.
the effects included burst pipes, exploding lightbulbs, dying pets, and sounds in the night. They turned out to be due to the location of the house, which coincided with the cross over point of radar pulses from four Ministry of Defense establishments (Guardian, 1987). A similar, but as yet unsolved case was the series of breakdowns which befell the London camera crew of the Channel Four’s ‘Network Seven’ when they tried to arrange a televised exorcism of a house. The occupant, and independent visitors had reported experiencing chills, mists, and having seen a full form figure. On the journey down from London to this house the camera crew had to change cars Four times due to unexpected electrical or mechanical breakdowns (Portsmouth Evening News, 1988).

Even if these kinds of anecdotes, and the universality of the belief in such machine anomalies is ignored there is considerable experimental evidence from independent parapsychological researchers which points to such possibilities. It may not be quite so preposterous to look more carefully at Steve Richards’s suggestion of the interference of sensitive modern electronic equipment by ‘mental action’. The possibility of such effects have been acknowledged by several parapsychologists (Puthoff & Targ, 1974; Nelson et al., 1981; Stanford, 1981; Cook, 1982; May et al., 1985; Morris, 1984 & 1985 & 1986; Dahlen, 1985; Jahn & Dunne, 1986 & 1987; Morgan, 1987).

The central figure within parapsychology who has driven the conceptual forefront of these speculations has been Robert Morris. He first laid out his speculations in a paper in 1983, and in more detail in the following year (Morris, 1984). Morris postulated that the increased sensitivity of modern technology provided more opportunity for the possible influence of machine functioning by as yet unknown means. There was a large body of evidence within parapsychology that suggested that sensitive labile systems could be influenced by the volitional action of observers. Note that as Broughton (1979) has speculated it is difficult to determine if it is the subject or experimenter (or both) who are responsible for such effects. This research involved the use of Random Event Generators (REGs). This is a loose term used to cover any device or process which produces individually unpredictable outputs, but whose overall behaviour should be statistically predefined. This allowed experimenters to test subjects for general extrasensory perception, (GESP) and later Psychokinesis (PK). Early examples of these REGs were purely mechanical in nature (Tyrrell, 1936; Kendall & Babington-Smith, 1939; Redmayne, 1940; Parsons, 1946), and usually involved some kind
of revolving drum and interspersed electrical contacts as the basis for their randomness. In many cases this made them subject to variations in their randomness, due to temporal behavioral characteristics of their subjects. These inherent weaknesses are graphically illustrated in an excellent review of the then current REGs by Wilson (1947).

These electromechanical methods of generating randomness were a logical development from the early card shuffling and dice rolling methods, since they were less subject to known methods of human intervention. Today the hypothetical factors of both ESP and PK are viewed as being equally pertinent to any anomaly, but from a historical perspective PK has been assumed to play the major role in any apparent machine influence. Indeed one of the earliest commentaries on the apparent psychic nature in such machine anomalies was entitled 'PK-Missing People' (Greenhouse, 1975). It is for this reason that any introductory overview of the origins of Morris's speculations must concentrate on PK. The apparent phenomena of Psychokinesis has been studied by parapsychologists in various ways since Rhine (1943) conducted his first dice rolling experiments at Duke University. We have already mentioned the electromechanical REG devices that were in use up until the late 1950s (for a review see chapter 2). Most of these studies assumed that they were investigating the phenomena of ESP. Investigations into PK concentrated mostly on mechanical dice throwing experiments (Rhine, 1943), the landing placements of blind die (Nash, 1946; Forwald, 1953), or the influence of some highly complex mechanical devices constructed by Cox in the late 1960s and mid 1970s. Much of this early work has been criticized by parapsychologists and skeptics alike. This is because of problems with recording (West, 1962; Girden, 1962), or in Forwald's case that he was both experimenter and subject (Gardner, 1985). Since the 1960's the study of PK has become split into two main areas of research. First there are investigations into gross material influences, such as metal bending, or table levitation, called Macro-PK (Batcheldor, 1982; Isaacs, 1985). Second there is the study of statistically detectable effects. These are usually upon random processes or more often random number generators (RNGs). These small scale effects are called Micro-PK\(^{22}\). The latter form of investigation has become much more popular in recent times (for

\(^{22}\) The use of the term 'PK' in this context has been called into question almost since the term was first adopted (Rhine & Pratt, 1957). Lately this has been questioned by the intuitive data sorting (IDS) model (May, 1986; May et al., 1986).
examples see Beloff & Evans, 1961; Schmidt, 1969; Jahn, Nelson & Dunne, 1985). This increase in popularity is due to such factors, as the ease with which computerised controls can be maintained (Hansel, 1967; Millar, 1977; Broughton, 1982), and the theorized properties that such supposedly quantum mechanically indeterminate events have until they are observed (Klip, 1967; Schmidt, 1975; Walker, 1973; Millar, 1986; Von Lucadou, 1987).

1.1.3.1. Parapsychology’s Use Of Quantum Mechanics.

Possibly the most promising theories which have emerged in parapsychology in the past 20 years are the so called 'Observational Theories'. These theoretical explanations are felt by many researchers to be providing a consistent framework within which psi phenomena might be accepted within mainstream science. The author of this dissertation is not competent to comment in any more than a passing way upon such a specialised field. Instead the next brief section will try to provide a background to parapsychology’s use of observational theories. However it should be taken with some degree of caution since many of the assumptions have been grossly simplified to allow explanation without the need for complex equations. Interested readers should refer to a reference text on modern physics.

Within mainstream physics the prevailing interpretation of quantum physics is called the 'Copenhagen Interpretation'. This was agreed upon as a standard to reduce possible problems with conflicting alternative models. The term 'quantum' is due to the discrete or measurable steps, found in the subatomic events that make up matter and energy. All types of matter have been found to be made up of the same constituent parts, namely nuclei and electrons. The measured activity, or 'quanta' found in such systems have been found to be in multiples of the smallest charge, 'e', the charge of a single electron. It was also found that all types of matter have what is called a 'wave-particle duality'. This means that they can be totally described as a wave or as the effects of multiple particles, but not in both ways

---

23 The quantum indeterminacy of such unobserved events has been called into question (May & Spottiswoode, 1988). Furthermore many mainstream physicists have many reservations about these parapsychological interpretations of observational theory (Gardner, 1985).

24 It has been said that only a dozen people in the world fully understand quantum mechanics and their associated theoretical constructs (Feynman, quoted by Aharonov & Vardi, 1980 p129).
Simultaneously. Such sub-atomic particle systems are in continuous motion, and these motions, or outcomes (given enough degrees of freedom), are equally probable. With limited degrees of freedom there are varying statistical probabilities associated with each possible final observed state. The actual outcome, or state is unknown until it occurs. Until that moment the possible outcomes exist only as statistical likelihoods. This state where the outcome of a such an event is as yet unknown is called 'quantum indeterminacy'. The complete list of probabilities associated with this indeterminacy is called a 'state vector'. These concepts form the basis of accepted modern physics, which has replaced so called 'Classic' physics. This is because quantum theory accurately describes observed sub-atomic activity\(^25\), and when the predictions made by quantum mechanics are applied to macroscopic (larger than sub-atomic systems\(^26\)) they agree with the predictions of classic Newtonian physics.

Some modern theoretical physicists have postulated extensions to the 'Copenhagen Interpretation'. Such people as Eugene Wigner & Bernard d'Espagnat have postulated that quantum indeterminacy remains in effect until the moment of human observation. They propose this extension because they postulate that the measuring devices themselves can be described in quantum mechanical terms, and are therefore in one sense undefined. This in turn makes the whole system of sub-atomic event and measurement only become defined at the moment of human observation (d'Espagnat, 1979). Such an interpretation is highly controversial within modern physics, and has certainly being challenged by the recent findings of May and Spotteswoode (May & Spotteswoode, 1988).

---

\(^25\) Recent experimental work has shown that Chaos Theory may be better at explaining subatomic behaviour than Quantum Theory (Gleick, 1984). Such Chaotic effects may have important implications for parapsychology. If the entire universe could be shown to be a chaotic system, progressively derived from a single function (on a single mass), non-local correlations might be explained as by products of that single function.

\(^26\) This point is usually assumed to be when the specified system is larger than the 'de Broglie wavelength'. In practice this the boundary of the subatomic system.
Certain parapsychological interpretations of these quantum mechanical theories take d'Espagnat's theory one step further. They postulate that consciousness influences quantum indeterminacy by the act of observation. Some theoreticians within parapsychology propose a reactive process that occurs during perception. This process occurs by special 'hidden variables'27, that can be postulated to exist between the random neuronal discharges of the brain and the state vector of the target system. It is proposed that both random neuron discharge, and target system share 'quantum indeterminacy' until consciousness has collapsed both sets of state vectors into reality. Such an interpretation is not generally accepted by modern physics. One of the main theoreticians (Walker, 1973) has proposed being able to determine the 'PK hit rate probability' by using Gridgeman's relation, and specific bit rates for separate systems in the human mind. Walker states the bit rates29 for the 'Subconscious mind' (10^{12} bits/sec), 'Conscious mind' (10^8 bits/sec), and 'Attention or Will' (10^4 bits/sec). He proposes that 'Will' is the channel for telepathy (since he feels it is independent of time or space), and 'Consciousness' combined with the 'Will' use the 'hidden variables' to determine state vector collapse (Walker, 1973). In a development of these ideas Mattuck (1977) proposed that this PK influence could be shown to 'put information into matter' in one millisecond pulses. He felt that the 'matter variables' changed at 10^5 times the noise level. Mattuck felt he showed this using Langevin and Fokker-Plank equations containing the extra PK term equal to the mind output (10^8 bits/sec). Unfortunately the author does not know enough about quantum physics to be able to comment on the validity of these extensions to quantum mechanics.

1.1.3.2. The Use Of REGs In PK Research.

Many micro-pk studies have been conducted using a variety of randomness sources, including radioactive decay, electronic 'noise' diodes, and pseudorandom algorithms. Significant results have been reported by experimenters using REGs of all three types. In 1983 at a symposium on the applications of anomalous phenomena Morris introduced the

---

27 Mainstream sub-atomic physics postulates the existence of 'hidden variables', but these are only 'local' and described as 'realistic'. Walker's (1973) 'hidden variables' are neither local or 'realistic', since they are not measurable, dimensionally infinite, independent of space and time, but must be agreed by separate independent multiple observers.

28 Truly random events are macroscopic events that are thought to have quantum mechanical indeterminacy.

29 It is uncertain where Walker obtained these bit rates.
idea of using the output of micro-pk influenced REG to perform a useful task (Morris, 1983 p148). He later called this use 'Smart Noise' and introduced the twin concept of 'Dumb Noise'. This would be when an influenced noise based system performed worse than base line chance (Morris, 1985 p7). Morris also introduced the 'Function Linked Person' (FLP), who would be someone who (consciously or unconsciously) directed their psi to enhance the functioning of computer systems with which they were involved. Morris also proposed the hypothesis of a 'Malfunction Linked Person' (MLP) who would do the reverse. To compliment these two hypothetical types Morris also proposed that working environments could be conducive to either function or malfunction linkage. He called these Function Linked Environments (FLEs), and Malfunction Linked Environments (MLEs), respectively. Morris would be the first to admit that these concepts can exist independently of their parapsychological connotations, and include (or perhaps can be eventually confined to) the basic factors traditionally assumed to be involved in human equipment interactions. These include psychomotor action, sensory organizing and cognitive abilities, physical and mental health, work habits, intimacy with the task, equipment, observer mis-intepretations, and the persons attitudes to the human-equipment interaction system they are using (Morris, 1985, p7).

Traditionally academic parapsychology has assumed paranormal interactions could be classified into one of two main classes of phenomena. The first describes an unknown influence made by an organism on its environment (called Psychokinesis by western parapsychologists). The second is an information transfer from the environment to an organism by unknown means (called Extrasensory Perception by western parapsychologists). These classifications may be valid in parapsychology, but when we look at the causal action that determines inexplicable events on a computer system, these labels assume more knowledge than we have. Instead the adoption is proposed of the notion of Function Alteration Activity (FAA). This term is preferred since it only says that some activity has taken place within the computer system which has affected the performance or actions of the computer system. In a later section the full theoretical background from which this term was derived will be elaborated. The term is introduced at this point since it will be used to describe any activity that alters the performance of a complex system. It should also be noted that this definition does not necessarily imply that such an effect is paranormal in nature.
1.1.3.3. Human Computer Interaction.

Obviously any experimentation into this area must investigate all the normal physical, psychological, and subjective factors involved with human equipment interaction. This area of inquiry already forms a domain of accepted interdisciplinary investigation between Psychology, Ergonomics and Computer Science, and is called Human Computer Interaction (HCI). For these reasons, and because the author is a computer scientist with an applied interest in HCI, a large part of the study asked empirical HCI questions. The areas of inquiry addressed by the HCI part of the study can be broadly split into two sections: those which looked directly at computer interface preference and performance issues, and a secondary investigation into the attitudes and backgrounds of computer users. The background and details involved with these factors as they were applied in this dissertation are discussed in chapter 4.

1.1.3.4. Computer Security.

The other concern of the author was the lack of clear definitive guides to the security risks on computer systems. Security is being used in its broadest sense to cover both intentional fraud, and unintentional artifact. These problems are particularly acute within parapsychology, where there is a rapid increase in the use of computers, often by experimenters with little knowledge of computer systems' inherent weaknesses. The dissertation examined and detailed the major problems areas, and developed practical and effective preventive measures. These are detailed in chapter 3.

1.1.4. Summary Of The Contents Of This Dissertation.


The security aspects used the details from the overview of all parapsychological use of computers since 1936 (chapter 2) to provide a comprehensive review of current security problems with parapsychology's use of computer systems. This security chapter outlines most known methods of influencing electric components.

30 Some methods are not covered in detail for reasons of national security.
1.1.4.2. HCI.

The HCI part of this dissertation explored several subject areas, three major aspects were:

First, which of two currently alternative computer interfaces, command line and direct manipulation, is superior in terms of work throughput, error rates and subjective preference, over three major user populations.

Second, are there individuals who have habitually good or poor experiences with high technology systems, and can any individual difference be found which correlates with this factor.

Third, do individual characteristics in background and personality influence a user’s preference for an interaction method.

1.1.4.3. Parapsychology (Anomalous Human Computer Interaction (AHCI)).

The parapsychological part investigated the following questions: Is there a paranormal element in human computer interaction. Can this hypothetical FAA be made to perform intelligent tasks, and do different types of REG have different susceptibilities to these hypothetical influences. Do individuals have any habitual pattern of influence, such that the FLP/MLP hypothesis can be confirmed. Do individual differences, either in behaviour on the computer system, or from a personality and background questionnaire, allow any consistent factors to be isolated that can allow the prediction of a persons success in machine interactions. Do different users have any individual differences in the way and means by which they influence REGs. Can the periods when these influences show most strongly be matched to some aspect of the users actions on the computer such as error rates, task complexity, or speed of entering commands. Alternatively do they match some other external environmental factor such as electromagnetic activity, gravitational effects, barometric pressure, time, experimenter’s mood, or weather conditions?

Murphy’s Laws: The Yulish additions.

1) Persons disagreeing with your facts are always emotional and employ faulty reasoning.

2) Enough research will tend to confirm your conclusions

3) The more complex the idea or technology, the more simple minded is the opposition.

From 'Murphy’s Fundamental Laws', Charles Yulish Assoc Inc., (1975).

2.1. Introduction 1.

Although there have been attempts to combine all parapsychological PK studies that have used a random number generators (Radin, May & Thomson, 1985; May, 1986; Nelson & Radin, 1989) there has never been a general review of the use of electronic machines designed to produce random sequences, or to control experiments in parapsychology. This review does not claim to be exhaustive (that would be impracticable), but will serve instead to indicate the general trends involved in the use of machines in parapsychological research. The review will not go back before the mid 1930s, since the digital computer was not truly born until Alan Turing dreamt of his 'Universal Machine' while lying in a Cambridge meadow in the early summer of 1935 (Hodges, 1983 p96).

The advantages of using a computer in parapsychological experiments have been stated frequently (Wilson, 1947; Girden, 1962; Hansel, 1966 & 1980; Millar, 1977; Broughton, 1982; Rush, 1986), and include reduction in recording errors, increased security against

---

1 The content of this chapter and that of chapter three form a sub-set of a section in the forthcoming 'Advances in Electronics and Electron Physics 1990 edition' (Academic Press). This will be entitled 'Electronic Tools in Parapsychology', and will be co-authored by Robert Morris.
fraud, and reduced experimenter involvement. In moments of unguarded enthusiasm parapsychologists have been known to go so far as to say that 'These machines eliminate target bias, recording errors, and opportunities for fraud by a subject' (Rush, 1986 p254). This could be interpreted as being a little optimistic, since there are unfortunately still numerous ways in which computer systems can introduce opportunities for all of these effects to occur, some of which will be shown in subsequent chapters.

The False Dichotomy.

Most parapsychological experiments which use some form of automated experimental design assume that subjects are either influencing the action of some part(s) of the system, or are in some way sensing the forthcoming action(s) of the system. These two assumptions form the two main areas covered in this review of automated experimental design, ESP and PK. This review will cover them separately for ease of presentation, however in reality the distinction between them is vague. A subject who can determine a future state of the machine, may well be influencing the system to produce the predicted state. This problem becomes even worse when we allow for retrocausation, since in this case the subject has almost limitless scope in interacting with the states of the system.

2.2. The RNG Devices.

Having seen there is some evidence for subjects being able to influence automated systems in what are apparently paranormal ways we can now review the basic methods used to generate uncertainty or randomness in an automated system. For the sake of simplicity these methods can be divided into five main methods, Natural, Mechanical, Electrical, Atomic, and Logical or Mathematical. Natural sources of randomness are rare, but have the advantage of being easily verified by multiple independent observers. The most famous of these suggested natural sources is the combined temperature readings from the worlds major capital cites for a specific day (Morris, 1986). Mechanical systems are not so frequently used in experimental parapsychology since the advent of the computer. These randomization methods were usually based upon the final resting position of some free moving part(s), or system, such as the fall of true dice. These had the disadvantage of being highly susceptible to the short term process interrupting behaviour of subjects. These mechanical systems also tended to be noisy (West, 1962 p160), and this gave the added disadvantage that the subject
could use or develop an ability to interpret subtle sound from the device. The early electrical methods of generating randomness used the same principles as the mechanical systems, except of course they were using electricity as the motive force for the free moving randomizer. However these electromechanical devices have been largely replaced by the so called 'zener' or noise diode. These systems take advantage of the random motion of a large number of electrons ($\sim 10^{23}$ according to Furth and Macdonald (1947)) in a specially designed circuit. This circuit is designed so the signal produced when it is read during a trial will be in one of two possible electrical states. If the circuit is properly designed, constructed, and operated the outcome of any one read operation (trial) should be completely random.

The final truly random source we will cover is that based upon the random disintegration of an unstable atomic nucleus, or radioactive source. This disintegration results in the liberation of sub-atomic particles which can be detected by the appropriately calibrated measuring equipment. This measuring allows the system to derive a random signal from the process.

Finally we come to the category of pseudo random sources (and preset sequences). For the sake of simplicity this section also includes the various tables of random numbers with computer generated pseudo random number generators (PRNGs). Mathematical purists regard the former as being truly random sources since they are usually the recorded results from a truly random source. However many have had their original sequences adjusted by their creators to remove the more 'wild' non-randomness which occurs in truly random sequences (as noted by Kendall and Babington-Smith, 1939; Spencer Brown, 1953 & 1957; Oram, 1954; Nicol, 1955; Harvie, 1973; Sargent, 1980; Palmer & Weiner, 1985)). For this reason the author feels it is safer to regard them either as a sub-category of PRNG (as Schmeidler and Borchardt, (1981) did), or as being in a class of their own. The author prefers the former categorization method (regardless of Radin's (1985) objections which were based upon the fact that these sequences were produced by true random sources.). The PRNG is based upon some deterministic logical operation which is normally performed by a computer. Various mathematical methods have been developed to generate such sequences, the more popular of which include the Feedback Shift Register Technique (Lewis & Payne, 1973; Knuth, 1981), The Linear Congruential Method (Knuth, 1981), and The Power Method (Hewlett Packard, 1975, p116; Tart, 1983; Radin, 1985). All these methods share the property of being totally deterministic, and will produce the same sequence of numbers given identical starting conditions. Pseudo random sequences are usually more tightly bound
(less labile) than those produced by truly random processes. For this reason parapsychologists have tended to feel PRNGs would be less likely to influenced by 'psychic' means. Radin (1985) outlines five limitations with PRNGs. These were that linear congruential methods produce high order regularities, limited precision arithmetic may adversely affect what would otherwise be a good PRNG (due to such things as rounding errors), many linear congruential methods produce regularities in their inter-word bit patterns (while feedback-shift register PRNGs produce random inter-word bit patterns but non-random whole word results^2), that PRNGs show less lability than true random sources, and finally that generic PRNGs supplied with systems are often extremely poor. Many of these problems, and others not covered by Radin are discussed in chapter 3.

To illustrate the diversity of electronic RNG systems developed to investigate different aspects of parapsychology the following section presents a summary of the major types of RNG. Each of the four main types of random number generation used in automated experimental design have the same sub-categorization. This numbering reflects the increased sophistication of each device, so the lowest numbers are least sophisticated, and vice versa. The basis of the numbering system, assuming 'X' represents the major category identifier, is: (Xa) single media feedback, then (Xb) Multi-media feedback systems, where Xb1 has EEG recording & 'Gating' capabilities, Xb2 has variable generation speeds. Finally Xb3 represents those systems which are integrated into a controlling computer system.


An example would be the PEAR researchers use of the Elgenco Model 3602A-15124 noise diode generator in their PK-RNG work.

The use of this kind of system has the advantages and disadvantages inherent in using a standard product as an experimental target. There should be an advantage in demonstrating that showing paranormal effects can occur on commercial products, since this evidence may be more convincing to critical evaluations. However this may not be the case since the

average commercial use of these products does not require the same stringent standards of integrity essential in responsible parapsychological research. It has also been pointed out (Morgan, 1988), that the typical use for commercial noise generators are either in the enciphering of computer transmitted information, or in the simulated testing of frequency sensitive systems. The size and reliability of the effects shown by parapsychologists have been such that they are unlikely to be of any interest to the users of these commercial systems. The disadvantages in using such a system is that its full operational specification is available to any interested party. This means that any inherent weaknesses in the secure operation of the device will be freely available. The experimenter will also have to design the experiment to fit the specification of the commercial product, and this will probably limit the range of possible experimental hypotheses. Commercially available sources exist for sub-categories 1a, 1b2, and 1b3 respectively.

2.2.2. Type 2. Specially designed Noise Generation Systems.

These are in many ways preferable to commercially available products since the experimenter can have the product built to fit some exact specification. The major dangers with such systems are those present from the experimenter, poor design, construction, or misuse the system due to lack of experience in the fields associated with the proper use of RNGs. The fact that these systems can be specifically designed to have specialised functions means that they have spawned a multitude of sub-categories of RNG. These sub-categories will be summarized:

2.2.3. Type 2a. Electronic Noise Diode Systems With Single Media Feedback.

Example:

The Edinburgh Electronic ESP Tester or EET

John Beloff and Timothy Regan (1969) described the so called Edinburgh Electronic ESP Tester or EET, a purpose built electronic randomizer and tester built by the engineers at the Psychology Department of the University of Edinburgh. Beloff and Regan took considerable time and effort to eliminate any possibility of fraud, and made extensive reliability checks with the apparatus.
Although these are the most basic RNG systems, the use of a computer can enhance the feedback provided by such systems to an almost limitless extent.

2.2.4. Types 2b Electronic Noise Diode Systems With Multi-Media Feedback.

2.2.4.1. Type 2b1. Electronic Noise Diode Systems With EEG And RNG Recording In The Same Apparatus.

Example:

PsiFi (EEG and RNG Recording In The Same Apparatus.)

May published a detailed description of the device he called PsiFi (May, 1976b). This was a piece of equipment that provided EEG and RNG recording in the same apparatus. It had the capability to give feedback in the form of music, white noise, or a polygraph needle. Its random number generator was based upon a noise diode, and produced an output in the form of a binary decision. It had the added sophistication of allowing either the experimenter or the machine (via the RNG) to decide the target condition for a series of trials. One of the major features of its design was that it allowed Ss EEG activity to selectively record RNG output during periods of specific brain states (called 'gating'). May also published a report with Honorton about Ingo Swann’s attempts to influence the PsiFi equipment (May & Honorton, 1976). Unfortunately Swann’s attempts were not reported as being significant, but May & Honorton felt there was evidence that Swann was more successful the farther away he was from the device.

2.2.4.2. Type 2b2. Electronic Noise Diode Systems With Variable RNG Generation Options.

Example:

Modular Communications Testing System.

Placer, Morris, & Phillips (1977) described a RNG-PK testing machine called the Modular Communications Testing system or MCT. The MCT was an advanced electronic system which was based upon an electric zener diode, with a two transistor amplifier with the output
converted via an aLM339 comparator. The resulting logic signal was further divided by two to ensure that an equal time was spent in a high and low states. The random logic signal oscillated at 200Khz, and was sampled and then clocked. The system had electronic counter totals for the number of trials, and hits; these were displayed via LEDs. The authors describe the MCT as having five modes of operation: GESP, Clairvoyance, Precognition, PK1 (a new random number is generated for each button press), and PK2 (in which targets are generated in a continuous stream), at rates ranging from one target every two seconds, down to 50 per second. The MCT was used in a lot of Robert Morris’s work in California.

2.2.4.3. Type 2b3. Electronic Noise Diode Systems Integrated Into A Controlling Computer System.

Example:

Psychophysical Research Laboratories 'PsiLab' with Computer Administered Psychometric Measures.

Rick Berger and Charles Honorton (1984) reported on the construction of the Psychophysical Research Laboratories (PRL) computer system called 'PsiLab'. This was an adapted Apple II, which included an electric noise diode RNG\(^3\) on an internal board, and special software packages to make use of the board’s facilities. The RNG used two independent noise sources (2 low voltage noise diodes, and two independent random bit generators) to achieve first order randomness. The system used the digital quiet time for thresholding. The noise diode outputs were digitised by using a reference voltage. The two resulting streams were combined in a digital half-adder circuit and fed into a shift register at a rate determined by the Apple II’s clock (250 KhZ). The entire system used internal timing circuits to ensure that all the logic levels were stable during essential operations. A read of the RNG results in a shift of the shift register to the processor, with a rate of 10KHz (1250 bytes/sec).

---

\(^3\) This was a modified version of the so called Bierman (RIPP) board, with modifications by Robert Chevako of Syracuse University Communication Studies Laboratory. These modifications removed some of the potential sources of induced artifact that could occur in the standard Bierman RNG board. These modifications include the removal and checking of oscillations and excesses of 1/f noise, particularly structured induced electrical noise. This 1/f noise was called in Chevako’s own language 'Popcorn noise'. 1/f noise can produce baseline shifts, or pulsing.
software supplied with the system included two PK experimental games, and two ESP games. The PK games were called 'Volition', which used fast assembler routines to access the RNGs, and displayed a moving line down the screen. The results were simultaneously written to floppy disc. In contrast 'Psi Invaders' was a modified version of the arcade game, but the RNGs determined for a given trial if the fire gun worked or not. In Psi Invaders the RNG was sampled 100 times and the MSD of these were compared with a randomly assigned target bit to determine if the laser gun worked, or misfired. The ESP games included 'Circles' where the subject moved a joystick to indicate which circle they thought was the target one. Finally the ESP game 'Dowser' looked at how close the joystick was placed to an undisclosed target. PsiLab also included a battery of computer administered psychometric measures (PIF, & Myers Briggs Personality Inventory).

2.2.5. Type 3 Radioactive Decay Noise Generation Systems.

Example:

Schmidt’s (Schmidt, 1969) Use Of The Decay Of Radioactive Strontium-90 Nuclei As A RNG-PK/Precognition Target.

The evolution of the use of these atomic sources is similar to that made by the electronic RNG sources (above), and leads to the same sub-categorization system. So 3a, would be a system with single media feedback, 3b would include multi-media feedback (3b1 having EEG recording capabilities, and 3b2 having variable generation speeds). Finally 3b3 would include those systems which had been integrated into a controlling computer system.

2.2.6. Type 4. Pseudo Random Number Generators.

Excluding random number tables all these sources are implemented on a digital computer, so fall within the category 4b3.

2.2.7. Type 5 & 6. Mechanical and Natural Systems.

These have been combined to save space since their current role in experimental parapsychology is purely historical. Both types could have the full sub-category ranges described for the previous sources.
2.3. Experiments Designed To Demonstrate An Effect Which Appeared To Show Subjects Guessing The Outcome Of A Random Event Generator (REG).

2.3.1. Guessing Mechanically Generated Targets

The Electronic Rotary Switch Methods of REG Generation.

The first use of electro-mechanical device within parapsychological research which falls within this review is Tyrrell’s (1936) use of an electronic pointing apparatus with an automatic recorder, electric lamp, and a mechanical selector, or randomizer source. When most experimental researchers were still using hand shuffling of targets (or random number tables), Tyrrell’s battery driven device was a major technological achievement. Tyrrell’s randomizer was a rotary switch with 25 contacts of five groups of 5, one of which the rotating switch made contact with to produce the random selection. These mechanical methods of deriving random numbers were the best that were then available, and were used as the basis for creating tables of random numbers, such as the table of 100,000 random numbers produced by Kendall and Babington-Smith (1939). These were widely used in parapsychology and the social sciences until the advent of the digital computer. The numbers used in random tables have often been ‘adjusted’ to remove any extreme biases which would be present in ‘real’ randomness. However such biases have still been found in sequences from such ‘adjusted’ random number tables (Spencer Brown, 1953 & 1957; Oram, 1954; Nicol, 1955; Harvie, 1973; Sargent, 1980; Palmer & Weiner, 1985).

A similar device to that created by Tyrrell (1936) was produced by Redmayne (1940), and used a battery powered motor driven rotary selector, with vanes mounted at equal distances around a cylinder. The motor was interrupted by the subject, and the position (vane) where the selector stopped determined the output of the system. The user of this machine and Tyrrell’s earlier machine needed to ensure that the battery was constantly checked before each use, since a dying battery would have biased the randomization process. Although the rotating switch method of generating randomness was the usual method adopted at this time, it should not be imagined that it was the only method available. Six years later another device
was created by Parsons (1946), but this used a different method of random event generation which was the movement of counters when they were shaken.

2.3.2. Early Use Of Punched Cards.

Kahn’s Use Of An IBM International Test Scoring Machine.

The major use of computers in these early years was in the role of automated checkers, and collators of data. The most notable user was David Kahn (Kahn & Neisser, 1949; Kahn, 1952). In these experiments Kahn used an IBM 'International test scoring machine'. This apparatus read forms which contained 300 rows of five data items (150 x 5 per side). Kahn generated targets using random number tables, and placed one target in each of the 300 rows. In one study Kahn used 100 Ss at differing distances from the target (0.5 - 500 miles), each of whom checked their responses on blank forms. In total the subjects generated 43,278 guesses, which were automatically checked by the IBM machine. Kahn also manually checked the responses, to remove any artifact that might have been present if the Ss had entered multiple guesses, or omitted to respond for any target. Kahn reported finding 3.14 percent more hits that would be expected by chance (p < 0.01). The distance effects revealed that the more distant subject scored better than those closer to the target (p < 0.01). However a predicted sheep-goat difference did not emerge. The major problems in such form based guessing was that subject’s might have a natural bias to one of the entry points on the form. This would cause many subjects to repeatedly chose the same target, called a stacking effect. To check against this Pratt reanalysed Kahn’s data using Grenville’s adjustment for this so called 'stacking effect'. This analysis revealed that Kahn’s results could not be explained by such an artifact.

Rhine’s Experiment With The Canadian Broadcasting Corporation And McLean’s Magazine.

Rhine (1961), took advantage of the computer’s large automated data handling capabilities when he collaborated with the Canadian Broadcasting Corporation and McLean’s magazine to conduct a large scale precognition test involving 30,000 people. The respondents completed their guesses on cards supplied by McLean’s magazine, which were attempts to match targets (playing cards) generated by a computer’s pseudo random number generator.
These targets were only generated by the computer after all the guesses had been submitted. At the time this was thought to be a test of precognition, however more recent theoretical developments in parapsychology (Schmidt, 1975; May, 1986; May et al., 1986) would have allowed PK or IDS\(^4\) to play a role in determining the outcome in this experiment. Rhine reported finding significant results, but these findings have been called into question due to the presence of a statistical artifact in the range of possible targets (playing cards) (Diaconis, 1978).

**Mihalasky’s 100 Digit Number Hand Punch Test.**

Another experimental protocol which took advantage of the computer automated data collection and manipulation was the system devised by Mihalasky (1966), to be an experiment in precognition. It took advantage of the fact that large numbers of potential Ss were available at public meetings, such as the meetings of the ASPR (a typical experiment was done at the ASPR meeting of November 20th 1965), and used the audience as subjects. The Ss had to key a 100 digit number (range 0 to 9) onto an IBM punched card, using the then standard hand punch machines\(^5\). Ss were allocated to either a fast or slow guess entry speed, to see if speed of guessing had any effect on the results. Once all the cards had been completed an IBM 1620 computer using a pseudo random number generator\(^6\) was used to generate an

---

\(^4\) Although May’s (1986) version of the IDS model assumes transtemporal effects (precognition), it is also possible to create real time alternatives which act upon existing probabilities. The author of this dissertation feels these real time explanations should be investigated and falsified before the investigation of the more complex transtemporal explanations.

\(^5\) The hand punch was a semi-portable method of keying in data to a punched card. It looked like a one-handed Braille machine, and had shift buttons to enable the few keys to double both as alphabetic and numeric buttons. They were not noted for their ease of use, since the key combinations required some time to learn. It is therefore likely that the Ss would be distracted to some extent by the very process of entering their 100 digit guesses.

\(^6\) As with most early uses of the computer to generate pseudo random numbers the experimenters do not mention the algorithm used by the computer to generate the random numbers. This makes it very hard to evaluate the significance of their findings. In this case however IBM UK Ltd have been kind enough to supply the source code of what would have been the then standard IBM pseudo random number generator in technical document c20-8011. Examination reveals that it was a power residue generator, which would have had a repeat sequence of \(2^{(b-2)}\), where \(b\) is the word length. Although this would give a reasonable repeat length (1.8\(^1\) with a 64 bit word), the random sequences produced by such an algorithm are not of a very high quality, and could have produced spurious results. These might take the form of biases, repeating sequences, or sub-patterns. The problems involved in random number generation will be discussed in a subsequent chapter.
identical number of target digits. Mihalasky reported finding that the fast method of data entry worked best for men (with the fast method being above chance, and the slow being below), but this trend was reversed for women. (Since this was reported in a pre-1970s PA proceedings there are unfortunately no data available as to the significance levels of these effects, or if they were predicted or post-hoc).

**Pratt’s Use Of PRNGs In Computer Punched Card Guessing Experiments.**

A similar experimental technique was described by Pratt (1967), where he used the University of Virginia’s Burroughs 5500 computer to generate pseudo random numbers using an algorithm written by the computing department of that University in the language Algol. Ss entered their guesses on punched cards in the same way as they did in Mihalasky’s (1966) study. The only difference in methodology was that Pratt generated his targets (the digits 1-5) using the computer before the Ss made their guesses. Pratt’s methodology is also of note since his system performed various forms of analysis upon the data, such as using Greville’s method of evaluation for random targets in an open deck, checking for displacement (+1) & (-1), and looking for a session decline.

**2.3.3. Electronic ESP Testers.**

**Schmeidler’s Early Work With Electronic Machines.**

Gertrude Schmeidler (1964) reported the results of a series of experiments that she had conducted in 1961 and 1962. Using a computer to generate and score 150 targets Schmeidler investigated the role of precognition and feedback with those targets. Fifty were seen by the 75 Ss after they had made their guess, 50 were only seen by the experimenter, and 50 were never seen (except in the overall results). Such an experimental design would not have been possible without a computer. Schmeidler found significant scoring, which she interpreted as giving strong evidence for precognition, and she also noted that psi missing tended
to occur if 'interesting information' (feedback) was withheld from the Ss. What is so interesting about this study is that it was made before any of the observational models had been made, and therefore the role of feedback, and computer based experimental control was still very rare indeed. In this case the experimental design had tried to investigate the roles of precognition and true telepathy. The fact that Schmeidler noted that feedback had an effect on the results has been interpreted as giving extra weight to the observational theories.

**Parker’s Use Of The Edinburgh Electronic ESP Tester or EET**

Adrian Parker (1974) reported on a use of the EET (Beloff & Regan, 1969). Parker reported on two cases of extended ESP success from a sample of 25 Ss. He speculated as to why the screening process of using the EET should only find two Ss out of 25 that had any success with card-guessing, but sadly could provide no concrete answers.

**Millar’s Work With Electronic ESP Testing Machines.**

Millar (1977) discussed electronic ESP testers, and the problems that could be associated with them. He defined two types of randomness, Causal and Non-causal, and noted that when using mechanical methods of random number generation one cannot be certain how much of the randomness falls into each category. Millar’s two randomness categories would be more frequently described today as truly random and pseudo random. Millar repeated the concern he raised in a previous publication (Millar, 1973), that ESP testing devices might only be detecting outstanding abilities of some subject’s reaction times. Many of the concerns that Millar expresses as problems with mechanical randomness methods have become obsolete by the advances made in electronics design. An example of these would be his concern over a S’s too rapid button depression disconnecting the oscillator as the ring counter is poised between two successive states (See Wilson, 1947). Millar’s concern in this case was because the final indeterminacy would depend on small differences in between

---

7 Morris (personal communication, 1989) has pointed out that the issue of retroactive observation effects were well known before the arrival of the observational theories.

8 Note that throughout this review the term Observational Theory will refer to a parapsychological extension to what mainstream science recognizes by the term observational theory. As such these parapsychological extensions may not be recognized by physicists as being valid interpretations of quantum physics.
units. In electrical terms such events are called 'races'. When such events occur it is usually the fastest device, or shortest distance that 'wins'. However much of Millar’s comments are as valid to day as they were 12 years ago. For instance his warnings of possible fraud methods included: Controls being worked in the wrong sequence, switches held on permanently, more than one guess (pressing all options), switch held so it only partially made contact, the switch may be operated too fast, and finally the rng being turned off and on during a session. These kinds of issues will be discussed in chapter 3.

Tart’s ’Apple-Adept’ ESP Tester.

Charles Tart (1981) reported a study at the 1981 PA convention in which he had attempted to see if any difference emerged when Ss had to guess targets generated at the moment of guessing, and those which were pre-recorded. Tart used the Apple II 10 choice ESP trainer games called ’Apple-Adept’. The computer randomly chose targets from an electric noise diode source, or from a series of pre-recorded numbers. Both the E and Ss were blind to the source during the experiment. Tart reported finding no significance in this study, but in a subsequent series with +1 precognitive targets (instead of live), he reported finding a suggestive difference. This was in favor of the +1 precognitive targets (p < 0.06). However some doubt must be cast upon these findings, since Tart’s technician found that the hardware RNG was broken during the experiment.

2.3.4. Atomic Decay ESP Testing Devices.


Helmut Schmidt (1969) reported on the tests he had carried out on a random number generator he had designed and constructed, which used the decay of radioactive strontium-90 nuclei as the basis for its randomicity. The device randomly generated the integers one to four as its output; the extensive tests which Schmidt carried out over a 100 day period (total
of 4.7 million numbers) yielded nothing to indicate non-random behaviour. In this work Schmidt also reported the results of some experiments in which Ss were able to predict the number the device would generate next to a statistically significant degree. In an attempt to minimize the risk of recording errors Schmidt used paper tape recordings of all the trials and conducted his analysis by a computer. These studies with Schmidt at the end of the 1960s marked a major watershed for parapsychological research with computers. From this time on the computer would play a larger role in experimental control and analysis.

2.3.5. The Use Of Pseudo Random Sources In ESP Testing.

Schmidt (1969b) also reported using the RAND tables as the source of the randomness being observed by his Ss. In this study he fed in 300,000 digits from the rand tables onto paper tape and then used these numbers to generate 15,000 trials for six Ss. In these experiments one of four lamps was lit, and the Ss task was to guess which of the lamps would light next. Schmidt reported very significant results from this study (p < 0.01), and also recommended that future work use a computer to test the randomness of the systems being used, since it removed much of that hard work involved.

2.3.6. Environmental Factors On ESP Performance.

Andre (1972) conducted two ESP experiments which looked at possible effects on psychic performance made by the time of day, weather, and subjects mood. Andre reported that morning sessions seemed to get more significance, while afternoon sessions got a slightly negative affect. There was also some indication of there being higher scores on days that had higher humidity (p < 0.05). Surprisingly neither the first or second sessions showed the subjects mood having any affect, in contradiction to the findings of Carpenter (1982).

---

9 It is interesting that Schmidt initially assumed that some form of precognition was involved. This is almost certainly because, as we have seen, most previous studies with a machine based REG had assumed that the effects being found were due to precognition. In some ways the IDS model (May, 1986; May et al, 1986), have reversed views back to the way they were before the micro-pk work began.

10 These were a table of one million random numbers produced by the RAND corporation.
2.3.7. ESP Tests With Varying Probabilities.

Elizabeth and Russell Targ (1985) presented a paper which investigated if psi differentiates towards probable or actual future events. They used the associative remote viewing (ARV) protocol to describe which of six unknown objects would later be shown to them. There was a preset high probability (50%) of one of the objects, and low probability (10%) that they would see one of the others. They used the electrical random number generator present in the Texas Instruments (SR-51A) hand calculator. Targ and Targ conducted 12 trials and had seven high a probability hits and five low probability hits out of the 12 trials. This was interpreted as evidence that it was not more difficult to perceive low probability events by psi.

2.3.8. Varying Of Feedback With ESP.

Russell Targ and Charles Tart (1985) reported on a study in which they investigated if clairvoyance could be due to Ss precognizing the state of their own brain after they receive feedback. Targ & Tart used a computer to administer a test of clairvoyance by getting Ss to guess random numbers one to 10 generated by a RNG, with and without feedback. The totals were recorded by the computer to preclude anyone examining them in the future. Eleven Ss took part in the experiment, eight of whom had no previous record of success, and three of whom were known to be 'talented'. The eight naive Ss were reported to show no significant results, but the other three Ss were reported to show success in accordance with their beliefs, significant hitting by the believer in clairvoyance, and significant missing by the other two non-believers in clairvoyance. The overall results were non significant.

2.3.9. Learning ESP On Electronic Devices.

Duplessis’s Use Of The IBM Porto-Punch.

An experiment which attempted to use the old method of computer punched card entry to allow psi to be learnt was reported by Duplessis (1973). Subjects made calls on colours and

---

11 The authors could have been using PK to influence the RNG which made the decisions.
12 This is also open to the use of PK on the RNG.
dice face targets by punching the guesses onto IBM punched cards using an IBM PortoPunch device. Duplessis reported that colour and dice face hits increased when the Ss had learnt to use the keypunch, and the action of recording the guesses became automatic (unconscious). However no mention is given of how Duplessis evaluated when subjects had achieved this degree of expertise.

**Learning ESP On An Electronic RNG.**

One of the first reported uses of an electronic RNG being used for Ss to learn psychic ability was reported by Targ and Hurt (1971). They used a specially built four target machine\(^\text{13}\), and a 10 year old girl as a subject. The machine lit one of four coloured slides for each RNG decision, and the subjects task was to guess (in real time) the slide which the machine had chosen on the basis of its RNG. Targ reported that the subject gradually acquired the ability to guess the correct slides, and that when she could do this quite well, he changed the machine’s design so it became a precognitive task instead. The S would chose the slide and then the machine would decide which of the 4 slides to select\(^\text{14}\). After the method had been changed to be come precognitive Targ reported that the girl’s talent disappeared and it took some time before she could again get any degree of success with the device in the new mode. Targ interpreted this as being evidence that some form of learning had taken place.

**The ESP Training Work Of Thouless.**

Another, but less successful attempt to demonstrate learning (that predated the well known attempts of Tart), was done by Robert Thouless (1971). Using a version of one of Schmidt’s atomic RNGs (which had become something of a standard) with four lights, Thouless conducted experimental sessions in which he varied the amount of feedback. Although Thouless reported some evidence of a learning effect, this never reached significance, and he considered the experiment a failure.

---

13 This was an electronic noise diode RNG based upon an electronic one mega cycle square-wave oscillator.
14 It is almost impossible to differentiate precognition of an RNG from PK, although Schmidt has tried (Schmidt & Pantas, 1972).
2.4. Experiments Designed To Demonstrate An Effect Which Appears To Influence The Outcome Of An REG.

2.4.1. Investigations Into Different Types Of Random Number Generation.

2.4.1.1. Electronic Noise As PK-RNG Targets.

Since descriptions of most of the widely used electrical RNGs have already been given, including examples of their use, (see types of RNG), this section will concentrate upon the two main alternatives: Atomic decay, and PRNGs.

2.4.1.2. Pseudo Random Number Generators As PK-RNG Targets.

Schmidt’s PRNG Work.

Schmidt (1977a) stated that pseudo random number generators on digital computers were unsuitable for PK tests. Instead Schmidt proposed using the measurement of atomic decay using a geiger counter. In a series of tests outlined in this paper Schmidt showed that the randomness produced by a simple external geiger counter was better than was needed by any current psi task. The trend in the 1980s has been to use these pseudo random number generators in PK tests (Schmidt, 1980; Schmeidler & Borchardt, 1980; Radin, 1982; Shafer, 1983; Puthoff, 1984; Jacobs, 1985; Babu, 1986; May, 1986; Vassy, 1986). Within two years of this publication Schmidt found highly significant PK effects using just such methods (Schmidt, 1980).
Three years after this bold statement Schmidt (1980a) reported using a pseudo random number generator\textsuperscript{15} (PRNG) on an Intel\textsuperscript{TM} 8035 microprocessor. Schmidt was interested to see if his observation of half the initial seeds before they were used to provide feedback to the Ss had any influence on the amount of PK\textsuperscript{16}. The Ss task was to try to influence the direction taken by a random walk on a circle of lights, so the lights went more in one direction than the other. Schmidt reported finding highly significant results overall (p < 0.001), and for the half of the trials based upon his pre-observed initial seeds (p < 0.01). Schmidt concluded his observation of the seeds made no difference and that PRNGs could be used in PK work since a “no more sophisticated truly random generator is required”.

Schmidt also discussed the theoretical implications of being able to detect psi influences on deterministic systems. He concluded that psi could well operate in an acausal manner. However he admitted that if this was the case it would be problematic since humans have very little experience in dealing with noncausal systems.

Schmidt (1983) described how the use of electronic equipment could be used to determine if mental striving can influence games of chance. Using his own experiments as examples he described how the evidence strongly suggested that volition could change truly random events. Schmidt felt it was therefore not unreasonable to assume that games of chance were also susceptible to such influences. He described how the reader could test these assumptions by using a computer to generate quasi-random sequences\textsuperscript{17} for use in psychic games on personal computers.

\textsuperscript{15} Schmidt used the liner congruential method of generation: \( r_{n+1} = 242293 \times r_n - 524287 \times \text{int}(242293 \times m / 524287) \), where \( r_{n+1} \) is the next number in the sequence and \( r_n \) is the current number.

\textsuperscript{16} The performance of a PRNG is deterministic after the initial seeds have been chosen, so that the same seeds will produce the identical sequence of results. Any psi (or observational effect) therefore depends entirely upon the selection of the initial seeds. This is identical to the problems in finding a starting point in a table of random numbers, and there is evidence that this can involve a large psychic component (Morris, 1968).

\textsuperscript{17} Schmidt suggested using the BASIC RND function supplied by the manufacturer of the machine. These generic PRNGs are generally poor, and would produce results that are grossly inflated by artifact (see chapter 3).
Lowry’s Use Of The Commodore Pet’s BASIC RND Function...

Lowry (1981) reported using a Commodore Pet to generate PRNGs for a PK RNG task. He did this by using the PRNG provided by Commodore in their Beginners All-Purpose Symbolic Instruction Code (BASIC) interpreter for that computer. Lowry acted as both experimenter and subject, and reported finding significant PK hitting, and a decline effect.18

2.4.1.3. Comparisons Between Pseudo Random Number Generators (PRNGs) And Other RNGs.

Schmeidler And Borchardt’s Comparison Of The RAND Tables And An Electronic RNG.19

Schmeidler and Borchardt (1981) reported a study which compared a PRNG (from RAND tables) and an electrical noise diode RNG. The electrical RNG used was May’s PsiFi (May, 1976). The Ss were split into sheep and supersheep, and in each session they tried to guess or influence the target. The authors reported finding no deviations from MCE.

Jacobs’s Comparison Between Electronic Noise And A PRNG.

Jeff Jacobs (1985) reported a study which looked at the difference between an electric noise diode RNG (The so called SRU generator), and a PRNG. Jacobs acted as his own subject during these 2 experiments. The feedback was the RNG controlled movement of a white dot on a computer screen around a circle of 16 grey dots. The S’s task was to make the white dot’s motion remain significantly more in one direction than the other. Each run consisted of 32 trials, 16 of which were in one direction and 16 of which were in the opposite direction. Half of the trials were controlled by a noise diode RNG, and the other by a PRNG. Jacobs reported finding a marginally significant effect for the noise diode source, and MCE for the

---

18 Given the very poor randomness that such generic PRNGs produce Lowry’s results must be taken with caution.
19 Radin (1985) Has Objected To The RAND Tables Being Termed Pseudo Because They Were Generated By A True Random Source.
20 Philips 2000 personal computer, which is an IBM XT clone.
PRNG. However subsequent control runs revealed that the noise diode source had a greater variance than theoretically expected. These wider variances cast doubt upon these findings.

Shafer’s Investigation Of The Differences Between A PRNG And Atomic Decay.

Shafer (1983) reported a study which investigated the differences between PRNG and atomic decay. In this study Shafer used a large circle of lights as the Ss target. The behaviour of these lights was controlled by a random walk driven by two different RNGs. These were a Schmidt atomic decay RNG, and a PRNG. The choice of which one produced each run was determined by the atomic source at the start of each session. Shafer reported finding only a weak (non significant) psi missing effect for the atomic source, but nothing for the PRNG. However a post-hoc analysis revealed a significant difference between the experimental and control runs of the PRNG. Shafer interpreted this as evidence that the PRNG was influenced by the Ss to miss, but less so than the atomic source.

Radin’s PRNG Work.

Dean Radin (1982a) reported the results of a series of experiments which investigated RNG PK. The first of these was 100 runs on an atomic RNG with the experimenter as the S. Radin reported significant hitting (p < 0.01), and chance controls. The second series was with 10 unselected Ss who completed 150 trials. Radin reported this as showing suggestive results (p < 0.07), but no control series was run. The third series used a similar method, but revealed no significance. The fourth series used a PRNG for 50 trials, but was found to be non-significant in both the experimental and control conditions. A fifth series used a PRNG, and the experimenter as subject, but the target was chosen randomly after each seed had been determined. Using this protocol Radin completed over 200 trials, and reported very significant hitting (p < 0.001), with a chance control run. In the final study the same protocol as the 4th experiment was adopted, except that the experimenter/subject did not know the randomly determined target until after the run was complete. Radin reported this to have produced significant missing (p < 0.05), and the control was reported to be within chance expectation. In addition to these results the author reported some interesting post-hoc observations. He felt that the experiments that had revealed no significance had all been
during periods where either the phone had rung, people were talking loudly in nearby rooms, or it was too cold in the laboratory during that session. However these observations must be taken with caution since they could have been subject to an unconscious bias.

In a second paper within that same year Dean Radin (1982b) reported the results form an experiment which looked at the attempted influence of a PRNG on a 32 bit UnixTM computer\textsuperscript{21}. Radin reported on four experiments in which two unselected Ss attempted to influence PRNG outputs\textsuperscript{22}. The PRNG was seeded using a method known as 'fast clock seeding'\textsuperscript{23}. This process involves using digits derived from the computer’s clock as the seed for the PRNG. The author reported that the control runs were with chance expectation, and that five of the 18 independent analysis were highly significant (p < 0.005). Radin inferred from these results there were only four possible explanations available. These are that they were due to chance (which is unlikely), an artifact in the RNG, Ss learned to respond at favorable times, or that they are due to 'an unusual human ability'.

**Radin & Bosworth’s Investigation Of Possible Explanatory Models For The PRNG Results.**

At the 1985 PA convention Dean Radin (1985\textsuperscript{24}) reported on a study in which he attempted to determine the process involved in PK PRNG activity. Radin hoped to use the distributions created by Ss when they interrupted the computers clock by pressing a button. The reading of the fast clock was used, via a transformation\textsuperscript{25}, to reference into an array of 5000 random numbers. The values in the array were created before each session using a mixed congruential

\textsuperscript{21} Such a machine derives its timing from the mains cycle. This is approximately 60 Hertz in the US (50Hz in the UK).

\textsuperscript{22} Used the algorithm called 'RAND', which is a standard 32 bit UNIX PRNG (Kernighan & Ritchie, 1978).

\textsuperscript{23} Fast clock seeding (FCS) has become widely used within the parapsychology. Unfortunately FCS makes the PRNG's output statistically undefined (Wichmann & Hill, 1982 & 1987). In such FCS applications the randomness is totally derived from the randomness of the seeds feed to the PRNG. This in turn is derived from the temporal button pressing activity of the Ss. These problems, and suggested solutions are discussed in chapter 3.

\textsuperscript{24} Also published in the same year under joint authorship with Bosworth (Radin & Bosworth, 1985).

\textsuperscript{25} Value := array(clock_value modulus 5000), where the array has bounds of 0 to 4999.
PRNG. Radin felt that this excluded the possibility of the Ss inducing artifacts by regular button pressing. Two Ss were used, one of whom worked from home via a 1200 baud line, and the other used a dumb terminal attached to a data switch and 9600 baud line. Radin proposed comparing the resulting distributions from this experiment against their theoretical counterparts. He hoped this would enable him to determine the process involved. The models he tested were, 'an ideal sensing ability' (ISA), an ISA filtered through a human information processing system, transtemporal inhibition, and a frequency domain filter. Through a series of complex analysis Radin felt that he could discard the second of these proposed models, since the results did not seem to support it. However these results, the resulting interpretations and the experimental method were largely discredited by Richard Hoppe (1987). In a very scathing critique Hoppe identified serious statistical (inappropriate tests), and methodological (optional stopping & selective data removal) flaws with Radin’s experiment. Hoppe claimed Radin and Boswell’s use of an optional stopping rule, coupled with regular feedback, and inappropriate use of statistics made their findings worthless. He even reported creating a computer simulation that outscored Radin & Boswell’s Ss. Not surprisingly Radin and Boswell (1987) disputed Hoppe’s allegations of having selectively removed Ss data after the results were complete, and of using invalid statistical methods. However they did admit Hoppe’s criticisms with regard to the optional stopping artifact, and corrected their finding to be non significant. However Hoppe’s objections mean that the findings from this experiment should be treated with caution.

2.4.2. Atomic Decay As A PK-RNG Target.

A major advance in the use of electronic random event machines was the use of a source of radioactive decay. One of the earliest of these attempts was at Edinburgh University by Beloff and Evans (1961). Although Beloff and Evans did not succeed in finding any apparent PK effect on atomic decay, subsequent researchers, such as Chauvin & Genthon (1965) reported finding that Ss could apparently influence what should have been a steady rate of

\[
\text{Array}_{n} = (1103515245 \times \text{array}_{n-1} + 12345) \mod m,
\]

where \( \text{array}_{n} \) represents the next random number, and \( m = 2^{32} \) to match word length (Knuth, 1969 & 1981).

Radin and Bosworth excluded one of the Ss data before they started analysis because the Ss reported she had disobeyed instructions, and tried to psi miss.

---

26 Array(subscript n) = (1103515245 * array(subscript n-1) + 12345) (mod m), where array(subscript n) represents the next random number, and m = 2^32 to match word length (Knuth, 1969 & 1981).

27 Radin and Bosworth excluded one of the Ss data before they started analysis because the Ss reported she had disobeyed instructions, and tried to psi miss.
an output from a geiger counter. The most frequent and successful experimental user of radioactive decay has been Helmut Schmidt.

Schmidt (Schmidt, 1969) had originally used the decay of radioactive strontium-90 nuclei as an apparent test of precognition (see section on ESP testing). However a year later Schimdt (1970) published some work based upon developments he had made since his precognitive studies of the previous year. He had decided that it was more likely that his Ss were influencing the decay of radioactive strontium-90 nuclei, rather than that they were predicting them. Schmidt had constructed a circle of nine lamps which the binary output of the RNG influenced to move in a clockwise or anti-clockwise direction. The S’s task was to make the display move consistently in a clockwise or anti-clockwise direction, each run comprised of 128 individual trials, and there were four runs per session. In both a pilot and main study significantly negative (psi missing) was reported (p < 0.001).

2.5. Investigations Into The Effect Of Varying The Behaviour Of The RNG.

2.5.1. The Effects of Feedback.

The Effects of a Majority Vote of Trials To Determine Feedback.

Heseltine and Kirk (1980) reported using a majority vote technique to combine trials into small amounts of audio feedback, so they could investigate the role of feedback. The Ss received audio feedback which indicated the size and duration of the majority vote cumu-

---

28 Schmidt outlined the technical details of this random number generator in a paper of the same year (Schmidt, 1970b). It was based around the same method outlined by Furth & Macdonald (1947).

29 Both options are, as has already been stated, equally unacceptable according to current scientific knowledge, unless one accepted Bohm (1957 & 1980) and Dudley’s (1976) suggestion. These two physicists have speculated that atomic decay is regulated by the impact of neutrinos within the atomic structure. This would then mean that Schmidt’s Ss were predicting or influencing the movement of neutrinos (depending upon your interpretation). However since neutrinos are postulated to travel faster than light, and to be able to pass through the earth with such velocity, as to have little chance of contact with other particles, it would be problematic to try to define the process involved in such an interaction, although some parapsychologists have tried (Chari, 1974 & 1977)).
lation of trials. Heseltine and Kirk reported that the first series of experiments showed psi missing on both Majority Vote and individual trials (p < 0.001 and p < 0.05 respectively). They also reported finding a decline effect over the length of the session. The second series was similar to the first except that an additional level of majority vote determined each digit in the target sequence. The authors reported finding no psi in this condition.

2.5.1.1. The Effect of Varying the RNGs speed.

Schmidt’s Variable Speed Trials.

Schmidt (1973) conducted an experiment to investigate the role of the speed of random number generation, in PK success. He compared two speeds of generation, 30 per second and 300 per second, the number of trials were adjusted so the length of time involved in the fast and slow sessions were the same (3 seconds). This ensured that Ss were blind as to which condition they were experiencing. The system gave the options of both acoustic and ink pen recordings as feedback, and Schmidt used all four possible conditions from the two feedback methods, slow acoustic & fast visual, slow acoustic & slow visual, fast acoustic & fast visual and finally fast acoustic & slow visual. Overall Schmidt reported finding significance over all four feedback conditions, on both RNG speeds, with a minor advantage being shown by the slower (30 per second) RNG.

Bierman, and Houtkooper’s Investigations.

Bierman and Houtkooper (1975) reported a complex computer based RNG study. They investigated the effects of random number generation frequencies of 100 and 1000 generations a second, and session lengths varying between lengths of three and 12 seconds. In this experiment they used 14 un-selected Ss and counterbalanced the experimental conditions for each Ss. Bierman and Houtkooper reported overall one tailed significance of (p < 0.05) when using trials as units, and (p < 0.01) when using the sign of the deviation per run pair. They concluded that the results were consistent with what would be expected under observational theory (which they both assume).

In a joint paper Bierman and Houtkooper (1978) reviewed all the fast RNG experiments that had been conducted in Amsterdam between 1973 and 1978. These studies were an attempt
to find a reproducible PK RNG experiment. The experiments usually consist of a set of runs in which the S tries to influence a randomly directed moving line that vertically bisects a computer display. The S’s intent is to make this downward line deviate to the left or right. The target directions are alternated in a control-left-right-control-right-left, sequence. The control\(^{30}\) situation is where the moving line is supposed to conform to MCE. Experimental conditions had manipulated such variables as the run length, trial rate, and feedback format. Bierman and Houtkooper reported finding that the overall results from over six million RNG based trials was non significant. However they reported that if they manipulated the data so they looked at run pairs they found that this ‘majority-vote’ technique produced significant hitting (\(z = 2.54, p < .02, 2\)-tailed). They also reported that they found the hitting ratio in the units of feedback given to the subjects showed significant hitting (\(z = 2.095, p < 0.04\)). Bierman and Houtkooper interpreted this as evidence for a strong consistency within the observed PK effects.

**Millar and Broughton’s Investigation.**

Another attempt to replicate Schmidt’s findings on the effect of varying the speed of the RNG, was made by Millar and Broughton (1976). They reported on an experiment where Ss had tried to influence and RNG while it was running at different speeds. The RNG was set to run at varying speeds, but all conditions would take the same amount of time (100 seconds). This allowed the Ss to be blind to which condition they were in\(^{31}\), and make the duration for each sub-condition equal. The speeds which Millar and Broughton used were: 1000, 100, 10, and 1 trial(s) per second respectively. The Ss task was to try to move a dot from a central point on the screen, Ss tried for 400 seconds, during which time they experienced all four speed conditions. Millar and Broughton reported recorded a total of 111,100 trials, but finding no significant results.

---

\(^{30}\) PK RNG experimental controls are difficult to obtain in parapsychology, since there is nothing to stop the E or Ss influencing the control run to be within MCE (Millar, 1977b; Broughton, Millar, Beloff, & Wilson, 1978; Broughton, 1979). Indeed Rick Berger (Berger, 1986), and Robert Jahn (Jahn & Dunne, 1988) have both noted that their control data fit chance distributions more perfectly than would be expected.

\(^{31}\) It is usually assumed that subjects do not use their ESP to monitor the task in which they are participating. This assumption is made because otherwise, as there is no known limiting factor to psi, any truly covert design becomes complex.
The Princeton Engineering Anomalies Research (PEAR) Group’s PK-RNG Investigations.

The Princeton Engineering Anomalies Research group (PEAR) group reported some results from a highly systematic investigation into RNG PK (Dunne, Jahn & Nelson, 1982). In this investigation the authors had built up a large database of the results from a number of Ss over many sessions. This publication reported on the analysis of the effect of trial sample size, and operators. The PEAR group used a commercial electronic noise diode generator (Elgenco Model 3602A-15124). This produced a random noise spectrum of up to several megahertz. Other logic circuits filtered down the noise to the a lower range of 50-20,000 Hertz. A sampler was then used to read a set of trials in groups of 100, 200, or 2000 respectively. These were sampled at a frequency of 1, 10, 100, 1000 or 10,000 samples per second. The results from each Ss session were recorded on strip printer, and simultaneously to an AIM-65 microprocessor. The AIM-65 then used processing routines supported by a TERAK (model 8510), which was used as a terminal to a DEC VAX 11/750 running Unix, or an IBM 3081 mainframe. These powerful systems then performed all the necessary analysis. These precautions help to minimize the risks of fraud, since all results were recorded independently to two different mediums. The effects that Dunne et al. reported were highly significant, in that the aim conditions +PK, -PK, and baseline all behaved in the Ss intended direction. It was reported that while the MCE for the number of significant series would be 4, PEAR reported having 10 significant series for +PK, eight significant series for -PK, and no significant series for the baseline condition. 3 of the 33 operators got overall significance in the +PK, -PK conditions. The overall results of these experiments show a significant hitting, and low aim (p < 0.01 and p < 0.01 respectively). The distribution of data from the baseline condition also shows a significant restriction from what would be expected by classic statistical theory (p < 0.05). The PEAR research into machine anomalies must be regarded as the most thorough and sophisticated experimental attempt to date. The Pear Group published a number of reports on the results from this large scale database they had established (Nelson, Dunne, & Jahn, 1984; Jahn, Nelson, & Dunne, 1985). The finding from the RNG elements of their work have been summarized above.
Morris and Garcia’s Investigation.

Morris and Garcia (1982) published a report on some PK RNG studies which looked more closely at the role of feedback. They used the MCT system as their source of randomness (Placer, Morris, & Phillips, 1977). In this study Morris and Garcia used two different displays. The first was a small box with two rows of eight LEDs, only two of which were used in this study. The second display was a circle of LEDs. Both displays were controlled by the output of the RNG, which drove the lights at two speeds (1 a second and 60 a second respectively). Within the experiments all the conditions were counter balanced. The authors reported that the slow rate revealed no significant deviations from chance. In contrast the authors reported that the faster speed revealed a significant difference between the displays. The circular display was non-significantly below chance, while the two LEDs showed significant hitting ($p < 0.05$). Morris and Garcia interpreted that the two light LED feedback at 60 changes a second provided little distracting feedback. They felt that this allowed the Ss to interpret the feedback in the most psi conducive way.

2.5.1.2. Varying Generation Speed And RNG Probabilities.

Bierman And Wout’s Investigation.

Bierman and Wout (1977) reported an investigation they had conducted into the effect of the different feedback and generation algorithms, to see if they had different 'psi sensitivities'. They used 20 subjects, all of whom were self-declared healers, and who had a positive belief in the existence of psi. Each S completed three series of 12 runs (20 trials per run). In series one and three they received direct feedback from a single RNG with a hit probability of 0.25, and a generation rate of four trials per second. The second series was different in that it used the output from two RNGs with different probabilities, and speeds. One of these RNGs generated numbers at 50 trials per second ($p < 0.01$), while the second RNG had a rate of five trials per second ($p < 0.1$). Although Bierman and Wout reported failing to find any significance, this study is still of interest since it allowed the Ss to access the RNG

32 A light flicker of 60 Hz would appear as a single uninterrupted glow.
remotely via a telephone link from their own homes. In the author’s opinion this method is under used in parapsychology, since it could produce the most psi conducive surroundings.

2.5.1.3. Remote Sources Of Randomness

In an experimental design that investigated the effects of distance on Pk performance Dickstein et al. (1979) varied the geographical location of the target RNG. One of the binary RNGs was located near the experimental room, and the other was several miles away. The remote RNG was linked to the coordinating computer by a telephone line. One of the RNGs was used to control the lighting of one of two different target lamps (p = 0.5). The decision of which RNG to use was made by the computer based upon a random sequence that was pre-recorded before the session began. The Ss used goal oriented imagery to try to make a designated lamp come on more often. In the pilot study 15 Ss carried out 100 trials each, and received immediate feedback (from the target RNG). The authors report finding no difference between hitting rates on the two RNGs. However they did report finding a significant difference between the target and non target RNGs (for the same time period), (feedback versus no-feedback) REGs (F1,14 = 8.806, p < 0.02). They conducted a confirmatory study with 37 naive and experienced Ss, but this did not reveal any significant effects.

The Syracuse Communication Studies Laboratory’s Investigation.

Another endeavor which investigated the effect of distance and PK was conducted at the Syracuse Communications Lab during the Mid 1980s (Personnel communications Mack, 1987, Morris, 1989). A major reason for developing the protocol was to investigate the effects of research between remote laboratories under mutually secure conditions. In these experiments the Apple II controlled a 300 Baud modem link to SRI in California, and this fed SRI’s RNG data to the minicomputer. In this setup Ss would be blind as to which RNG (local or remote) was feeding the experiment. Unfortunately this equipment was never used experimentally.

33 See also Bierman and Wout (1977) for remote computer based telephone work
2.5.1.4. The Effect Of Feedback Variations On PK-RNG Performance.

PK-RNG Feedback As Degree Of Task Difficulty.

Schechter, Barker and Varvoglis (1983) investigated PK RNG effects on a game called 'Psi Ball'. In this test Ss tried to keep a small ball inside a computer display by using a bat, controlled by a joystick. The difficulty of the game was controlled by an RNG which was sampled five times a second. At each of these five second intervals ten samples were taken. If the sample had less than five hits the game became harder, if there were five or more it stayed the same. Each change in difficulty was not directly noticeable by the Ss. Each S rated the amount of attention they had been focusing on the game on a post session questionnaire. Ten Ss participated in 10 games. Half these games used the RNG contingent difficulty, the other half used an automatic scale. The automatic scale involved no RNGs, and should have produced about the same gradual increase in difficulty. Each S also completed a battery of personality questionnaires. The authors had hoped to look at the difference in game length between the RNG controlled difficulty and the automatic scale version. However an unforeseen artifact in the program made the RNG version significantly longer, so no conclusions could be drawn from the comparison. Unfortunately the authors reported that even after extensive analysis, they could not find any significance. They did however find some correlations from within the questionnaire data. From these analysis they reported that extroverts seemed to do better on psi hitting. Ss who where rated as 'feeling' types on the Myers Briggs Type Indicator had greater deviations from chance. Attention rate was reported to be significantly correlated with PK hitting, so Ss who were PK hitting gave more attention to the task. The authors concluded there may have been some artifacts from multiple analysis. A follow up study to this conducted in the following year (Schecter, Barker, Varvoglis, 1984) failed to replicate any of these findings.

Virtulli’s Investigation Into Feedback And Target Symbols.

William Virtulli (1983) reported on a study which investigated if differences in feedback and target symbols affected the number of correct guesses in a PK RNG test. Virtulli used 60 undergraduate psychology students as Ss, and randomly assigned them to one of three levels of feedback, and one of two types of target symbols. The types of feedback were no feedback, feedback of every incorrect trial, and feedback for every correct trial. Regardless
of feedback condition Ss received their overall results at the end of each run. In a highly controlled and well designed study Virtulli had each S complete their sessions in an electromagnetically shielded room. Inside this chamber Ss entered their guesses into a dumb terminal which was monitored by the main computer system. Each S completed three sets of 25 trials, and was only exposed to one of the three types of feedback. Unfortunately, perhaps due to some unknown factors associated with the stringent controls, Virtulli reported being unable to find any significant results.

**Investigating The Use Of RNG Feedback In The Form Of A Strobe Light Which Induced A State Of 'Fascination' In Subjects.**

In a variation of the feedback methods used on his previous work Schmidt (1978) reported using a strobe light (6-8hz) as a reward for his Ss. The strobe light was used to give feedback on both live and pre-recorded targets produced by atomic decay. Schmidt reports that the effects of such stroboscopic feedback was that Ss either went into a state of happy relaxation, or became literally fascinated. Using an IMSAI micro computer to monitor and control the strobe lights the Ss aim was make the strobe effect more pleasant (or hypnotically fascinating). The exact rate at which this effect was at its maximum had to be determined for each subject before the trials began. Schmidt reported finding that the light was significantly biased to each S’s most enjoyable level, and that this occurred in both pre-recorded and live modes of RNG action. This method has the weakness of relying on the subject’s own subjective account of the point at which they became ’fascinated’ by the light. It is possible that subject may not be the best judges of their own ’hypnotic’ state.

**High And Low Aim Conditions With Multi Media (Audio And Polygraph) Feedback.**

In an attempt to replicate Schmidt’s RNG work Honorton and May (1976) used May’s (1976) PsiFi RNG equipment. In this experiment 10 Ss each had two conditions (high and low aim) for their influence of an RNG. They got feedback through an audio sound or a polygraph needle reading for all of 10 second trials (10 RNG generations per second). Honorton and May reported that this experiment revealed only hitting on the high aim condition (high aim P < 0.01), and a significant decline effect over the ten second period.
Afro-Brazilian Shamanic Cults.

In an investigation into the PK RNG abilities of Brazilian shamans Patrick Giesler (1985) described the results of three studies. In each of these experiments 20 Ss (10 cultists, and 10 controls), were tested using one of Schmidt’s atomic decay RNGs. The three experiments were done with targets of a cult possession god, or meaningless symbol. The cultist Ss were from the three main Afro-Brazilian shamanic cults, the Candomble, the Caboclo, and the Umbanda. Giesler hypothesized that cultists would demonstrate more PK than controls, and would perform better with gods as targets than with meaningless objects. He also hoped that the differences between cults would show a linear trend in PK performance, matching the degree of training required in each cult. Giesler reported there was significant hitting overall, and in the cultist Ss data. Performance with a cult god target did not however reach significance (although it was suggestive). These trends with the cult gods were reversed for the control groups. However Giesler reported that the expected difference between cult groups was not supported, only the order of two of the cults were in the expected direction.

A Large Moving Circle of Lights As A Target.

Morris (1985) reported on some evaluative pilot studies with an electronic RNG controlled feedback light display. These used the same basic method of computerised experimental controls described in Talbert and Debes (1982). The target comprised of a three foot ring of 128 lights mounted on a black octagonal wood frame, the RNG controlled several parameters which determined the behaviour of these lights. The Ss task was to control some aspect of the lights behaviour, be it speed, direction, or intensity. This proved to be quite successful, and Morris reported positive (but not significant) results from a pilot study with this apparatus.

2.5.1.5. Hidden or Silent Trials (Experiments In Which Subjects Received No Feedback.).

William and Lendall Braud’s Investigation.

In an investigation into the effects of feedback Braud and Braud (1979) varied the extent to which Ss were informed of their PK-RNG success. They conducted two experiments. In
the first 10 Ss attempted to influence a Schmidt RNG, with half their trials being in the feedback, and the other half in a non feedback condition. The authors reported that significant hitting only occurred in the non-feedback condition. In the second experiment 20 Ss tried to influence the RNG under what Braud termed no feedback conditions. In this last experiment the Brauds reported a significant hitting effect (p < 0.05). In their conclusions the Brauds suggested there might be only a limited amount of psi available, and that psi that had been expended at a certain time or place might no longer be available for use elsewhere (Braud & Braud, 1979 p29).

Honorton et al’s Investigation At Princeton.

Honorton, Barker and Sondow (1983) investigated the effects of various levels of feedback. They used a Cromemco Z80 microcomputer running BASIC to read from two independent RNGs. The two RNGs were assigned randomly at the beginning of the session to provide feedback and silent runs. Ss got live feedback from the feedback RNG, and overall totals from the silent runs at the end of each run. The feedback was via a coloured thermometer. This indicated the level of the trials, and the cumulative scores, by colour variation, and the scale. Honorton et al. reported that series one with 200 trials showed a non-significant (but suggestive) hitting rates in both the silent and feedback conditions (Honorton et al. reported that some Ss achieved significance in the non-feedback condition). Series two and three revealed no significance in either conditions. The authors concluded that Ss should be screened to remove psi-missers from experiments.

Berger, Schechter, and Honorton (1985) reported on a series of experiments which used the PsiLab (see Berger & Honorton, 1984) games of Volition, Psi Invaders, and PsiBall. The authors reported finding that the PK RNG trials under the silent condition (no feedback to the Ss) had significantly larger effect sizes than those of the feedback RNG trials. Furthermore the hidden trials correlated with the Ss psychological variables. An example of this was that Introverts performed better on the silent trials than Extroverts. Berger et al.

34 The Experimenter got gross feedback of each Ss end of session results, and the Ss got the gross overall results for all subjects. From an observational point of view this made the Experimenter the psi source. This is because it would be the experimenter’s observations which would be assumed to have collapsed any state vectors.
interpreted this as strong evidence that the hidden trials had been influenced by the Ss and not the experimenters, although observational theorists might suggest that these effects are caused by the experimenter’s observation of the silent trials. In this interpretation the correlations would be due to the experimenter’s own ESP ‘reading’ the personality type of the subjects.

2.5.1.6. Investigating The Role Of The Experimenter.

Millar’s Randomness Test.

Millar (1977b) described an original covert method for detecting PK RNG effects. He generated 300,000 digits in 30 minutes which were thought by the subject (Richard Broughton) to be the control runs for his experimental use of a RNG. As such it is certain that Broughton would be most eager for these numbers to be within MCE. However unknown to Broughton Millar had adjusted the RNG output to have a hit probability of 0.1. Broughton’s task was therefore to bias the RNG to reduce 1st to 5th order deviations from theoretical randomness. Millar reported that only the third such test (if considered in isolation) indicated any significant deviations. Millar considered that this was a very promising way of investigating PK RNG effects, since the Ss would be highly motivated. However it could be considered unethical.

Software Exchange Between Experimenters.

In a study which provoked some controversy Dick Bierman (1985) (RB) reported on two experiments which looked at the PK RNG abilities of three ten month old human infants. This study was unique because it involved a software exchange between a parapsychologist (RB) who conducted the first session, and a more skeptical parapsychologist Susan Blackmore (SB) who conducted the second. Before each session a target (integer number) was generated. The computer also generated a series of numbers from a RNG, and these were stored in the RAM memory of the computer. These pre-recorded numbers were used as targets for each trial in succession, and compared with numbers which were generated spontaneously during that trial by the PRNG as guesses of the target number. Two conditions were used: odd numbered trials had feedback for a hit (a laughing face, and a tune), while even numbered trials had no feedback. In between feedback situations that computer screen
was filled with a nonsense block pattern. The results of both series of experiments were split in half and jointly analysed by RB & SB. Bierman reported that in the first series (performed by RB) the feedback hypothesis was not confirmed, but that the second series showed significant hitting in the feedback condition. These results were hotly disputed by SB (Troscianko & Blackmore, 1985), who claimed that artifacts were produced in the RNG in the transfer from RB’s computer to her micro computer\(^{35}\). SB suggested that the randomness involved when using computers should be tested with Ss present. SB further argued that randomness should only be assumed after tests specific to the target machine had been conducted. RB replied to these accusations (Bierman, 1985b), and felt that SB’s points were weak. He argued that the RNG used was immune to the effects SB had postulated, and that the randomness tests conducted had been sufficient.

**Bierman’s Investigation Of The Paranormal Experimenter Effect.**

Dick Bierman (1978) reported on an RNG study which was a further test of the observational theories. Bierman was interested in investigating whether the two contradictory results obtained by two different experimenters, Sanford and Davis were due to a psi experimenter effect. Sanford and Davis had both conducted experiments in to the psi abilities of rats. Bierman conducted a study in which rats were not used, and the results of the series RNG trials were only observed by Sanford & Davis. Bierman split the 40,000 trials between the two experimenters, and predicted that each would bias the results retroactively so they matched their earlier rat based results. Bierman reported that the results did not turn out the way he expected. There was an overall positive deviation \((z = +2.58; p < 0.01)\), such that both\(^{36}\) observers seemed to have biased the RNG in the same (positive) direction.

---

\(^{35}\) It would have been better if SB could have raised these objections before she completed her half of the trials.

\(^{36}\) As has already been stated in a previous footnote, under the observational theories it is very hard to determine what constitutes the act of observation that would cause state vector collapse.
2.5.1.7. Varying Internal Complexity Of The RNG Or Targets.

2.5.1.7.1. RNG Complexity.

Schmidt (1974b) reported an investigation into whether differences in the internal complexity of the RNG had any effect upon the overall psi effect size. The two generators were designed so the simplest version was based upon a one step binary output of the RNG, and the second more complex one was based on the majority vote of 100 binary outputs of the RNG. The 35 Ss obtained their feedback through two different coloured lights, which acted as indicators of hits or misses. To ensure that both the Ss and E were blind to which experimental condition was in effect, the choice of which generator was active for each trial was determined randomly. Schmidt reported significant PK effects, with no differences between the two methods of generation.

2.5.1.7.2. The Effect Of Similar Targets In A Target Pool.

Rex Stanford (1982) reported on an investigation which looked at the effect of similar targets in a target pool on RNG PK performance. In this experiment Stanford arranged that both the E and Ss were blind as to the target and similar non-targets while a random process selected the target. The random process was that the computer cycled through the digits 0 to 9 by counting an undefined number of digital pulses. The result of this counting process was displayed to a liquid crystal display (LCD). This was then used with the RAND tables to select the target. The experimenter used both the RAND tables and the REG to remove any bias that might be present in both. He reported the results gave supported the notion that non-target arrays which provide at least one non-target similar to the real target reduced the probability of the RNG selecting the target.

2.5.1.8. PK-RNG Performance And The Degree Of Lability In The RNG Source.

Braud and Schroeter (1983) reported on some studies that used a computer oracle which they called Algernon, which was based upon a similar system called Alice built at the

---

37 This result is exactly what would be predicted by the observational theory, which is the interpretation Schmidt assumes.
Maimonides medical centre. The responses which the computer made to 16 questions made by Ss were derived from the behaviour of some RNGs. Braud and Schroeter wanted to see if RNGs with a greater lability were more responsive to the PK of the Ss. They tested four RNGs. In order of perceived lability they were; atomic decay, a PRNG seeded 16 times (once for each question), a PRNG using a single seed, and finally a pre-selected file of responses. In the first experiment the authors reported that the degree of significance Ss attributed to answers from each of the methods was: atomic decay, multi-seeded PRNG, Fixed file, and single seeded PRNG. This almost matched the order of lability, although the differences between these groups did not reach significance. In a second experiment the time of a subjects key press determined 100 PRNG digits, which were converted to a binary form. If the number of ones produced by these 100 digits exceeded 50% then the Ss got an answer to their question, otherwise they had to try again. Braud et al. reported that this produced significant missing (p < 0.05).

2.6. Investigations Into PK-RNG Conducive States.

2.6.1. The Effects Of Meditation.

Matas and Pantas (1971) conducted an RNG study which compared meditators and non-meditators attempting to influence a RNG. They used 50 Ss, 25 of whom meditated, and 25 of whom did not. The task the Ss were given was to keep a globe with an electric light inside alight for as much of the four minute session as possible (there were 256 trials in each session). The amount of time the light inside the globe remained on was decided by a Schmidt atomic RNG. Overall the Ss did 6400 trials, and Matas and Pantas reported finding that the meditating group was the only group that achieved any significance (p < 0.05), such that the meditators did significantly better than the non-meditators (p < 0.01).
2.6.2. Comparison Of RNG-PK Performance And Brain Hemispheric Activity.

The Effect Of Listening To Audio Tapes Conducive To Left Hemispheric Activity.

Andrew (1975) reported on an experiment which tried to determine the effects of left hemispheric conducive audio tapes on RNG performance. In this experiment Ss first listened to a 10 minute muscle relaxation tape. Half of the Ss listened to a 23 minute tape which demanded that they counted letters in words. This was presumed to invoke left hemispheric activity. The other half heard a 23 minute tape which had music and other sounds thought to invoke right hemispheric activity. The RNG used in this study was a rather old fashioned high speed electric motor and a clock. The motor was activated for a five second period, after which the motor stopped and from its resting position indicated one of eight lights in a row. The Ss goal was to try to make one of the four lights on the left, or right of the bank of eight come on. The Ss aim was counter balanced half way through each session to remove any bias. Andrew reported the subjects with right hemispheric tapes obtained significant hitting ($p < 0.05$), while those with the left hemispheric tape got significant missing ($p < 0.01$).

Heseltine And Mayer-Oakes’s EEG Brain Hemispheric Investigation.

In an RNG EEG study Heseltine and Mayer-Oakes (1978) investigated the role of brain hemispheric dominance in successful RNG PK. In this study 23 Ss completed two series where the predominately left hemispheric activity initiated the trial, and another two series where the trial was activated by predominately right hemispheric activity. Ss received auditory feedback of a majority vote decision made by the RNG. Heseltine and Mayer-Oakes reported finding that the right hemispheric runs gave a significantly positive deviation ($p < 0.01$ post hoc), and that this deviation was primarily associated with alpha activity. In the left hemispheric runs Heseltine and Mayer-Oakes reported a lower level of significance ($p < 0.025$) in the opposite direction. However Heseltine and Mayer-Oakes reported that they had experienced some data recording problems, and that there was some variability in the task. These considerations make these results less definitive.
2.6.3. Attempts To Find A Correlation Between A Recorded Electroencephalograph (EEG) State And PK-RNG Performance.

EEG Alpha Activity.

There have been several attempts to investigate the possible link between RNG influence and an EEG state. The use of brain state monitoring had become highly popular in psychology, and physiology for many years. Parapsychologists hoped that they might be able to find some correlation between a mental state and psi. In the first of these EEG/RNG studies Heseltine (1977) reported some exploratory research. He claimed it differed from previous attempts by the fact that it involved PK, it required a motor response, and it utilized an auditory target. Sixteen Ss participated in two series, in which a specially designed RNG attached to a computer was used. The Ss were told that the task was an unconscious one, and that a certain brain state would influence the outcome of the RNG. Ss got feedback in part of the trials in the form of a high or low tone (and no feedback for the other part of the series), only one of which was the target for that run. The output was such that the more consistently the RNG produced hits the more of the target tone was produced. Heseltine reported that the results of series one were significant (p < 0.05) for the low tone runs, but were at chance for the high tone series. In the second series the low tone was made to be consistently the target. This made the effect reported by Heseltine to be much stronger (p < 0.01) in this second series. Heseltine reported that the non-feedback runs were all within the chance expectation. He also noted that in series two deviations were found to be correlated with EEG alpha activity, and to an extent certain beta frequencies, but not significantly.

Schmidt and Terry’s Investigation Of Alpha Wave Brain States.

A study conducted by Schmidt and Terry (1977) looked for specific brain wave states to correlate with psi activity. This experiment used an experimental design which forced the subject to be in a specific brain state before the trial could begin. Sixty un-selected Ss, of both sexes between the ages of 17 and 62 took part in 80 sessions each. Forty of these were with an alpha brain state, and 40 in beta. When Ss had achieved the right state according to a preset psychophysiological criteria they received four trials per second of a low or high
pitched noise. These noises were produced by a RNG. The S’s task was to make the low noises disappear. Schmidt and Terry reported finding significant PK effects under both brain states. In a second series Ss had only one tone of feedback, and had the task of reducing it. Unfortunately this second study revealed no significance, however Schmidt felt this was his co-experimenter’s fault, since the final series was done without Schmidt, and in some considerable hurry.

**EEG Theta Activity.**

Another of these EEG studies was reported by Honorton (1977), in which he used May’s PsiFi equipment (May, 1976b). Honorton reported putting his subject (a meditation instructor) in a steel chamber which was also sound attenuated. The Ss then had a 500 trial practice run, before starting the 10 experimental runs of 100 trials each. The hitting feedback was reversed between the 5th and 6th runs, so the target hit direction was switched and became missing and visa versa. However under these conditions Honorton reported being unable to find any significant results. The second series used the EEG gating mode which May had built as a capability for the PsiFi equipment. During these runs (which were additional to similar sessions as used in experiment 1) Honorton reported that RNG activity recorded during the Theta brain wave states (deep meditation of the second series) approached significance (p < 0.06), while RNG output recorded during non-theta brain states approached MCE. Honorton later found by a post-hoc analysis of this data that the Theta state RNG variability was significantly lower than during other non-gated periods. The third series of experiments the Ss were given a post-meditation PK task which was identical to the premeditation task. In this case the high aim condition was found to be significant at p < 0.024; while the low aim was found to be at MCE. Honorton reported that the difference between these two states was highly significant (p < 0.01). He concluded that this was evidence that PK could be directed by directional feedback.

38 Steel (ferrite metal) shielding or buried chambers exclude various bandwidths of electromagnetic radiation. The merits of using them are that they exclude certain interpretations of any influences that might be recorded on the RNG. Unfortunately (at least in the author’s opinion), they may have other undesirable psychological side effects. For example many Ss may find such surroundings unnatural, and even claustrophobic, thus reducing the chances of them performing at their best. The use of shielding will be discussed in chapter 3.
2.6.4. Competitive Long Duration Target Periods.

Debra Weiner (1978) reported using a computer RNG generated horse racing computer game as a test for RNG PK. She used horse racing because she wanted to test if psi was greater when in a condition in which the events took a long time to reach an overall goal. The display that the 19 Ss experienced was four columns of numbers, such that the total in each column advanced to the next number in sequence, or remained unchanged on the basis of an RNG decision \( p = 0.5 \). Each race was over 50 such decisions, and Ss had to choose which of the 'horses' would win (by having the highest number). Ss were split so 8 of them were tested in groups, and the other 11 were tested individually. The computer assigned either a $5 or $25 bet for each race. These bets were counter balanced over a 10 race series. Weiner hypothesized that Ss would want to score better on high risk bets, since that would maximize their winnings. However Weiner reported that none of the major planned analyses revealed anything significant, but some secondary analysis did find a significant difference between individual and group testing \( p < 0.01 \) in favor of individual testing), and a sex difference \( p < 0.01 \) post hoc.

2.6.5. Competition Verses Cooperation.

Broughton and Perlstrom (1984) reported on a series of covert competitive PK-RNG computer games at FRNM. Broughton noted that

Psi tests can be incorporated in and camouflaged by situations which many people find challenging, appealing, fun, and most importantly a common feature of everyday life (Broughton & Perlstrom, 1984 p390).

The authors tried to use a modified version of the computer game called 'Oink!', a competitive computer animated dice throwing game. They amended the program and called it 'Poink!'. In a series of experiments they used 'Poink!' to examine the effect of simulated competitive situations upon Ss. In the first pilot series Ss who had already completed a sports competitiveness questionnaire, took part in a game of 'Poink!', in the full knowledge that it involved a psi element. The overall scores were suggestive but not significantly missing. The authors reported data from one game was lost, due to a computer malfunction. In the second experiment student Ss from FRNM's subject pool were told they were competing against students from another local university, but the psi component was not emphasized.
In fact the Ss were playing 'against' the RNG. Broughton and Perlstrom reported finding no significant deviations from chance. They suggested that

Although it is worth investigating the value of removing the implied request to produce something paranormal for the experimenter it is also possible that it might be counterproductive in some situations.

Further replications by Broughton and Perlstrom (1985) found only MCE in their results. They concluded that stress was not psi conducive.

**Hansen’s Replication Of Weiner’s Horse Racing Study.**

George Hansen (1985) reported an experiment in which he tried to replicate, and improve on a simulated horse racing study conducted by Debra Weiner (1978). Hansen felt that Weiner’s study might not have captured the full imagination of her Ss, due to the unrealistic feedback they were given. He was also interesting in looking at the effects of competition and cooperation in PK RNG tasks. Ss were told to try and psychically help a particular computer animated horse to win a race against other animated horses. The performance of the horses was controlled by an electric noise diode RNG. However Ss did not know if the horse they were ‘backing’ was the same as the horse being ‘backed’ by another S in an adjoining room. Hansen reported finding that, under these blind conditions, evidence for PK was strongest under competition (scores deviated more from MCE during competition), rather than cooperation (which achieved only MCE).

**2.6.6. Volitional Activity.**

Honorton and Tremmel (1979) investigated the dualistic hypothesis put forward by Sir John Eccles. Eccles had suggested that PK and ESP were the methods used by the mind to control the brain. Honorton and Tremmel felt that this theory would predict that PK RNG influence would be highest when Ss succeeded in some volitional activity. In this experiment the authors decided that volitional activity would be getting specific EEG patterns. Honorton and Tremmel reported finding highly significant PK-RNG hitting during the target Alpha

---

39 The logic behind this assumption is hard to follow, since there are alternative explanations if a correlation was found between EEG activity and PK RNG performance. It is also certain that dualists would be able to find equally valid explanations if they did not find such a correlation.
brain state (p < 0.01). However they also reported that they had not monitored the RNG activity when the Ss was not in the target EEG state. This meant that the significance could have spread throughout the entire session. To try to remove this objection Honorton and Tremmel conducted a second experiment. In this they recorded all the PK RNG activity, and included a second baseline condition where Ss were just relaxed. From this second experiment Honorton and Tremmel reported that they found highly significant PK RNG hitting only in the gated EEG periods (p < 0.01). The PK activity for all other conditions was non significant. Honorton and Tremmel took this as supporting evidence for Eccles’ dualistic theory, but they noted there were other interpretations.

2.6.7. Comparisons Of Various Imagery Strategies.

In a reported use of the PK RNG apparatus device called the MCT, the details of which have been described in the section detailing RNGs (Placer, Morris, & Phillips, 1977). Morris, Nanko and Phillips (1979) looked at the effectiveness of visual imagery strategies upon PK RNG performance. Such strategies are advocated by some of the popular books on psychic development. In this study Morris et al. used the MCTs electrical noise diode RNG to control a seven cm ring of 16 light emitting diodes (LEDs). The RNG directed a random walk of these lights around the circle, at a rate of one shift every 1.75 seconds. Morris et al. compared Goal Oriented (GO), and Process Oriented (PO) strategies with a control group (who were free to try any technique they wished). In the first series of experiments Morris et al. reported finding significant hitting (p < 0.05), such that the direction of walk had apparently been successfully controlled by the volition of the Ss. Morris et al. reported that most hits had been with a Goal Oriented (GO) strategy. A second series compared Ss who had previously participated in some form of mental practice such as meditation with subjects who had never done so. This time the GO scored significantly better than PO for both the experienced and unexperienced subjects. The authors also reported that unexperienced Ss who used the PO tended to psi miss, while experienced Ss using PO tended to score at chance.

Morris (1980) reported on a follow up study in Goal Oriented imagery (GO), and RNG PK success. In this study Morris and Reilly used a single light emitting diode (LED) as feedback. The Ss task was to try to influence the LED to glow or dim. The amount of time that the LED remained alight was determined by the RNG (using the MCT RNG (Placer, Morris, & Phillips, 1977)). The RNG had four conditions each with hit probabilities of 50%, 25%,
12.5% and 6% respectively. The assignment order of these conditions were all counter balanced for the Ss, and ensured that the dim or glow target conditions were also counter balanced. Morris et al. reported there was a trend towards hitting but it did not reach statistical significance.

2.6.8. Muscle Tension Versus Relaxation.

Honorton and Barksdale (1972)\textsuperscript{40} conducted three studies which investigated the effect of waking suggestion for muscle tension versus relaxation on PK performance, using an electronic noise diode based RNG. In the first study six Ss attempted to exert a group Pk effect upon an RNG after receiving waking suggestions of either relaxation or muscle tension. The authors reported a significant positive effect (p < 0.04), and a significant difference between the two (tension/relaxation) conditions (p < 0.02). Honorton and Barksdale also reported that runs following muscle tension suggestions were independently significant (p < 0.005), and subjects obtained stronger results following passive concentration instructions (as opposed to active concentration). The second of these studies was an attempted replication with 10 Ss who worked independently, but no effect was found. However a third study using a single ‘star’ subject revealed muscle tension runs were significantly above chance (p < 0.01), and relaxation runs were significantly below chance (p < 0.01), The difference between the two conditions was also highly significant (p < 0.01).

2.6.9. Striving Versus Non-Striving.

Debes and Morris (1982) reported on an experiment which looked at the effects of striving and non striving upon RNG PK performance. Thirty two undergraduate Ss were instructed to try to influence a series of dots that were trailing from the top to the bottom of a computer screen moving randomly one step to the left or the right as they descended. These target conditions were split so half the Ss had one direction as a target and the other half had the opposite as a target. Independently of these groupings half the Ss were self rated as highly-competitive, while the other half rated themselves as low-competitive. Within these four sub-groups half the Ss were encouraged to use striving strategies and the other half

\textsuperscript{40} Also reported by Honorton in the previous year as a PA paper (Honorton, 1971).
non-striving strategies. Each S completed eight runs in feedback and non-feedback conditions. In the non-feedback condition Ss received their results at the end of each run. Debes and Morris reported that Ss who had been given striving instructions scored significantly below chance, while Ss with non-striving instructions scored significantly above MCE. The authors reported that the difference between the two conditions was highly significant. They concluded that non-striving strategies were most appropriate for both competitive and non-competitive Ss.

2.6.10. Release of Effort.

Stanford And Fox’s Investigation.

Stanford and Fox (1975) investigated the release of effort effect. This is where psi appears after the subject has transferred attention to another task after seeming to fail to produce any effect. Twelve Ss tried to influence the electrical resistance of a light stimulated photocell housed in a light proof box. Stanford and Fox used three feedback conditions, false feedback which showed a negative trend, false feedback showing a positive trend, and no feedback at all. At the end of each trial Ss were told to stop trying to influence the cell, as soon as they heard a tone. To ensure they followed this instruction, Ss were given the task of reading aloud from a magazine until the next trial. Stanford and Fox reported significant effects in the non-feedback ($p < 0.05$), false successful ($p < 0.001$), and false unsuccessful ($p < 0.05$). They felt that the significant difference which they found between the trial and end of trial scores confirmed the release of effort effect that had been observed so often in the literature. Later experiments by Stanford (1985) confirmed these release of effort findings. In this later studies a tone disrupted the Ss concentration on the RNG. The disrupted group showed a significant decline effect.

Broughton & Millar’s Attempted Replication.

In an attempted replication to the successful release of effort work of Stanford (1975), Millar and Broughton (1977) reported using a high speed RNG in a covert study. Ss completed a computer RNG session lasting 100 seconds, and then as the Ss were being given their results a second series was run to look for the expected release of effort effect. The RNG used was an electric noise diode, and it generated trials at 100 per second. After the covert session
was completed the Ss were asked to do a second (this would actually be their third) session. The data was then analysed using a split half technique, where one half of the data is used for exploratory analysis, and any trends found in that half of the data are postulated to be in the second half as well. Broughton and Millar reporting being unable to find any results.

**Palmer & Kramer’s Investigation Of Release Of Effort, Startle And Decline Effects.**

John Palmer and Wim Kramer (1984) reported the results of a large scale PK RNG study, which attempted to look at release of effort, startle and decline effects. The experimenters used 48 Ss, who each completed one session of three blocks of 2,500 trials, without any immediate feedback. Palmer and Kramer reported finding a significant difference between 'the absolute critical ratio of the experimental and control runs'. However the RNG baseline control condition was significantly below chance, so it may not indicate any psychic functioning. Palmer and Kramer interpreted this control bias as being due to an experimenter effect, although it could also have been due to a glitch in the RNG. They also found that the absolute critical ratio scores confirmed a release of effort hypothesis.

**2.6.11. Motor Skill And PK Performance.**

In an investigation into the correlations between motor skill and RNG PK performance Debra Weiner followed up on an experiment done earlier in the same year (Morrison and Davis, 1979). In Weiner’s (1979) follow up experiment the Ss manipulated a dial to try to keep a bar in the centre of a computer screen. Motor skill was measured by the amount of time the bar could be kept on the screen. The increments of difficulty in the motor task were decided by the RNG ($p=0.5$) thus implicitly defining the psi task. Each Ss completed three four run sessions in a single setting. One of these four runs included the PK component; the rest were just measures of motor skill. At the end of the final session, but before the Ss were given their scores, they completed a nine item questionnaire. Weiner reported that an Anova on PK scores on best and worst motor skill performance revealed a suggestive ($F_{1,24}=3.94$, $p$
<0.06) effect for male and females. This was such that males scored higher in PK (p < 0.05, two tailed), and substantially better in motor control than females\textsuperscript{41}.

2.7. Learning PK.

2.7.1. Conditioning RNG-PK.

Camstra's Use Of Silence As A Reinforcer.

Camstra (1973) reported an attempt to use involuntary conditioning\textsuperscript{42} to increase PK performance on an RNG task. In two studies Ss had to listen to loud music and white noise randomly interspersed by silence (the amount of silence was controlled by a RNG). It was hoped that this would condition the Ss into producing a bias of RNG, to reduce the amount of loud noise. Camstra reported finding significance in the first study, but failed to replicate the effect when he changed the experimental setting by introducing an additional video screen, and removing the RNG from the S's immediate vicinity.

Broughton et al.'s Attempt To Remove Psi Missing By Negative Reinforcement.

Richard Broughton (1979) designed several studies using a computer game called 'Kop van Jut' ('The Head of Jut'), an RNG controlled device that looked like one of the show of strength games in a fun fair. It had a column of 32 lights, alongside which were a series of signs which provided a guide to the Ss performance. For example 17 was rated as 'terrible'; 25 as 'average' and 32 or above was 'outstanding'. A bell at the top of the column rang when 32 or above was recorded. It used an electronic noise diode as its source of randomness, and the output from this was scaled so Ss saw the lower part of the scale light more quickly than those higher up the column. All the experimental data and the behaviour of the device was controlled by a DEC PDP-11 mini computer. One of the most novel experiments that

\textsuperscript{41} This was probably due to the increased practice that males have in using computer hand eye coordination tasks (video games). See chapter four for a review of the gender bias in video game use.

\textsuperscript{42} Some researchers have cast doubt upon whether this experiment can be regarded as truly testing conditioning (Personal Communication Morris, 1989).
was carried out with this device was that which investigated the effects of aversive electric shocks on negative PK performance (Broughton, Millar & Johnson, 1979). In this study the three experimenters used themselves as subjects, and programmed the PDP-11 to give them electric shocks every time the RNG performed below chance. The theory behind this experimental design was that of behaviorist aversion therapy. Aversion therapy tries to remove undesirable behaviour by conditioning it so it becomes associated with unpleasant stimuli. Ideally the undesired behaviour then becomes extinct. Broughton et al. hoped to be able to do the same with PK Missing\(^{43}\). However, after a number of trials they found no evidence that this technique was effective (except some evidence that Brian Millar might have reduced his tendency for missing). They concluded by suspecting that psi abilities might be uncontrollable.

### 2.7.2. Home Training Devices.

Tart (1982) reported using a Hewlett Packard HP-41c pocket calculator to train ESP, and called the program Professional ESP (PESP). The program, designed by Puthoff and Tart, used a PRNG written by May\(^{44}\), which Tart claimed produced output rectangularly distributed between 0.00000 and 0.99999. The initial seed was determined in this program by taking the day and year (as entered by the S or E), and adding the absolute value of the natural logarithm of the current seconds. The value of the seconds were again entered by the S or E. In respond to the potential risk that Ss might engage in fraud by entering favorable initial values Tart stated:

> You would not expect them (Ss) to figure out that they could cheat by entering these same values every time.

Tart argued that since the Ss actions were recorded on a printer by the program:

> cheating would be readily apparent when the day, year and time values on printout were inspected.

---

43 It is arguable whether a better experimental design would have been to use a reinforcer on PK hitting. The only potential problem with that is there are a limited number of ethically acceptable physiological reinforcers for human beings.

44 The algorithm was to take the seed add PI, and raise the result to the 5th power. This result has its integer part removed and the remaining decimal places are taken as the random number.
As an unofficial training device such a system has obvious benefits. It allows Ss to practice in their own time, and convenience. However the author of this dissertation has some reservations about such a system’s use in formal experimentation. If Ss amended the program, or produced their own version on their own calculator and printer they could easily falsify results. There are also potential problems involved in the limited range of arithmetic available to such calculators in generating the PRNG. Finally the number of instructions the program performed when a hit was registered was more than when a miss was generated. This difference amounts to 0.01 of a second when the program was run, and is within the cognitive threshold of human Ss.

2.7.3. Objections To The Use Of RNGs As PK Learning Devices.

Isaacs (1982) published a review of some of the PK metal bending work he conducted as part of his doctorate, and the reasons that he felt militated against the use of RNGs. Isaacs pointed out there were considerations derived from learning theory which made RNGs unsuitable for training purposes. The low percentile rates of success, result in Ss being presented with a confusingly high level of hits which are due to chance rather than PK. According to Isaacs Ss could not match PK hitting with their internal states because of the high level of false positive hits. To support this Isaacs quoted Tart (1982), who had said that improvement was only possible at an REG task if hitting was 10% above MCE. In Tart’s review of the PK literature over 90% of both dice and RNG reported below this 10% hitting rate. In fact 57% of Tart’s review were in the range of 1 to 3%. Isaacs’ second objection to RNGs was that they provided only a binary feedback, and that an analogue form is much more informative to the Ss. Finally Isaacs detailed the hardware that he had developed for his own PK metal bending experiments. This was the environmentally shielded piezo electric crystal, which had deformation sensors placed upon it surfaces. Three chart recorders registered the activity from the crystal, and were reset to zero volts by an internal clock which output a 5ms pulse every 800 milliseconds (ms), although this reset clock could be delayed for a further 400ms if it detected the start of a PK deformation. The device was fully shielded against various static effects which could have produced false positive results.
2.8. Animal PK.

2.8.1. The Controversial RNG-PK Work Of Dr Walter Levy.

The subject of Animal PK brings us, sadly, to the controversial RNG-PK work of Dr Walter Levy, former director of the Institute for Parapsychology who was caught falsifying his data. Although there are good grounds for ignoring all his work, this would be a loss, since it remains a fact that many of his experimental designs were both novel, and of great interest. Before briefly overviewsing his work however it must be noted that we will not be quoting any of his reported results, due to their suspect nature. Instead we will devote a brief paragraph to the automated experimental designs which he and his co-authors produced.

Levy and Andre (1970) published an account of an experiment where a lamp that gave heat to chicks was connected to a RNG. The light came on every 24 seconds by chance expectation, but this rate of exposure was such that the heat would be insufficient for the chicks to survive. It was postulated that the chicks would influence the RNG so the light came on more frequently, thus providing them with more heat. In his follow up work Levy changed the animals that he was working with from chicks to small rodents (Swiss-Webster mice, and Golden hamsters). Levy produced experimental situations where a testing cage was divided into two halves by a barrier that was low enough for the animals to jump over quite easily. The RNG selected once every minute which side of the cage would receive a low voltage aversive shock (Levy & Davis, 1973a; Levy, Davis & Mayo, 1973). It was hoped that the animals would show a tendency to avoid the electric shocks. However this system was improved (Levy & Davis 1973b) by the introduction of a moving treadmill, which was divided in two with a low barrier. The RNG selected (every ten seconds) which of the two halves would have an electric shock, and the animals were supposed to avoid the shock by jumping over the barrier to the non electrified side. In a further development to this methodology (Levy & Terry, 1973), the rate of RNG shocks was reduced even further to five second intervals. The major advances produced by Levy and his co-workers was the

---

45 This experimental design was originally designed by Rene Chavein and Jean Mayer under the pseudonyms Pierre Duval, and Evelyn Montredon (Personal Communication Morris, 1989).
introduction of a totally automated experimental design. A computer was used to control and coordinate the experiment and data collection.

2.8.2. Braud’s Fish Studies.

In a very original and interesting study Braud (1977) investigated the apparent Pk abilities of aggressive fish. Braud used the ingenious technique of controlling an aggressive fish’s exposure to a mirror via a RNG. Such highly aggressive fish show a classic fighting posture when they see another fish of their own type. Theoretically the mirror therefore allowed the fish to have as much exposure to such a threat as it wished. Braud compared the aggressive fish with a non aggressive species and reported finding a 2% (p 0.05) higher mirror exposure for aggressive fish than for non-aggressive fish. Braud also reported finding a significant decline over the four day period. He interpreted this to either show that fish had become conditioned to this new environmental stimulus, or that the novelty had worn off for the experimenter.

In a follow up study to his previous investigation with the PK RNG abilities of aggressive fish William Braud and James Kirk (1978) looked at the effect of a human and fish trying to influence the RNG at the same time. Braud and Kirk termed this joint observation a human-fish team. The human observer tried to increase the number of mirror presentations to the aggressive fish. The first two experiments were conducted with 10 Ss, each of whom completed 100 trials, with a 20 second mirror presentation. The third experiment increased the number of trials to 200, and reduced the mirror exposure time to 10 seconds. Braud and Kirk reported significance overall hitting, but experiments two and 3, if considered separately, revealed no evidence of psi. The scoring levels for the human-fish teams was approximately the same as that of the fish on their own. This can be interpreted as indicating that the effect shown in the previous experiments could have been due to the human experimenter and not the fish.

2.8.3. Plants.

Hoyt Edge (1978) reported what must be one of the most unusual RNG PK experiments to be published in the literature. He investigated the claim of Cleve Baxter that all living organisms have some form of perception. To do this Edge used plants and an RNG controlled
light source, which he assumed, the plants would wish to keep lighted. However Edge reported that over a seven hour night period the plants appeared to have influenced the RNG so the light was turned off significantly more than would be expected (p < 0.05). Edge conducted four confirmatory studies, the first two of which revealed no effect. In contrast, when the subject plants were replaced for the next two series, Edge reported highly significant effects (p < 0.0001 & p < 0.05 respectively for series three and 4). Edge postulated that the missing results that he had found might be due to some plants disbelieving in psi (a Sheep/Goat effect (Schmeidler & McConnell, 1958 & 1973)).


Schmidt conducted several animal psi experiments over a number of years. Many researchers regard him as the pioneer of these kinds of studies. Initially he conducted them to see if animals could influence some part of their environment. However after he developed his own theory of how psi works (Schmidt, 1975) the rationale behind the studies changed to an investigation into the potential of consciousness in organisms other than humans. It will be recalled that consciousness is important to Schmidt’s observational theory. In his theory Schmidt postulates that the act of observation directly influences reality. Therefore he was interested to find at what point in the chain of evolution consciousness started. To do this he tested various species of animals in experimental conditions where the outcome of an RNG determined an important aspect of their environment.

The first of these studies involved a cat and a heat lamp (Schmidt, 1970c). This was in a similar experimental situation to that which Levy had produced for the chicks (Levy &

46 As already noted in this review, it is difficult to determine the source of influence in RNG PK studies that use non-human subjects.

47 The author assumes that Dr Edge was being humorous in making this suggestion.
Andre, 1970). However Schimdt this experiment predates Levy’s chick experiment (Personal communication Morris, 1989). Schmidt reported that the cat, when placed in a severely cold environment, apparently\(^{48}\) influenced the RNG to make the heat lamp come on more frequently than would be expected by chance (p < 0.01).

Next Schmidt (1970c) reported on an experiment in which a batch of cockroaches were placed on a test grid connected to a binary RNG, which sent an electric current through the testing grid at random intervals. The objective, as with all these rather macabre shocking experiments, was to see if the cockroaches could influence the RNG so they did not receive so many electric shocks. Unfortunately, for the cockroaches\(^{49}\), the test gave them a significantly larger number of shocks than would have been expected by chance alone, a deviation of 309 more than expected by chance (p < 0.001).

Four years later Schmidt (1974) reported on an experiment where he placed a small colony of brine shrimp in a tank, and passed 600 volts DC through the tank with every incorrect RNG decision. Although Schmidt stated that he had pre-tested this method with human Ss (it is presumed this was without the electrical feedback!), and found significance, the shrimps produced no such deviation. Schmidt, who was working with an observational theory of psychic functioning (which might\(^{50}\) require consciousness to produce a state vector collapse), concluded that this experiment might be evidence that brine shrimp did not have consciousness.

---

\(^{48}\) It is always hard to know where the source of psychic intervention is supposed to be in these Animal Psi experiments (or any other psi experiment). It can be argued that it is as equally likely to be from the experimenter as from the animal subject. In most cases the animal is unlikely to understand very much about the ordeal it is suffering, apart perhaps from the fact that it could do without it. Such experimental designs assume that psi allows an organism to understand and control its environment (Stanford, 1974) for its own benefit. In making these comments it should be noted that the author of this work does not feel that causing distress to animals is justified.

\(^{49}\) Schmidt noted that he disliked cockroaches intensely.

\(^{50}\) Doubt has been cast upon the necessity of consciousness in state vector collapse by the recent publication of May & Spotteswoode (May & Spotteswoode, 1988).
As we have already mentioned (above) Schmidt had already conducted such experiments on cockroaches (Schmidt, 1970c), and reported significant psi missing. Given Schmidt’s theory this would be hard to explain, unless it was considered an experimenter effect. However he noted that such electrical shocks were not part of a cockroach’s normal experiences, and would therefore be difficult for the cockroach to understand. To try to take account of this Schmidt re-ran the experiment with the cockroaches (Schmidt, 1979), but this time using electric shocks of a much shortening duration51. The frequencies of these shocks were controlled by a series of different RNGs with different hit probabilities52. Schmidt reported that over 32 sessions there was no deviation from the number of shocks expected by MCE.

Schmidt (1979) reported on some experiments which he had conducted with algae, and yeast, both of which showed no significance. Another similar experiment was one which Schmidt conducted with a thousand fruit flies. In this experiment Schmidt took each fruit fly out of a container and on the decision of an RNG (p = 0.5) either killed the fly or released it53. Schmidt reported this fly killing experiment showed no significant deviation from MCE. He concluded that these findings could be taken as evidence that these lower organisms had no consciousness.

Finally in this series of Schmidt’s Animal PK experiments, Schmidt (1985a) reported the results from an investigation into the effects of an animals prior observation of an RNG

---

51 It is difficult to see how shorting the length of the shock would make it more understandable for the cockroach.

52 If PK was in any way connected with the electrical activity of the organism’s body, such an electric shock could quite well disrupt any psi functions in the body. This would especially be the case when very small organisms were given relatively large electric shocks.

53 If one considers a retroactive effect in this experiment it is hard to see how, once the flies had been killed, they would be able to influence the RNG retroactively to avoid the situation, of course this objection is only justified if the death was very quick. If Schmidt gave each of the 500 flies he killed a slow, and painful death they may have had an opportunity to influence the RNG retroactively. However this would require Schmidt ensuring he did not damage any organs or functions that might be responsible for such influences while he was torturing the animals. If one considered the possible use of precognitive powers, then the flies would simply avoid being the next out of the container until a ‘safe’ RNG outcome. In this case the ‘psychic’ flies would be the ones which Schmidt released. Using such deductive reasoning Schmidt should have re-tested the flies he released, but this would make the organism’s death certain and there would no point in using psi to avoid it.
output, before it was viewed by a human. If an animal could remove any PK effect, then it might be doing so via collapsing the state vectors. This would in turn mean that under Schmidt’s model the organism had consciousness. In these experiments Schmidt first tried using his pet dog as the observer. The dog was shown 128 bits (or trials) of data produced by an atomic RNG. The dog was given a biscuit for each of the hits in the sequence, and nothing if there was a miss (it is uncertain how the biscuits were administered). After this process was complete a human S tried to influence the same sequence. Schmidt reported finding ambiguous results from this procedure. He felt these could have been due to the dog not fully ’observing’ all 128 bits (biscuits). So to ensure that the animal did observe all the bits Schmidt used his pet goldfish, and wired a nine volt DC electrical charge from the RNG through the goldfish’s tank. This was designed so for every ’miss’ the goldfish received an electric shock. Schmidt reported that he found the goldfish appeared to have removed the psi from the data, so a subsequent human observer could not influence it. Schmidt concluded that goldfish could be conscious.

2.9. Allobiofeedback

Braud (1978a) reported on an investigation into the potential of Allobiofeedback in which someone influences another person’s biological functions when given feedback about them. In this study an RNG was used to create either relaxing or stimulating sounds. The S’s task was to influence the person experiencing the RNG to become more stimulated, by influencing the RNG to produce more stimulating sounds. Braud reported that he found significant increases in physiological activity in both the pilot and main experiment (p < 0.01, and p < 0.01 respectively). He also reported that the RNG showed a significant bias (p < 0.01). This could be taken as evidence that the Ss were influencing the RNG and not the target person. William Braud (1978b) also published a brief review in the same year concerned with the role of feedback and egocentric effort in PK RNG studies. Braud concluded that his previous study had shown that trial by trial feedback gave good results, but there was evidence that even without feedback such effects could be demonstrated. He felt instead that the evidence pointed to two factors, and these were that effort and awareness reduced PK RNG results.
2.9.1. The Defense Mechanism Test (DMT).

Martin Johnson and John Hartwell (1979) investigated the kind of effects reported by Braud (1978a) in trying to influence the biological functions of another individual. In this study Johnson and Hartwell used 27 Ss who had already taken the defense mechanism test (DMT). This is a test designed to see how defensive a person’s perception is when subliminally presented with a threatening stimuli. These stimuli take the form of pictures depicting a scene where a foreground figure is accompanied by a threatening background figure (called the ‘hero’ and ‘threat’ respectively). This picture is subliminally presented to the S and their reaction noted. This process is repeated with the same picture at slower presentation speeds until the S can perceive the picture perfectly. The results can be converted into an overall rating of perceptual defensiveness in the range one to 9. The Ss took part in 30 trials each. During each of these trials one of the experimenters (M.J.) viewed a picture which had been rated earlier by the subject as pleasant or unpleasant. M.J viewed each picture for 20 seconds, and attempted to communicate the impression he derived from the picture to the S. The type of target used was determined by an RNG (p = 0.5). During each trial the electrical potential of the S’s skin was recorded by a computer at 20 samples a second. These recordings were then converted to an activity measure by calculating the samples root mean square. This measure was used to try to detect any unconscious psi. As a supplement to this recording each subject made a guess as to the type of target being viewed. Johnson and Hartwell reported finding no support for any of their main hypotheses. These were that less defensive subjects would show greater guessing accuracy, greater skin potential activity, and that all Ss would show different degrees of activity for pleasant versus unpleasant targets. However Johnson and Hartwell reported that an extensive post hoc examination of the data revealed suggestive (non significant) evidence that the RNG had selected an excess of pleasant targets.

2.10. Theories Of PK.

This section represents example uses of RNGs in the investigation of higher level theoretical issues.
2.10.1. Cyclic Psi.

Price (1974) published the results, and a curious observation from some of the RNG work he had been doing with a Schmidt atomic-decay RNG. Price had designed the system so the RNG drove a circle of lights to move either clockwise or anti-clockwise, but when he came to analyze the data he reported that his research team had:

Found some very interesting periodicity in some of our data. We immediately wondered whether the periodic effects might be inherent in the machine in some way; however given the quantum mechanical processes governing the machines function this did not seem probable54 (Price, 1974 p42).

Price then went on to speculate that psi might be in some way cyclical, since he had found two main cycles in his data that were each 180 degrees out of phase. Price’s speculations may have later reemerged in Jacobs’ theory of a cyclic nature of psi (Jacobs, 1987), which inspired the so called adaptive speed trial methodology. There is of course nothing new postulating that ‘luck’ runs in cycles; Paul Kammerer’s ideas are one famous example (Koestler, 1972, p137).

2.10.2. Schmidt’s Mathematical Model Of Psi.

Schmidt (1975a & 1975b) proposed both a mathematical model of psi, and demonstrated apparent evidence for retrocausation. In the first paper Schmidt (1975a) proposed a theoretical basis for psi, which would prove to be one of the most successful theories yet proposed in the field, it’s success being measured in amount of empirical experimental verification it had received. It is not the place of this review to comment at any length upon theoretical developments. Interested readers should instead refer to specific reviews of this model (Braude, 1979; Chari, 1975; Houtkooper, 1983; Millar, 1978; Rush, 1986 p288; Gissurason, 1989).

54 See Furth & Macdonald (1947) for comments on possible regularity artifacts that could occur when using atomic decay as a source of randomness.
2.10.2.1. Retrocausation.

Schmidt’s Experiments.

Schmidt’s (1975b) second paper is of equal importance. In this paper Schmidt reported on an experiment which looked at possible Pk effects on live and pre-recorded targets. Schmidt believed that it would be possible to find such retroactive effects because they had been postulated in his own mathematical theory (see above). In the experiment he got Ss to listen to audio feedback of sounds, the rate of which had been determined by a RNG. Although Schmidt had designed this as a covert psi task, he noted that the Ss may have guessed they were in a psi test. This was since they were visiting a world famous psi lab, and being tested by a parapsychologist. Ss were told to try to listen to clicks on a pair of headphones which were made just at the Ss’s hearing threshold. Schmidt reported finding significant results (p < 0.05) which he interpreted as indicating that the clicks had been influenced to occur more quickly. These results showed a similar effect size on both the live and prerecorded targets. Schmidt interpreted this as evidence that some form of transtemporal (retrocausative) influence had occurred. This interpretation, if correct, is problematic since it introduces the possibility of such effects as the intervention paradox. For these (and other) reasons Schmidt’s theory has produced a considerable amount of controversy; see Rush (1986 p287-89) for a balanced review.

In the following year Schmidt reported on more experiments which seemed to confirm the existence of transtemporal phenomena, in accordance with what was predicted by Schmidt’s (1975a) own theory. In a series of three RNG-PK studies some of the targets to be influenced were generated and recorded automatically by a computer in the absence of both Ss and E. Schmidt reported that the Ss only came into contact with these targets after they had been

---

55 One of the best examples of this is where you could retroactively inhibit your own conception. At which point you would cease to exist, along with the effect which prevented your conception, thus making you re-exist. This would continue ad infinitum.
The first series reported by Schmidt in this paper involved using an electronic (noise diode) RNG with a hit probability of 1/64. The Ss task was to concentrate on the sounds of weak clicks produced every time the RNG produced a 'hit' such that they were waiting (and presumably wanting) to hear such clicks frequently. The clicks were generated and recorded onto magnetic tape, and were then re-played to the subject. Schmidt reported that he had found significant effects from both conditions (p < 0.001). In the second experiment Schmidt changed the design so Ss had the clicks randomly channeled into the left or right ears, while the Ss were instructed to try to increase the number of clicks appearing in the right ear (probability 0.5). During this session half the clicks came from a live source and the other from a prerecorded source. Schmidt had further devised a system where the recording would be played to the Ss four times (one for each of the four sessions). This was to ensure that the prerecorded clicks were exposed to four times as much observation as the live clicks. Schmidt’s theory predicted that the targets that had been exposed to four times the observation would have a larger effect than the momentarily generated targets. This theoretical prediction was confirmed, when according to Schmidt, he found the scoring rates from the two conditions for prerecorded and live were 52.9% and 50.8% (p < 0.005 & p < 0.05) respectively.

In a development of the concept of using prerecorded targets Terry and Schmidt (1978) gave Ss prerecorded tapes of clicks, which they were asked to influence in accordance with assigned instructions in their leisure time. In a short pilot study the targets were generated onto two tapes for each S. These tapes were marked 'High' and 'Low', and each side had six test runs separated by 10 second time intervals. The Ss task was to make the sounds increase or decrease depending upon the specified aim (H or L). The authors reported that they found that Ss were successful in both aim conditions (p < 0.001). On the basis of these results Schmidt and Terry used the same method on a much larger number of subjects.

Schmidt would have assumed that the act of recording the data onto magnetic tape would not have collapsed the state vector of the targets involved (in a sense they would still not exist). In a recent victory for common sense this assumption may have been shown to be invalid (May & Spotteswoode, 1988).

Although Schmidt uses the term effort, from the little that the author has understood from the quantum mechanics involved, the concept of effort is very hard to quantify in such quantum mechanical models. Instead it seems to concentrate more upon the act of observation, hence the use of the term observation.
tapes were mailed to Ss with a simple explanatory note, such that Ss had much less contact with the experimenter. Schmidt and Terry reported getting no results with this confirmatory study. They suggested that the loss of contact with the experimenter may have been to blame for this failure to replicate the earlier success.

2.10.2.2. Multiple Observation.

Schmidt (Schmidt 1985b) reported on three experiments using human Ss where one output series from an RNG was observed by different human subjects. The Ss tried to see if two consecutive PK efforts by different observers had an effect on results. In the first experiment Schmidt acted as S and E. and tried to influence the output of an RNG in a target direction, and then in a subsequent session tried to influence the same sequence in the opposite direction. The presentation was controlled by a computer so Schmidt remained blind to the direction and sequence of the trials. In the next two experiments two more Ss joined Schmidt (male & female). They participated in identical experiments to the first series. Schmidt reported finding evidence from all three studies which indicated that the first observation by a Ss prevents other influences by different subsequent observers.

Morrison And Davis’s Investigation.

Another attempt to investigate Schmidt’s claims of an increase in transtemporal PK RNG effects the more they had been observed was made by Morrison and Davis (1979). They used a DEC PDP11/20 minicomputer which read from an electronic RNG. This controlled the motion of a dot on a computer screen, and provided a feedback sound when a hit was detected. The Ss task was to try to move the dot to the right of its central position on the screen. There were 3 Pk conditions: one where the RNG’s actions were live, and two retrocausative conditions, delay one and delay four respectively. In delay one the RNG output was recorded before the session and played back to the Ss once. In delay four the prerecorded series was played back four times to the same Ss. Ss completed a motor skill test before each session. This comprised of trying to control the movements of a vertical line with a control knob. The control became more difficult throughout the motor skill test, until the line reached the bottom of the screen. This event ended the motor skill test. The duration between start and termination was taken as an indication of the Ss motor skills. After they had completed this pretest session Ss completed two sets of four runs. Each half consisted
of 400 live RNG trials, 400 delay one, and 100 delay four trials respectively. The authors reported finding no significance in any condition. Further analysis which compared PK effect against motor performance also revealed no effects. The authors reported that extensive post hoc analysis only revealed a large amount of variance.

**Talbert And Debes's Investigation.**

Talbert and Debes (1982) from the Syracuse Communication Studies Laboratory reported on the use of the so called 'horizon' software program. This displayed a jagged vertical line of 192 dots which traversed from the top to the bottom of a computer display under the direction of a RNG. The RNG moved the line left or right one character for each binary decision it made. The Ss task was to make the line deviate to either the left or right in accordance with prearranged counterbalanced instructions. In this study 16 Ss (8 male and eight female) sat in a sound attenuated room with a 19 inch colour monitor, which provided their feedback. The two bit noise diode RNG\(^{58}\) was controlled by a DEC LSI-11 PDP11/03 (with dual eight inch floppies and an operators console). This mini computer was connected to an Apple II microcomputer, by a parallel interface. The Apple II was used to control a monochrome screen in the computer room, and the S's colour monitor in the attenuated room. The randomness used to control the target line was split into two conditions, those which were derived either as the Ss saw them\(^{59}\), or were prerecorded 30 seconds before the session started. Ss were blind as to which of these two conditions they experienced. The other condition was a variation in the feedback. Half the sessions were conducted with only half the screen displayed, while the remaining sessions used the whole screen. The Ss did eight runs of 192 bits, with a non-experimenter making the assignments between full and half screen conditions. Although the authors reported finding no overall significant deviations from chance, the prerecorded targets were significant \((p < 0.05)\). However they concluded there were real time alternatives which could explain these findings.

---

58 The authors noted that this RNG produced randomness with some 3rd order dependencies \( (3^{240}) \), but that these were irrelevant to this task.

59 Even these real-time trials would be within the variations involved in allowing for the PDP to read the RNG, the hand-shaking between the 16 bit DA interface and the 8 bit Apple II, and the delays in displaying the feedback.
2.10.3. The Psi Mediated Instrumental Response (PMIR) Theory.

In a paper by Stanford et al (1975), the investigators hoped to investigate the validity of the Psi Mediated Instrumental Response (PMIR) theory of Stanford (1974). This model stated that organisms used psi in the service of their own needs, even when they are not consciously intending to do so. In an attempt to test this Stanford et al. used an RNG based study with 40 male college students. The experiment was designed so the Ss first participated in a conscious RNG study, and then were introduced to a boring task which could potentially take 45 minutes to complete. However unknown to the Ss the RNG determined when they would be released. The electronic noise diode RNG generated one trial per second, with a probability of success of one sixth. The Ss were released from the boring task when or if they had accumulated a total of seven or more such hits in one of the successive blocks of 10 trials. Stanford et al. reported a significant number (8) of the subjects managed to escape the boring task (MCE = 2.9 subjects, p<0.01). The overall bias of the RNG was also reported to be significant, (t = 2.19; df = 39; p<.05). There was no significance reported for the intentional RNG PK task that the subjects participated in before the covert study. However Stanford et al. did report there was a positive (but non-significant) correlation between the scores on the conscious and unconscious PK tasks.

2.10.4. The Intuitive Data Sorting (IDS) Model.

In one of the major theoretical papers of the eighties May et al. (1985) put forward an alternative explanation for the PK interpretation of much of the RNG PK work published over the previous 20 years. May et al. put forward the theory that using psychically obtained information Ss selected locally deviant sub sequences from longer truly random sequences. To support this theory May et al. used the results from Radin’s meta analysis of the RNG PK studies (Radin, May & Thomson, 1985)60. They reported finding that only a Goal Oriented (GO) physical force, or their Intuitive Data Sorting (IDS) model could account for the overall statistical distribution. They also reported that the GO force would have to have

---

60 This was a meta-analysis of 56 reported RNG studies between 1969-1984, which produced an overall significance level of 10^-43, against the reported effects being caused by chance.
adjusted its effect to match a $1/\sqrt{n}$ factor decline found in the meta-analysis data. They felt this evidence, combined with the problems in invoking a 5th force in nature\textsuperscript{61}, made the IDS interpretation the most likely one to account for the substantial RNG-PK effects.

In a subsequent SRI internal report produced by May (1986), he reported on an attempt to use the PEAR database to further test the IDS model. In these publication May was more explicit about the weaknesses he felt were inherent in the GO force (termed Remote Action (RA) in this report) interpretation of the PK RNG data. These were that the force had to be controllable to with one millisecond (ms), and interact equally with weak and strong atomic forces. May stated that

because there has been no evidence to date to support the idea that computer hardware is susceptible to a putative RA interaction\textsuperscript{62}, we assume that a purely pseudorandom number generator (PRNG), which is seeded by a computer clock, constitutes an environment that precludes RA\textsuperscript{63} (May, 1986 p20).

May reported the results of a series of experiments using PRNGs conducted at SRI which May claimed empirically demonstrated the IDS model\textsuperscript{64}. Using a Sun 3/160-c computer May used the FCS\textsuperscript{65} method to feed a Kendall shift register feedback PRNG (Lewis, 1973). May also reported the results of an analysis of the PEAR RNG data, but found it was unsuitable since it had not been recorded specifically with the IDS model in mind. May stated that more research needed to be done using PRNG, to enable the IDS model to be proved. He concluded that if the IDS model was proved to be correct:

---

\textsuperscript{61} Morris (personal communication, 1989) has noted that the IDS model would also seem to invoke new forces.

\textsuperscript{62} This statement is open to debate. Isaacs's work with Macro-PKMB, Schmidt's work with atomic decay, and Cox's mechanical devices would all be examples of such RA effects on hardware that is similar to that used in computer construction.

\textsuperscript{63} This statement is also open to dispute. Morris (personal communication Morris, 1986, & 87) has pointed out that the act of pressing a button is open to a PK influence on all the biological functions involved.

\textsuperscript{64} There is an obvious paradox with an experimental design that sets out to demonstrate that a paranormal Goal Oriented force does not exist.

\textsuperscript{65} This used the low order 15 bits from system clock (recorded on each Ss button press) as the seed for PRNG.

---

94
We would call into question any experimental results from any discipline that claim cause-and-effect relations based upon statistical inference (May, 1986 p29)\textsuperscript{66}.

\subsection*{2.10.5. Experiments Trying To Determine The Difference Between PK And ESP (Real Time PK Verses Precognition)}

Schmidt and Pantas (1971\textsuperscript{67}) reported on a study in which two RNG machines looked and performed identically, but were internally very different. One of them was a normal precognitive or PK test, the other tried to force the subject to perform what Schmidt claimed was a pure PK test. Schmidt reported finding that both systems produced similar results (normal PK or precognitive system \(p < 0.05\), and the PK only system \(p < 0.05\)), Schmidt concluded that the two paradigms made little or no difference to psychic functioning.

\textbf{Heseltine’s Investigation.}

Heseltine (1985) reported on some studies which used a single female subject, to look at the differences between real time PK and precognition. Heseltine was also interested in observing if structured runs produced by an RNG showed any difference to similarly produced non-structured ones. In the feedback condition the S’s task was to observe a horizontally moving random walk across a computer screen. This was in contrast to the non-feedback condition of a flat line. Both conditions were counterbalanced. The S was instructed to press a button if she felt that the line was about to move downwards, and to do nothing for any other condition. Heseltine reported finding a significant PK effect on both real time and precognitive trials, but found no difference between the scoring of structured and unstructured runs (both conformed to MCE).

\textsuperscript{56} Of course this would include the IDS model itself (personal communication Morris, 1989) \textsuperscript{57} Also reported as Schmidt & Pantas 1972 under a slightly different title
2.11. Review Of Discussions Of The Use Of RNGs.

2.11.1. Problems With RNGs.

In 1947 the problems involved in the generation of random numbers or events were detailed in two excellent papers.

Furth And Macdonald’s 1947 Review.

Furth and Macdonald (1947) looked at physical ways of generating random numbers for statistical purposes. They reviewed both of the then novel methods that used electric noise diodes and the counting of the particles from sub-atomic decay. Much of what they had to say about the problems involved in using such processes are as valid today as they were 42 years ago. The noise diode was at that time a thermonic delay valve, but such apparatus still used the random motion of large samples of molecules and electrons (in the range of $10^{23}$ particles involved in a noise diode) that are used in modern noise diodes. The only major differences are in the expected operational life span, speed of operation, amount of current required, and the size of the physical component involved. Some of Furth and Macdonald’s comments on the use of radioactive decay have survived the rigors of time less well, since many of the problems they describe such as availability, or speed of measuring apparatus, are no longer valid. However their warnings against using insufficient scaling currents, or inadequate amplifier bandwidths in the recording process are still valid. Either of these two faults can induce a bias into the output of an atomic RNG so the sequence seems regular or repeating. Obviously this would be disastrous in a parapsychological experiment.

Wilson’s Review.

The second paper in that same year was more directly concerned with the parapsychological use of random number generation (Wilson, 1947). In this paper Wilson reviewed the use of random event generation in ESP experiments, and the problems that can be encountered in their use in parapsychology. Wilson criticized the design of many of the machines used in previous research (particularly Tyrrell’s machine), since he felt that they were too dependent upon the subject’s temporal behaviour in pressing the button to derive their randomness. Wilson rightly felt that it was essential for responsible parapsychological research to be using truly random processes; otherwise the use of the assumption of randomness implicit
in statistics would be invalid. Wilson also agreed with Tyrrell’s (1936) statement that a finite deck of ESP cards could not be assumed to be random, and that the use of statistical analysis upon such nonrandom events was incorrect. To correct these errors Wilson ended his paper with the specification of an improved randomness source based upon a noise diode to select one of four lamps, which automatically recorded the selections made by the subject. A version of Wilson’s machine was subsequently built by the SPR.

Eighteen years after this work on random number generators (Wilson, 1947), Richard Wilson again discussed the use of random number generators in ESP research, and stated there was no such thing as the perfect random number generator. However he proposed that by improvements in the design parameters (which he outlined), a generator with a precision of 0.03% could be constructed, which he claimed would be some 30% better than those currently being used within parapsychology.

May’s Review.

May (1976a) discussed the sources of possible bias in electronic random number generation. He stated that these artifacts were unlikely to have affected the PK RNG work that had been done to date. It was May’s opinion that as equipment became more complex and the speed of the required random number generation increased the field would run the risk of these speed and tolerance artifacts. The artifacts associated with the use of high technology equipment is discussed in chapter three.

2.12. Alternative Hypothesis For These Effects.

Some critics would rather suggest that parapsychological findings are in some way direct consequences of the role of chance. This can either be in the form of parapsychologists

---

68 It has been pointed out that good experimental design can get a round the problem of having a based generator (Personal communication Morris, 1989). However the author feels that such measures might provide critics with more scope for writing misleading articles

59 It should be noted that PK based work with random event generators was still very much in its infancy. Much of the PK work was based upon dice (Rhine, 1943; Forwald, 1953), or on novel and highly complex mechanical devices such as those devised by the eccentric, but highly innovative W. E. Cox, a good example of which is the so called 'clocks machine' (Cox, 1966).
misunderstanding chance expectancies (covered in chapter one), or more radically that the phenomena itself is an aspect of chance. The latter would imply a flaw in the current statistical theory, the results of which could have strong implications for methodology throughout the sciences.

2.12.1. Misunderstandings Of The Role Of Chance.

In his book Alcock (1982 p156) reports the research findings of, two skeptical researchers, Pitkin and Mulholland (1938), who made an investigation into the role of chance in ESP card experiments. According to Alcock they did so using 200,000 coloured punched cards (a form of early computer input), half of which were red and half of which were white. The cards were further split into smaller subgroups of 20,000 each, and the first of each of these groups was punched with a number one, each of the next group were punched with the number two, and so on right the way through to the digit ten, which was punched on the last group of 20,000. The two sets of red and white cards were then mechanically shuffled and fed through the IBM computer, comparing the orders of the red and white decks of cards. While chance expectation would be a match for every 5th card, some of the sequences Pitkin & Mulholland reported finding were as long as 32 pairs without a match, and runs of eight matching pairs occurred 780 percent more than would be expected by chance alone. Runs of seven pairs occurred 59 percent more often than would be expected by chance, and three times as many runs of five occurred in the first 40,000 pairs of cards, than in the remaining 60,000 cards, which they claimed were similar to the decline effects found in the ESP experiments. The results reported by Mulholland, and quoted by Alcock, sound extremely disconcerting. This is because if the biases reported by Mulholland occurred 'by chance' than it would seem to indicate a flaw in generally accepted statistical theory. However all is not as it might seem, for a detailed re-analysis of Mulholland’s data by Donald West (1945 & 1962), which was omitted in Alcock’s review, revealed that they were in close conformity with MCE. It would appear that Mulholland (and Alcock) had misrepresented these findings, so it would appear that ESP results could be explained by chance alone.
2.12.2.  Spencer Brown’s Theory (Patterns Within Randomness Itself).

Spencer Brown’s (1953) theory was that naturally occurring patterns in random number sequences were the cause of significant ‘psi’ results.\(^{70}\)

**Oram’s Investigation.**

To investigate this possibility Oram (1954) used the Kendall Babington-Smith random tables to generate a whole series of simulated subject guesses and responses. He reported finding no significance in over 98,000 simulated trials, except a slight decline effect. However Spencer Brown (1957) criticized Oram’s investigation, and claimed he should have used the quartile split analysis method currently popular at the Duke laboratory since this revealed a significant quartile decline in Oram’s data.

**Harvie’s Investigation.**

Harvie (1973) reported the results from an experiment designed to test the theories put forward by Spencer Brown (1953a, 1953b, 1957). Harvie tried to simulate ESP results by matching random numbers from different sources against one another. Using 50,000 numbers from the RAND table of a million random numbers and the Kendall Babington Smith tables, he compared them against numbers produced by a computer’s pseudo random number generator (PRNG). He reported finding a highly significant result (p < 0.01) in the missing direction. He concluded that:

Perhaps some apparent successful ESP experiments are indeed pointing to an unknown property of that insubstantial territory which we call randomness? (Harvie, 1973 p136).

In a JSPR review of the book in which this experiment was reported Robert Thouless said that Harvie’s finding..

Would seem to point to some uncertainty as to the expectations based on orthodox probability theory.

\(^{70}\) Within parapsychology it has been admitted that Spencer Brown’s observations could explain some deviations from chance found when using PRNGs. However it felt that it is unlikely that they could produce deviations of the size and consistency reported in parapsychology (Soal, 1953; West, 1962 p121).
Ward’s Attempted Replication.

Ward (1979) reported an unsuccessful attempt to replicate Harvie’s finding using the Edinburgh parapsychology lab’s live RNG to generate the sequences, and the University’s mainframe computer to analyze them. However it is obvious from reading Ward’s paper that he was biased against finding any deviations from MCE. Since he was using a live RNG to generate his sequences this does leave open the possibility that he influenced the source to conform with his expectations. Of course it is equally possible that those experimenters who have found significance using Spencer Brown’s method used ESP.

Shallis’s Investigation Into PRNGs.

Michael Shallis (1982) published a book titled ‘On Time’, which as well as investigating time, and physics, detailed a computer based replication of Harvie’s (Harvie, 1973) experimental evaluation of Spencer Brown’s randomness theory (Spencer Brown, 1953, & 1957). Shallis noted that both Oram (Oram, 1954), and Harvie (Harvie, 1973) had investigated Spencer Brown’s claims. Shallis conducted some experiments using computer based PRNGs. In the first he compared a PRNG against itself for a simulated 175,000 trials, and found significant missing (p < 0.0001). With such strong results Shallis was concerned that he should take steps to remove any bias that might be effecting the results, so he used a bit shuffling technique. This method was to take the least significant digit (LSD) of a generated 11 digit number as an offset from the most significant digit (MSD), to give the actual digit. An example will probably save considerable space. If the digit was 62947629971 then the LSD is 1, and the first digit from the MSD end gives the result 2. Using this system on

Shallis does not mention the experiment by Ward (Ward, 1979).
sequences derived on three different machines. Shallis simulated a further 85,000 trials and reported finding significant missing (p < 0.01). Shallis interpreted these results as supporting Spencer Brown's argument.

2.13. Discussion.

This chapter has tried to provide a background to the use of electronic tools in parapsychological research, it is closely linked with chapter three, which should be regarded as the discussion which results from this review.

---

72 Shallis gives no details about the algorithm used in these experiments, or the initial seed values. If he had not controlled for these factors his findings would be invalid (see notes on PRNGs).

73 During testing of the KMDB system the author evaluated many different RNGs (n > 10), and conducted a series of intense statistical evaluations upon the RNGs he finally selected. For example in testing PRNG Algorithm AS183 (Wichmann & Hill. 1982 & 1987), he conducted over 40 million simulated trials. This is orders of magnitude greater than the combined number of trials in the experiments outlined above. These simulated trials showed no evidence in support of Spencer Brown's theory.

74 As noted at the beginning of this chapter the content of this chapter and chapter three form the basis of a section in the forthcoming 'Advances in Electronics and Electron Physics 1990 edition' (Academic Press) entitled 'Electronic Tools in Parapsychology', to be co-authored by Robert Morris.

Glib’s Laws of Unreliability.

1) Computers are unreliable, but humans are even more unreliable.
2) Any system which depends upon human reliability is unreliable.
3) The only difference between the fool and the criminal who attacks a system is that the fool attacks unpredictably and on a broader front.
4) A system grows in complexity until the resulting unreliability becomes intolerable.
4) Self checking systems have a complexity in proportion to the inherent unreliability of the system in which they are used.
5) The error detection and correction capabilities of any system serve as the key to understanding the type of errors which they cannot handle.
6) Undetectable errors are infinite in variety, in contrast to detectable errors, which by definition are limited.
7) All real programs contain errors until proved otherwise - which is impossible.
8) Investment in reliability will increase until it exceeds the probable cost of errors, or someone insists on getting some work done.


3.1. Security Of Parapsychological Computer Systems

The personal computer has made a large impact upon parapsychological research, such that it is now rare to find any experimental papers in the main parapsychological journals which

---

1 The content of this chapter and that of chapter two form a sub-set of a section in the forthcoming 'Advances in Electronics and Electron Physics 1990 edition' (Academic Press). This will be entitled 'Electronic Tools in Parapsychology', and will be co-authored by Robert Morris.
have not used a computer in some way. The affects of the arrival of the computer on parapsychological experimental method has been shown in the review in chapter 2. Although the computer has reduced the potential for Barber’s (1976) experimenter\textsuperscript{2} and investigator\textsuperscript{3} effects in some ways it has increased them in others. Rather than attempt to categorise aspects of security into investigator or experimenter effects (this would be too situation dependent) we will adopt the terms experimenter or researcher, which will assume the possible dual roles of experimenter, and investigator.

The modern parapsychological researcher has the unenviable task of having to gain mastery over multiple disciplines to be sure that no artifact produces a false positive or negative result. If an experimental study uses a computer then this required expertise includes the areas of computer science, and electronics. This section will try to outline the problem areas involved with the use of such electronic equipment, and some potential solutions to these problems. The security measures designed and implemented in the project described in this dissertation exceed any realistic threat. They were used and are described here because one of the aims of the dissertation was to develop the strongest methods of practical security to help future parapsychological computer based research. No where in the parapsychological literature has computer based security methodology and standards being comprehensively covered. It was hoped that in developing such standards and reporting them here the problems of replication, and artifact would be reduced, allowing a clearer picture of any actual psi processes involved.

### 3.2. The Most Fundamental Security Problem.

Sadly those involved in fraudulently creating anomalies have sometimes been suspected to include not just the subjects but the parapsychological experimenters themselves (for references see West, 1962 pp128-30; Rhine, 1975; Barber, 1976 pp38-41; Palmer, 1978;

\footnote{Barber (1976) defines the experimenter as the person who conducts the experiment.}

\footnote{Barber (1976) defines the investigator as the person who plans and coordinates the experiment.}
Palmer, 1986 pp214-17). Some parapsychologists feel that experimenter fraud is so bad within the field that based upon their knowledge of psi they created a 'rule of improbability'.

Reports of extremely high or extremely stable psi performances are in contradiction to general experience and warrant the conclusion that they are most probably due, at least in part, to error, delusion, or deception. I have named this the 'rule of improbability (Timm, 1983).

The reasons why an experimenter should wish to falsify a paranormal event are numerous (Gregory, 1980). They include the advancement of their career (or perhaps the justification of their area of study), financial security, enhancement of their reputation, or the advancement of their personal belief system. There is no effective way to prevent experimenter fraud. In all security systems there has to be some point at which something or one has to be trusted, and in experimental design that is the experimenter. There are ways of reducing this possibility, by involving multiple independent coordinators. The military have found that two independent trusted controllers, (each with equal system expertise) give the best trade off between expense and dependability (Durham, 1987). It was through just such a system that the Levy affair was uncovered (Rhine, 1975). Indeed Rhine recommended dual experimenter methodologies in his report on the Levy affair. Unfortunately the numbers of active parapsychologists, and the distribution of their computer expertise, make the universal adoption of such methods unlikely. It must therefore be kept in mind throughout all this chapter that the use of these suggested computer security standards will only help reduce non experimenter induced artifact or fraud, not eliminate it. It is also important to note there is no such thing as a totally secure system. Any security system can be breached, provided enough resources are made available. The job of the responsible experimenter is to make such a breach unjustifiably expensive for the potential fraudulent person. It will be seen that it is cheaper to make a system detect such attempts than it is to stop them.

4 Morris (personal communication, 1989) has noted that experimenter fraud is more likely to be of concern in fields where replicatability is very low.

5 This review will also provide general coverage of many methods and details of breaching or disabling electronic systems which will not be described in detail, since to do so would be against the interests of national security.
3.3. Anomalous Human Computer Interaction.

Previous parapsychological publications in the area of explaining possible effects that can influence computer systems have been scarce (Wilson, 1947; Girden, 1962; Hansel, 1966 & 1980; Millar, 1977; Broughton, 1982; Rush, 1986). In a previous publication Morgan (1987), the author of this dissertation, outlined a schema for evaluating possible computer security threats, however this was far from ideal. It lacked any technical detail of the problem areas, and outlined few truly effective solutions. While its overall schema of breaking down anomalies could be regarded as a reasonable starting point, it spent unjustified amounts of effort on unlikely, or unimportant aspects of computer system security, and concentrated on non-experimental paradigms. Since that paper is reproduced in appendix 1, this review will pay particular attention to the specific problems that face experimental parapsychology when it uses a computer in its experimentation. To do this each problem area will be described and then potential solutions will be discussed. This will be done with direct examples from the methodology developed and used in this dissertation. This approach allows a full description of the security used in this experiment, however it does mean that the proposed problems and solutions provided in the examples will be derived from the target experimental computer systems. A possible problem with this method is that it will not cover weaknesses in the other main computer systems that have been used in parapsychology. To try and redress this footnotes and separate summaries will be used detailing any major differences that are relevant for any of the main alternative computers covered in the review in chapter two.

---

The term security will be used to cover both intentional (fraud), and unintentional (artifact) breaches of the integrity of a computer system.

The scenarios described in the paper were derived, in part, from a popular anecdotal account of supposed paranormal human influence of computer systems (Geller & Playfair, 1986). The author however does not claim to have shown how the phenomena in this book were in fact produced.

IBM clones running the operating system MSDOS V3.20.

In this review we will split experimental security into three main temporal divisions, before the experiment, during it, and after it. The infiltrator will probably decide to attack at the point where the target system has the smallest amount of protection. It will be at this time that the most effective (minimum) action can be taken to influence the outcome of that system. The actions required in each case will be different, and will reflect the nature of the study. However they will be a commonality among all situations, and it is this we will outline in the following sections.

3.4.1. Pre-experimental Influences.

3.4.1.1. The Shared Resource Problem.

As a piece of experimental laboratory equipment the computer is unique. The average parapsychological experimental computer system will be used to fulfill several roles, this is unlike such items as stopwatches, or targets. These will range from clerical procedures to multiple separate experiments. This overlapping of use means that many people will have access to the same physical hardware, they may even share files. Any scenario that allows a computer system to be accessed by many people for varied purposes has a high vulnerability to being attacked. Research has shown that the majority of computer fraud is committed by people within the target organisation (Wong, 1988, Hamilton 1989). Although steps can be taken to try to protect certain aspects of the experiment, if other users have unrestricted access to the target system the fact is the security of the experimental data cannot be guaranteed.

3.4.1.2. Preparation.

Every computer based experiment requires preparation time of some sort. In most situations this involves some form of programming in a computer language, unless the proposed system is being prepared by a second party. Obviously in the case of systems created by second party's most of the security problems during the systems creation would be the second party's responsibility. However it remains the experimenter's responsibility to check that the system performed to within the required experimental specifications, without introducing any
security breaches of its own. It is good practice for the experimenter to check that all second party software performs correctly before it is used in an experiment. For example ‘off the shelf’ packages for statistical analysis should have their results checked against proven results from sample data. It is ideal to have a test suite of data which has already been subjected to a trusted form of the processing the second party’s system provides. This can be available specifically for this testing purpose. For the remainder of this section we will deal with the more standard situation of the experimenter creating their own experimental system. It should be noted that although this review uses the term experimenter, this could be a team of researchers; the singular form has been adopted to simplify explanation.

The experimenter designs some form of controlling logic which will coordinate and record all the relevant events which takes place during the experiment. This design process is usually protracted, and several stages and versions of the system have to evolve through a development process. This is often the time when the weakest security surrounds the experiment. Most experimenters do not guard the system during this development stage, and yet they will assume the final system is totally secure. This confidence could be unwarranted, since every detail about the functioning of the system would have been available in the development versions. Several things could happen if an adversary obtained access to the system at this stage. Details of the file structures could allow false data files to be created. It is also possible that security features and experimental details such as randomization procedures could be learnt. It is therefore important to ensure that some minimum security

If the experimenter wishes they can give the test data to the second party to help in the de-bugging process. However it is sound practice to ensure that the second party does not have all the test data. Allowing the second party to have all the test data could mean that the system becomes ‘patched’ to be able to run the test data correctly. A separate suite of such data allows the experimenter to have an independent check.
standards are adopted. For example, printouts of previous versions should be destroyed when a new version is created, and the current listings should be kept secure, so they are accessible only by the developer, or other trusted individuals. Similarly, the machine being used to develop the system should ideally be made secure when the development is not taking place. These precautions are of course only possible when the machine is the sole property of the developer, and the development is being conducted at a secure site. In reality, most laboratories will have an open access policy towards their machines, meaning that many people are authorised to use the machine, since computers are a valuable resource. This is particularly true if the lab is in an academic environment, as was the case with the system described in this dissertation (which will be referred to as the KMDB project, or system). This system development took place in an open-plan office, with no method of isolating the development machine from access by other persons. This meant that other security techniques had to be used to ensure that the development was reasonably secure. Although the techniques described cannot guarantee that the development system is secure, the odds against such infiltration can be explicitly calculated.

3.4.1.3. Methods Adopted During System Development.

The basic parts of the KMDB system were designed and coded away from Edinburgh during a summer vacation at the authors' parent's home. During this time the handwritten code was stored in a secure area to which only the author had access. On return to Edinburgh the code was then typed into source files on one of the Koestler Chair's IBM XT286 (AT type) personal computers.

10 Even in the most secure systems it is always good practice to have some individual, other than the experimenter, who has access to the experimenter's work. This is especially true for projects involving a great deal of expense, or effort. These precautions are in case the experimenter is in some way unable to continue, either through ill health or some other sudden event. The trusted individual can then at least retrieve the work.

11 Standing for the rather egotistical, but certainly distinctive Konrad Morgan DataBase system (KMDB). Any security or operational features mentioned as being part of the KMDB system, and not involving some third party's software or idea (these will be identified by a trademark, registered product name, or publication reference) must be assumed to be a result of the author's own innovation and design. These innovations form part of the work conducted in the study outlined in this dissertation, and are copyrighted intellectual property.
3.4.1.3.1. Cryptography.

General Principles of Cryptography.

It is beyond the scope of this dissertation to explain cryptography in any depth\(^2\), however some basic principles are necessary since the use of encryption is vital to computer security. Encyphered text is known by the term ciphertext, while non-encyphered text is called plaintext. The best cryptographic systems make use of the difficulty in factoring large numbers, and use a large random encryption key only once. These are called holocryptic systems, and can be mathematically proven to have a fixed probability of solution by analytical methods. Other less complex systems such as simple letter substitution can, in comparison, be easily compromised by computer analysis. Holocryptic systems are obviously preferable for parapsychological use, but even the strength of the best techniques can be invalidated by improper use such as nonrandom keys, or the compromise of the physical generating system. The difference between a cipher and a code is that ciphers act upon plaintext elements of regular length, while codes act upon plaintext elements of irregular length. Codes are categorized as either one part, or two part. One part codes are where the same system is used to decrypt the ciphertext as was used to encrypt it. Two part codes, which are more secure, require a different algorithm to decipher the message. Most computer algorithms are based on ciphers rather than codes, because the computer uses bytes (8 bits) to store plaintext, making the plaintext already in regular units. Caution needs to be taken when adopting a cryptographic system, since it must be assumed that everything but the original key is known to those who are going to try and decipher it. It is for this reason that most military and commercial cipher systems use a trusted standard system such as the Data Encryption Standard (DES\(^3\)), and use a random process to determine the key. The DES system uses a variable 56 bit key to control a repeated substitution and transposition of the plaintext. This method has subsequently been found to be vulnerable to the intense factoring capabilities of new super computers.

---

2 Interested readers should consult standard texts such as Hershey and Rao Yarlagadda (1986).
Parapsychology’s Use Of Cryptography.

Parapsychological use of ciphers has been minimal, apart from the survival research conducted by some researchers (Thouless, 1948 & 1949). Thouless reported using a quasi-random 74 letter key with the Vigenere letter square for the first of his tests and a double application of the Playfair system for the other (Stevenson, Oram, & Markwick, 1989). These encryption systems are based upon alphabetic substitution or matrix transformation, and are vulnerable to symbol frequency analysis, or by knowing part of the plaintext. Once some relationships have been determined the entire plaintext version can be obtained, since the relations apply to other sets within the message. More modern encryption methods (Yu & Yu, 1989) use second order time reversal transformations as the encryption process. This system makes ciphertext which is currently unbreakable, since it has in the order of $10^{155}$ possible solutions (larger than the age of the Universe in seconds), and total independence of ciphertext data items.

Yet the computer and ciphers should be recognised a valuable tools for parapsychology. The encryption of computer digitised images could make valuable clairvoyance targets. The use of encryption methods for data storage on computer based experiments is another of the many improvements that parapsychologists should consider adopting. These uses will become especially important in the future when there is an increase in situations where Ss remotely interact with a computer controlled experiment. This would allow the Ss to access the system in highly psi conducive surroundings, and at times of their own choosing (Personal communication Morris, 1986). However such remote accesses to un-monitored systems are especially vulnerable to infiltration. The use of encryption and write once read many times (WORMS) optical discs to log data and actions could help to ensure such future research remains secure. Encryption has other advantageous uses which are unique to parapsychology. Some of the current weaknesses in experimental design can be reduced

---

14 Thouless stated he had used an extract of English literature. From the example of the text that was de-ciphered by a cryptoanalytic expert in the 1940s it is reasonable to suppose that the extracts are about death, or survival, and are from the 'classics'. Much of the classics are now stored on computer readable compact discs, so an exhaustive search for target text areas of similar sizes would not be impossible.

15 See the section in this chapter on hardware RNGs.
by the use of encryption. These possible uses will be outlined under their relevant headings in the section on security during the experiment. However a note of warning should be made at this point. Experimenters should not view an encrypted data file as having been proofed against fraud forever. A more responsible attitude is to view encrypted files as being 'protected' for a specific 'life'. The time involved determining such a 'life' will vary, but it is good practice to assume it to be half the estimated time it would take the experimenter to break the cipher from scratch, assuming they had lost the 'key'. This amount of time should give a reasonable indication of the length of time it would take a competent cryptoanalyst to break the code. Beyond this time period the file should be presumed to be no longer secure. If it is a target then the Ss should be made to respond well within the encryption life, and procedures be devised that ensure that data files are processed, analysed and the resulting data duplicated to secure sites within that time period.

The Use Of Cryptography In The KMDB System.

All files were created with the public domain American Standard Code for Information Interchange\(^{16}\) (ASCII) text editor MicroEmacs\(^{TM}\) (Version 3.8d & later 3.9e & 3.10 respectively), using its encryption option. This allows the source programs to be created and edited in the normal manner, but uses a key (29 bit) to encrypt the ASCII file when it is read from or written back to the disc. This encryption key can be of an arbitrary length, in the case of the KMDB system different 11 digit random numbers\(^{17}\) were used for each of the four initial files. The MicroEmacs method also encrypts the key while the editor is running (in RAM memory). This avoids potential security breaches from terminate and stay

---

\(^{16}\) Alternative name for the International Telegraph Alphabet No. 5.  
\(^{17}\) Generated using the hardware random number generator built into the Casio FX-450 scientific calculator. The use of random numbers as cipher keys increases the difficulty of breaking the encrypted message.
resident (TSR) programs. These semi-virus programs could be installed and sitting in memory waiting to 'grab' and record the key. This method is also recommended since if the development is on a multi-user system it avoids the risk of another user scanning the memory space to find the key. Once the user has the key they would be able to gain access to the code. The MicroEmacs encrypt routine is based on the 'Beaufort Cipher'\textsuperscript{18} and has a cipher alphabet of all the ASCII printable characters\textsuperscript{19}. This encryption system is reasonably secure. It uses a variant autokey, derived from a weighted sum of all the previous clear text and cipher text. Any cyclic behaviour is removed by a separate counter, which acts as 'salt' for the routine. Original key feedback is used preventing decryption attacks on the last part of the message, which is the weakest part in pure autokey systems. Although direct methods of attack are known for this method of encryption, they are complex, and computationally expensive\textsuperscript{20}.

3.4.1.4. Compaction Of The System.

All the work files that are created during the development phase can be combined into one file at the end of each work period. This combined file can then be moved to backing storage, and kept more securely than if it remained on the development system. It is important that the development files are deleted from the hard disc of the computer system. In deleting such files it must be remembered that computer operating systems do not actually destroy the contents of deleted files, and that destroying file involves some extra steps. In Microsoft's Disc Operating System (MS-DOS), which was used in the KMBD project deleted files are retrievable using special programs. Indeed a way of secretly storing files has been

\textsuperscript{18} DLH-POLY-86-B Cipher by Dana L. Hoggatt (c) 1986. The MicroEmacs version of this cipher has an error in the 'salting' process. The source line in module Crypt.C reading "key = key + key + cc + *bptr + salt;" should be amended to "key = key + key + (cc ^ *bptr) + salt;", and the resulting source code recompiled. This correction has been incorporated in all MicroEmacs versions above 3.10.

\textsuperscript{19} All high bit and control characters are left unaltered by the encryption process. This allows the encrypted file to be sent via standard communication systems without fear of corruption.

\textsuperscript{20} On average to decipher each separate file would require over $10^9$ separate attempts. For all the development files this would involve over five US billion detailed decryption attempts. Given that the combined length of the initial development code was 10,000 lines this would involve the intensive use of powerful computers, by cryptanalysis experts.
to delete them and then write protect that disc (Campbell, 1989). Fortunately software also exists which actually ensures that the file is destroyed, an example is Peter Norton’s WipeFile™ Utility\(^{21}\).

**KMDB System’s Use Of Compaction.**

During the development period the encrypted files were also compacted and 'garbled' using System Enhancements Associates (SEA) public domain archive program 'ARC™' (V 5.20 & 5.21, and later PKARC™ V 3.6). This process reduced the size of the development source code so it could be stored on a floppy disc and removed at the end of each day's work. These floppies were then taken home by the experimenter and made secure each evening. In addition to reducing the size of the files SEA's archive program uses its own propriety encryption system. SEA call this 'garbling' the file; this is not as sophisticated as the encryption method used by MicroEmacs. The contents of the ARC compacted\(^{22}\) file are encrypted using an exclusive (logical) 'or' between each byte of the compacted file and each byte of the operator supplied encryption key. This encryption key can be of any length, but since each of the key’s bytes are used in rotation during 'garbling\(^{23}\)', it is obvious that multi-byte keys are better. The KMDB project used 11 digit random numbers\(^{24}\) for this process. As was the case with the encryption key used with MicroEmacs these had to be committed to memory, and not recorded anywhere. These precautions were intended to make this part of the development phase as secure as possible.

\(^{21}\) This conforms to an US DOD security standard, which involves writing numeric rubbish over the file contents, and then deleting the file. A more advanced version of this method was used in the KMDB system.

\(^{22}\) SEA's ARC uses one of two methods to reduce the size of files. First is repeated-character compression, where repeated sequences of the same byte are collapsed into a three byte code sequence. The second is called 'Dynamic Lempel-Zev compression', is where variable size bit codes represent character strings.

\(^{23}\) The key is converted to upper case before this process.

\(^{24}\) Generated using the same method as used for the MicroEmacs.
3.4.1.5. Printouts And Compilations.

The problem with the precautions that have been outlined above is that at some point in the development phase files have to be decrypted into an ASCII format for compilation, and or printed out on a line printer\(^{25}\) (decryption is necessary because most commercial language compliers take ASCII format source files to produce their executable programs). These actions could invalidate all the precautions taken while the source code was being typed into the computer. To minimize this risk the KMDB project adopted the philosophy of keeping all the new systems file format definitions\(^{26}\), and other security sensitive aspects in one file. The security measures were then concentrated on that one file, while the other files could be handled with fewer precautions. This method does impose some modular design constraints on the system, such that every module of code can only access sensitive areas through specific secure routines. Since this is part of good software design anyway it should not prove a problem, unless the experimenter is unfamiliar with accepted computing practice\(^{27}\).

Unencrypted ASCII source code files were only copied on to the disc for compilation or listing purposes. At all other times these files were encrypted. However the source of low risk files only had ARC garble protection\(^{28}\), while the one high risk file was kept under the MicroEmacs and ARC encryption. All listings of the sensitive module were carefully destroyed after use by systematic hand shredding. The resulting paper waste was mixed with identically shredded 'dummy' waste from other low risk code modules. This was disposed of in one of several public waste bins. This was probably the weakest point in the security of the KMDB project. It would have been more secure to use a mechanical shredding device.

\(^{25}\) It is possible that the cables connecting the computer to its printer can be monitored remotely, either by 'bugging' or by reconstituting the electrical signals remotely (see RINT).

\(^{26}\) Knowledge of the file formats would be important not only to allow the alteration of data, but would be needed before the encryption process could be broken.

\(^{27}\) It is not within the scope of this discussion to teach such practice, and interested readers are referred to computer science texts on software design. A list of such texts are provided by Allworth (1985).

\(^{28}\) This was with the same 11 digit encryption key for all files in the archive, but this key was changed by rotation on each day of the week. This was because the number of files had increased to over 60, and it exceeded the author’s memory to recall sixty different 11 digit numbers. The object of this encryption was to remove the chance of the code being altered, not to prevent it being read, so its 'life' only needed to be until the next working day.
and burn the resulting waste. Unfortunately such facilities were not available to the author so he had to improvise. However it was felt that the effort required to monitor which waste bin was used, and then reconstitute the listing was a sufficient deterrent. No security was used for obsolete listings from the remaining sections of code. These were left as scrap paper in the laboratory.

3.4.1.6. Compilers And Languages.

During the development of experimental systems a decision has to be taken as to the computer languages which will be used. There are a wide variety of brands, and it is might be assumed that all these options are equally secure. However, this is not the case. There are different kinds of security risks which affect a system. The ones which have been considered so far have been explicit attempts by other individuals to infiltrate a system. There are also risks from the software which is used. These risks can split into intentional and unintentional aspects. As a rule the unintentional ones are the hardest to control. These include errors in coding and operation, known in the computer world as 'bugs'. All software must be assumed to have bugs in it, unpleasant though that thought may be. No matter how much the code is tested, it is impossible to test every possible aspect of operation or data path. This is why the choice of software used in developing a system is vitally important. The developer will be creating bugs in the new system, so the fewer bugs there are in the tools used to create that system the better the overall project can be. This is of great importance to scientific research, where a bug could result in the experimenter proposing new forces, or effects.

Languages.

There are basically two main language options open to the experimenter, interpreted or compiled languages. The difference between the two is basically that the interpreter reads a source file a line at a time, and then converts that source line into instructions which it then obeys. A compiler converts the entire source file into a file of instructions (called an executable file) which can then be directly run by the machine. Obviously compiled code
is much faster than interpreted code, since the translation of the source file to the instruction has already been done. compiled code is also more secure, since the interpreter keeps the source code, which can then be copied or altered by a knowledgeable person. The compiled code can also be copied, and there are products called dis-assemblers available which covert the raw instructions (which are binary numbers) into their equivalent assembler instructions. Assembler code is the set of mnemonics that represent each instruction. It is possible to derive the detailed actions of a compiled file from this process, and this should be kept in mind when using computer languages. This was not considered a major risk in the KBDB project since to derive the working of the system would require the dis-assembly and analysis of well over 550,000 bytes of executable code. It was considered unlikely that anyone would be inclined to do this, since it would involve the best part of six man-months work, by a highly skilled programmer.

The Advantages Of Using Certified Products.

The computer industry has various official bodies which issue what are called compiler validation certificates. These organizations have a series of tests which comprise an agreed international standard which a language compiler must pass before it is validated. In the UK the organisation responsible for this is the government run British Standards Institute (BSI). Experimenters should ensure that they use a language compiler which has been issued with such a certificate. This will cost more, but will have two advantages for experimental parapsychology. First the compiler will contain fewer errors, and be of a high standard. Secondly, the code that the experimenter has developed will be known to act in identical ways by other experimenters when they use a similar standard compiler. The compilers used in the development of this system were all BSI validated (ProPascal V3.0, ProFortran77 V1.1).

Some of the better interpreters do allow a protection mode which encrypts the source code while it is on the disc, and decrypts each line as it needs it. Microsoft GWBASIC, and IBM BASICA are examples of this kind of system. However the files can still be copied, and the same encryption key is used by all interpreters so the source could easily be run on another machine.

Morris (personal communication, 1989) pointed out that it would be an advantage to be able to suggest methods of protection against the possibility of disassembly attack. Although there are no currently recognised methods of protection, it would be possible to encrypt the executable file, and produce a small controlling program which decrypted, ran, and then re-encrypted the system for each experimental session.
Tools.

It is equally important that any other computer tools used are not only of a high standard, but are also internationally known, or can have their source code supplied.

Parapsychology’s Use Of Standards.

It would be an advance if the various parapsychological laboratories around the world could standardize the computing aspects of their research. This would help in their attempts at replication. The first steps towards this would require the professional bodies to set up committees to define standards, which would then be adopted by all the members of these bodies. This would also help in journal reporting, since experimental protocols could be fully described by merely stating which standard the experimenter adopted. In the creation of such standards it would be usual to have various levels or grades, so less well equipped researchers could still comply with some minimum level of security, without the expense of implementing the full standards, which might require such things as shielded rooms. These standardisation bodies could also provide validation suites of data and programs (or program specifications) to test aspects of the protocol. For example suites of test data for statistical packages would help to ensure that the results were not being affected by bugs in the analysis. Another example would be validation programs (or specifications) for testing RNGs. Without these steps experimenters cannot begin to conduct replications of computer based experiments.

3.4.1.7. Hardware.

The hardware used to develop and construct the system should also be recorded. This is in case some aspect of the hardware design has an influence on the system being created.

There are many possible influences that can be made to the target computer system before the test, both to the hardware and the software. Since most of these methods are the same or very similar to those adopted in such attempts that take place during the experiment they will be discussed in that section.
3.4.2. Security During The Experimental Period.

3.4.2.1. The Problems Of Using A Deterministic Machine.

As we have seen from the review of parapsychological use of computers, most experiments use some form of randomness in their experimental design. The digital computer is largely deterministic in its behaviour, depending entirely upon inputs and logic for its behaviour. This makes the generation of randomness a problem, and specialised subsystems are used to generate true randomness. This is as opposed to pseudo random processes, which are based on an algorithm, and are therefore totally deterministic. Any apparent randomness in such a quasi-random process is caused by the initial state of the system.

In order to discuss some of the security problems associated with the period of time during the experiment some background information about computers is necessary. The modern personal computer used in such experiments is based around what is called "The Von-Neumann Architecture". This is a concept where the instructions and data are both held in the computer’s fast internal memory or RAM. This allows the same machine to have a fixed hardware configuration and yet be able to perform any logical operation. This feature was developed by the Allies during the Second World War, for the early computing device called the Electronic Discrete Variable Computer (EDVAC). This machine was designed for three dimensional aerodynamic and shock wave simulations, and was probably used in the development of the atomic bomb.

---

31 At the end of the last century it was believed that every action in the universe was deterministic, and the apparently random behaviour of systems were the result of environmental interactions. Modern theoretical and experimental advances such as the work of Feigenbaum in Chaos theory have proposed that this may not be the case (Gleick, 1984).

32 This is excluding special processors that make use of parallelism, and at the time of writing are becoming more common. Such transputers derive their speed from executing many instructions at the same time. The more traditional architecture has a strictly serial processing capability.

33 These ideas were based upon the concepts put forward in Alan Turing’s paper ‘On computable numbers’ (Turing, 1936). In which he proposed the concept of a ‘Universal Machine’, which had instruction and data independence, combined with the ability to perform any operation.

34 The draft report on the EDVAC (De-Classified Report - March 1945). Von Neumann was ‘adviser’ on this project. This report also drew upon the early neural network publication of McCulloch and Pitts which proposed the brain as a form of ‘Universal Machine’ (McCulloch & Pitts, 1943).
This concept of the personal computer being a 'Universal Machine' (UM) is important when we look at the problems of induced artifact. The perfect UM would perform an identical operation in an identical way, no matter when or where the UM was being operated. This is implicit in the digital computer, which is largely deterministic. However, the personal computer is not a perfect UM as some parapsychological experiments have assumed (examples where this assumption is dangerous will be given later in this chapter). In general, use the small variations in performance in a computer are unnoticeable. The extra millisecond of drive access time, or the slowing of the system cycle by a nanosecond are not of importance to most users, but can be to parapsychologists. As we shall see the imperfections in the computer's imitation of an UM only becomes apparent when you take the system to the limits of its performance.

3.4.2.2. Micro Computer Components.

This section will detail the elements found in a computer system and the possible artifacts or security weaknesses that are inherent in these systems. Examples will concentrate upon the two main systems used in the KMDB study, the Amstrad 1512 and the IBM XT-286. Similar problems to those described here are equally valid for the other makes of computer systems used in parapsychology, and a summary of these will be given where necessary.

Both target machines used the IBM personal computer industry standard which uses the Intel 80x86 series of chips. They were chosen for the study because of the large amount of cheap support available for them, and because they were able to fulfill the HCI requirements of the dissertation better than any other currently available machine. The Amstrad 1512 series use the Intel 8086-2 processor (@8Mhz four wait states (WS)) and the IBM XT-286 uses the Intel 80286 processor (@6Mhz zero WS). The 80286 processor is a more advanced version of the 8086, but can run identical software, and for the purposes of this discussion they can be assumed to share similar properties (where the differences are of importance for the KMDB project this will be noted).

KMDB Security Measures.

The first level of security during the KMDB experiments was that the psi part of the study was totally covert. That is the Ss and one of the experimenters (who stood in for SG) were completely blind to the psi element in the task (although, as Morris (personal communica-
tion, 1989) has pointed out it is possible that the covert nature of the study could have been detected by some people). It can be argued that any security beyond this is unnecessary, however the security measures included in the KMDB study were not just for the security of the KMDB study, but were intended to provide a comprehensive guide to security methods.

3.4.2.3. Security Of Computer Task Environments.

There are some major problems involved with trying to run a task on a micro computer with a well known operating system (O/S) like MSDOS. These problems were minimized by the use of a covert part to the study and the use of security features. These features allow the experiment to be run with a limited degree of certainty that the data being collected will be uncontaminated by fraud. Even with the extensive security features that were built into the system the users were never left unattended with the computer system. The importance of this factor of constant observation cannot be over emphasized. The security protocols required when Ss are allowed unsupervised access to the target computer system are of a different order of magnitude to when Ss are constantly observed. The extent and complexity involved in the security for unattended Ss are highly specific to each situation. Experimenters who wish to use such experimental designs, and use the results with a reasonable certainty of their integrity would face problems that exceed the scope of this thesis.

3.4.2.4. Hardware Errors.

Under the term hardware we will cover every piece of physical equipment that is involved in the experiment. We will ignore the problems in defining the exact boundaries of this definition (such as are electrons hardware or software?), and instead look at the range of problems that can be associated with hardware malfunction. Hardware malfunctions can be induced by design flaws, such as the sinking of HMS Sheffield in the Falklands. In this case the sophisticated radar and surveillance systems were jammed by an outgoing radio phone call that used the same frequency. Alternatively hardware faults can also be due to breakdowns caused by wear and environmental influences, an example of this was in the

---

35 The security systems necessary in such work would probably become obtrusive to any favorable psychological factors involved.
2nd armed test of the Tomahawk nuclear cruise missile. In mid flight a bit of data became corrupted and the weapon disarmed itself, opened its parachute and fell harmlessly to the ground. These environmental influences occur naturally, but they can also be induced to simulate a paranormal anomaly. Excluding all these possible influences are important in the security of computer systems. However there are problems in doing this; unlike traditional engineering the design of software and hardware have not reached the stage where proven techniques can guarantee to exclude all possible malfunctions (Durham, 1987, Joseph 1988). These techniques can only be developed where the principles involved have some degree of stability. The problem is that computing is continually working at the frontiers of human knowledge, and no such stability is possible. A sobering example of this is that some of the solutions and problems outlined in this dissertation will be obsolete within a few years from the date of its writing.

Hardware Problems With The KMDB System.

Hardware based malfunctions or influences are one of the most difficult problems the KMDB system faced. The nature of the interaction of any piece of software with its supporting hardware is usually implicit. Within the KMDB system there were some implicit assumptions that had to be made. It was assumed that the target machine had the sole function of actioning the instructions issued to it by the KMBD system's logic. A potential weakness would be if specialised hardware had been inserted into one of the target machines, so it monitored and actively changed the workings of the system. This would require that the detailed workings of the KMDB system were known before the experiment, and the some components in the target system had been exchanged.

Hardware Checks.

The KMDB software checks the integrity of the system clock and all the actions of the O/S that have any active part in the tasks that the system is undertaking. However there is very little that the software can do if it installed on a specifically designed piece of hardware which will preempt the logical activities of the software and then changes the performance of the system. To guard against this it is good experimental practice to record the serial and identifying marks of each of the target computer components before each experimental session. This will guard against components being switched before the session begins. A
standard session log with pre-session checklist can help experimenters remember to record the system component serial numbers. Although a hardware exchange scenario was unlikely given the covert nature of the KMDB task, it is possible in some experimental situations. The next sections will outline the various components of the computer system item by item, and detail their respective functions, and the types of security problems associated with them.

3.4.2.5. Time

A major influence in the behaviour of a computer system is the timer, or computer clocking cycle, since it determines the interrupt behaviour of the entire system. Interrupts are of great importance since they determine what activity occurs at what time on the computer. In parapsychological experiments and theory (May, 1986; Radin, 1986) it has been proposed that Ss may be able to guess the most favorable time to initiate an action. You will recall that earlier the importance of the UM concept was stressed, and it was stated that the digital computer was only an approximate UM. We are about to discuss why such factors as the proposed minimum of a 20 millisecond psi timing capability of human Ss (Radin, 1986) are probably gross underestimates, because the experimenter assumed he was using an UM. Before we can do that we have to summarize the role of interrupts on a personal computer (PC).

3.4.2.5.1. Review Of Role Of Hardware And Software Interrupts.

In the idealised UM machine the processor is totally dedicated to running only the users program instructions. However this not the case; the average computer system is composed of many separate and independently active parts. The processor has to ensure that all the activities remain as a coordinated whole. To do this the processor has to be kept informed of the various system functions as they progress. It could do this by a fixed routine which it checked regularly to see if events had occurred. This would ensure that parts of the system never had to wait, but it would be wasteful of processor time. Instead when a component needs the processor to do something, such as read a buffer, or update the system clock, it sets what is called an interrupt flag or bit. When this is set the processor temporarily stops
doing whatever it was doing at that time, and diverts its attention to servicing the interrupt. When it has completed that the processor will resume its previous activity. There are usually different levels or priorities of interrupts on computer systems. The Amstrad 1512 has 9 levels of hardware interrupt, and the 80286 has 16 levels. The highest is reserved for system calamities like memory failures, or power loss. Then come aspects like servicing timers, reading full keyboard buffers or serial ports. Lower down the priority list are things like updating the real time clock, and accessing disc drives. None of these activities can occur without the processor servicing these interrupts, unless it is done explicitly under the control of the users software. So when someone presses a key on the computer keyboard (and fills the keyboard buffer), the processor will be interrupted to empty the keyboard buffer. However if the system received another interrupt at the same time, which was of a higher priority, such as the timer, it would service that first, and then the keyboard, before it finally resumed its previous operations. The resolution of these events is the time it takes to set the processor’s hardware interrupt flag.

Obviously these events could be detrimental to some time critical aspect of the users code. An example of this would where a user is monitoring or controlling some real time event with a computer. For these reasons it is possible to disable the hardware interrupts by a special instruction from within the user’s software (called CLI on the 80x86 processor). Most time critical uses of a computer should be done with hardware interrupts disabled, for obvious reasons. The 80x86 family also provides 256 possible software interrupts, which is how the user’s program accesses the low level services of the computer. This is called the INT instruction, and is supplied with a specific number, which identifies the interrupt routine which will be serviced. The processor treats these software interrupts in a similar way to hardware ones. On the 80x86 processor, a vector of 256 code start addresses exist at the low end of the computer’s memory (called the Interrupt Descriptor Table (IDT)). When an interrupt occurs the processor multiplies the user supplied number by four and looks at the resulting memory location for the memory address of the software routine that handles that interrupt. It then executes that routine (while saving all the relative working environment of

---

6 On the DEC PDP11 series of computers (which have been widely used in parapsychology), most models had four levels of interrupt, while the LSI-11 (the micro version) had only 1, and allowed any I/O device to interrupt the CPU, thus suspending program execution (DEC, 1984).
the interrupted process). If this was being done by an UM, all the times taken to perform these operations could be specified and would always remain constant. Intel and Microsoft do provide such timings, in the form of the number of clock counts required for each operation, and these can be converted to nanoseconds \((10^{-9} \text{ of a sec})\). This is accomplished by dividing one microsecond by the number of megahertz (Mhz) at which the processor is running. So for the Amstrad 1512 running at eight Mhz the clock cycle is 125 ns \((1.25\times10^{-6} \text{ of a second})\). The software interrupt instruction INT takes 51 clock cycles on the 8086, so each software interrupt takes a minimum of \(6.375\times10^{-5}\) of a second. However both Intel and Microsoft provide the following exclusion clause to such specifications.

The clock counts are for best-case timings. Actual timings vary depending upon wait states, alignment of the instruction, the status of the pre-fetch queue, and other factors. (Microsoft, 1987).

To illustrate the problems which could arise if the experimenter assumed that this process was being handled by an UM, let us imagine that an experimenter has created a parapsychology experiment. In this experiment a S is supposed to press a key to get a random number from an external live RNG. Since this is a time critical process we will assume the experimenter has disabled hardware interrupts. Traditionally parapsychologists have assumed that such an influence on a RNG could either be done by influencing the RNG by some remote action (PK), or by selecting the best time to press the key (IDS). We can further assume that the system gives feedback on the screen for each of these trials. This feedback would be where observational theorists believe any influence occurs, since they propose that the S’s observation of the result collapses the state vector associated with the RNG’s output. This is a problem since from the systems point of view the moment the processor reads the serial port input signal the RNGs output is defined as a one or a zero and the state vectors will have collapsed.

Now we will look at the processes involved from where the S presses a key on the keyboard. The first action is that the keyboard will use firmware routines in keyboard ROM to decode
the keypress signal and store the data as a byte of information in the keyboard's own RAM memory. This will take approximately a few hundred ns\(^37\). The resulting character will be kept there until the user's software reads from the keyboard (assuming hardware interrupts have been disabled). For the sake of ease of explanation we will assume that the code is attempting to read the keyboard in a short loop (or set of repeating instructions). We can assume these loops take a minimum of 17 clock cycles (the length of time required to execute the loop instruction). The actual transfer from the keyboard will be done in under two milliseconds (the keyboard aborts if it takes any longer), and the processor will then have detected that the S has commenced a trial.

The processor has to read each instruction in from RAM before it can obey it. A further delay is inherent here since the processor cannot access these instructions at any time. The RAM has what is termed a refresh cycle, which is needed or the contents of memory would be lost as the electrical potential decays. This refresh cycle causes the processor to have a time delay each time it requires its next instruction. This is approximately 500ns on the Amstrad 1512. There will be a further delay as the serial port is read, and its contents moved to RAM. We need not be specific on this operation since hopefully by now it has become clear there is a time lag involved in the whole process.

The final stage is that of displaying the result to the screen. This involves writing to part of memory called video RAM, and this causes a whole series of timing problems. The video RAM and processor timing on the Amstrad (and most microcomputers) are derived from separate reference frequencies\(^38\); therefore any displays must be synchronized by the VDU controller. This is done by adding an arbitrary number of CPU wait states. The delay between the processor being ready to update the display and it actually being able to will be dependent on the VDU controller and what other processes are being executed (such as Direct Memory Access (DMA), which is a process by which memory can be updated without processor

---

\(^37\) In all these timing cases it is impossible to be absolutely precise, since the event will be within the tolerance of the internal timing circuits, which normally have a 50% duty cycle (which means that the regularity of the signal is not guaranteed). These internal timers are used for synchronization, and are not intended for uses requiring a fixed frequency, or extreme timing.

\(^38\) A reference frequency is a regular signal which a process can use as a basis to determine the passage of time.
intervention.). On the Amstrad this delay will range from 12 to 46 (125ns) wait states, the final number being determined by the number of screen refresh cycles since the machine was switched on, the current accuracy of the governing clock cycles, and what actions the computer has previously executed since being turned on. This maximum delay of 46 wait states will occur approximately every 63.7 milliseconds. Finally, it should be noted that the S’s next trial cannot be started until the processor has written the result to the screen. The length of each trial will therefore be varying as a result of numerous independent and interdependent processes.

This brief scenario should have highlighted a number of dangerous assumptions that could be made if the system was assumed to be an UM.

First there is variation in the time delay between the S requesting a trial and the RNG being read on one occasion and on the very next trial. Parapsychology hypothesises that something associated with the S is interacting with targets (otherwise why use Ss?), and that this factor is highly precise both in terms of its locality, and temporally\(^39\). Since it is apparently presumed that the effect can be directed to within electronic time resolutions, it is remarkable that parapsychologists so rarely comment upon the variable temporal delays involved in each trial. In effect Ss have to endure a different experimental situation on every trial; when this is considered is it any wonder there are problems in replication? In effect the variability inherent in the hardware makes any exact replication impossible. A difference measured in nanoseconds in the time the machine has been turned on would make the test different, just due to the minute timing variations. These kinds of differences between each trial would also make any learning very difficult, however the next factor to be discussed makes these problems seem trivial.

The second factor is more serious. The complexity of the task Ss are being asked to attempt in a standard computerised RNG test can be shown to be beyond the capabilities of any

\(^{39}\) The numerous studies which have generated control runs at very close time intervals to trials must have assumed this (see the PsiLab experimental protocols). Morris (personal communication, 1989) has pointed out that if psi has transtemporal properties this observation would be questionable.
known human biological (cognitive) process. The fastest human neural discharges (myelated) achieve speeds of 0.4 milliseconds \(^{40}\) (single cell proprioception; somatic motor), the fastest reflex (the knee jerk), is within the range 19-24 milliseconds, and in contrast a typical brain cell’s fastest discharge speed is around 2 milliseconds (Ganong, 1973 p67). These timings represent the minimum human inter-nerve cell communication times. These specifications will be taken to approximate the minimum times in which biologically based cognitive awareness can take place \(^{41}\). It is important to note that these times do not include any account of the delays involved in the sensory input that is assumed to precede 'awareness', or any form of reaction to that 'awareness'. To review the effects these timing problems produce, we will briefly summarize the problems each of the major current RNG influence theories experience if a real time biological brain is assumed. Then we will discuss the various alternative explanations that could explain how such a process could be controlled.

3.4.2.5.2. Problems For Parapsychology’s Theoretical Models Of RNG Influence (Assuming A Realtime Biological Brain).

IDS.

Proponents of the IDS model have assumed that the S was timing their selection to the resolution of the system clock, we have seen that this is an invalid assumption. The actual resolution which a S is being asked to 'guess' within may be as small as a nanosecond, and

---

\(^{0}\) The author is grateful to Dr(s) Burton (Glasgow University), Macleod (Edinburgh University), and the senior librarian at Edinburgh's medical library for their help in determining the timings involved in these fastest human biological processes.

\(^{1}\) In response to these timings Morris (Personal communication, 1989) has pointed out that some parapsychologist's might argue that chemical or sub-atomic activities within the human organism operate at higher speeds than those used in these examples. The author admits that if cognitive activity was derived from sub-atomic activity, or was non-biological/physical, these timing objections would be invalid. However, it is argued there is some evidence to support the hypothesis that cellular electro-chemical activity is directly related to both physiological and cognitive activity. Therefore the author proposes that the realtime biological model should be falsified before investigating the alternatives.
is certainly much smaller than the 20 milliseconds quoted by IDS proponents (Radin, 1986). Since during some trials the human biological responses might be able to cope, we would expect hits to be rare, occurring only when the processes involved were slow enough for the brain to monitor. A realtime biological IDS model would predict that experiments using microcomputers which use processors faster than 33 MhZ\(^2\) should reveal no psi effects.

**PK.**

The direct physical force proponents (PK) have the same sorts of problems as those faced in the IDS model. In their scenario the Ss must be able to 'know' when the RNG has to be influenced, and not influenced\(^43\) to within the same kinds of time resolutions. However the IDS subjects only have to 'know' about the RNG output for the trial, for them the control data does not have to be monitored. In contrast subjects in the PK paradigm have to switch the effect off and on at a very specific time period. This is probably not possible for a realtime biological process on computer systems with fast clock speeds.

**OT.**

Observational theorists have a different set of problems. Their theory can cope with the timing resolution problems, if they can get around the fact that the computer will have collapsed the random number's state vector the instant it defines it to be a one or a zero.

**Timing Conclusions.**

These timing and complexity issues in modern microcomputers have raised some problems which are beyond the scope of this chapter to discuss, devoted as it is to security and artifact\(^44\). Issues such as the assumption that real time biological processes are involved in psi events are very much open to question, however for this time we can conclude that even allowing for any errors in approximation that might have been involved in the 500ns RAM monitoring example the problems raised are still valid. Modern processors have much

\(^{42}\) A 33 MhZ clock would have a 30.3ns clock cycle, and could perform the interrupt instruction in 1.5ms (well inside the 20ms IDS 'window').

\(^{43}\) Assuming 'simultaneous' control runs are used.

\(^{44}\) An example of the timing problems involved assuming a realtime biological brain are provided in appendix 14.
higher speeds than 500ns (15ns is not unknown), and there would be more pieces of equipment to monitor than those in one megabyte of RAM. The experimenter must therefore be aware that it is hard to make definitive assumptions about the processes involved in a RNG-PK study. It was for this reason that terms such as Function Alteration Activity (FAA), and Function Alteration Potential (FAP) were adopted by the author.

3.4.2.5.3. Timing And Recording Activity.

Recording the actions that have taken place on the system are important if the experimenter hopes to correlate some aspect of the Ss behaviour with any FAA that is going on in the system.

3.4.2.5.3.1. Time Based components of a computer system.

Real Time Clock & Parameter RAM (Battery Backed).

The clock accessed by the operating system to record the actual time can be derived from a battery backed quartz oscillator time keeper. These usually have an accuracy of 0.1%, providing that the batteries are not exhausted. This battery backed clock and RAM are used to provide the date and time along with various configuration information that the system can retain when the power is lost. It is important to check that the batteries do not become exhausted; otherwise the operating system default values will be used.

3.4.2.5.3.2. Dangers of using the timer on a computer system.

It is often tempting to assume that the timers provided by the operating system actually provide the resolution they imply (another of the UM problems). This may be a problem, particularly if you are using the computer to generate randomness via some process associated with the clock. In common with other operating systems MSDOS provides a system clock which can be referenced by the users in the form HH:MM:SS:TH (where HH is the hours, MM is the minutes, SS is the seconds, and TH are tenths and hundredths of second respectively). Unfortunately these minimum resolution values are in fact misleading. The computer actually only updates its use of the clock on the basis of an interrupt
signal (based on the channel one update). This results in the highest resolution available to user being approximately + or - 15~18 milliseconds.

It is possible to change the rate of the timer responsible for this (channel 1), by software patches, but these are unreliable, and have undesirable side effects. The experimenter either has to live with this reduced resolution, or implement a hardware 'patch'. This reduced resolution may be important if the clock is being used to generate the seeds for a PRNG.

3.4.2.6. Processing

The computer system uses electrically controlled logic devices to perform its operations. In modern computer systems these logical devices are constructed from silicon derived compounds, and involve very large scale integration (VLSI) of their components in devices known as semiconductors. The major chips on a microcomputer are the Processor (80x86), Direct Memory Access (DMA) controller, Maths co-processor (if present), and the Basic Input Output System (BIOS) ROM which is the lowest level of the operating system.

In order to discuss the various security problems which are associated with the use of these electronic components we must briefly discuss their construction.

3.4.2.6.1. Semiconductors

A semiconductor is a substance with resistance between conductor and insulator. This resistance is determined by the number of electrons that occupy the valence band (or energy

45 Depending on the various factors outlined in the section on UM.

46 Such as speeding up the processor beyond its makers specification, or in the case of the so called 'Norton fix', the extra resolution is lost during long interrupt service routines.

47 In MSDOS systems this can be done by using an Analog to Digital converter channel to read a one second alternating voltage (between one an zero volts). By reading the input from this channel the program can calibrate the exact time by a division of the voltage by 1, providing it reads the channel at least once a second. The user also has to make allowance for the slight voltage loss produced by the PC measuring the voltage.
level) of the substance’s atomic structure. The largest possible number of electrons is eight, which is an excellent insulator, since it is reluctant to exchange electrons with other atoms. Examples of atoms with eight valence electrons are the inert gases such as helium, argon, and neon. In contrast atomic structures with one or two electrons make excellent electrical conductors, examples being silver, gold or copper. This is why they are used so frequently in electrical components. Good semiconductors on the other hand have four electrons.\(^8\) When energy is applied to these atoms electrons become dislodged from the valence band and move away from the atom. This is an electrical current.\(^9\) It is possible to create a co-valent bond between atomic structures with different valences, so one or more structures bond together. In the process of covalence these structures receive and donate electrons, which in turn make the resulting semiconductor carry a positive or negative electrical charge, called p-type\(^{50}\) and n-type\(^{51}\) semiconductors respectively. Such semiconductor devices can be used to restrict the flow of electrons (or electrical current) in a specific direction. This forms the basis of switching devices that make up gating circuits, which in turn can be made to perform logical operations. Unfortunately it is beyond the scope of this dissertation to cover solid state physics and chip design in any greater depth. Interested readers should refer to standard texts on the subject. Modern chip construction techniques have enabled such integrated circuits to function at levels and dimensions that approach the

---

\(^8\) This can also be combinations of elements with three and five valence electrons respectively in the use of so called ‘doped’ semiconductors, where another substance is mixed with the semiconductor to change its electrical behaviour.

\(^9\) The flow of electrons is called the current.

\(^{50}\) When a covalence occurs between silicon and a dopant (such as aluminium, or gallium) the resulting semiconductor material can be made to have a so called positive valence ‘hole’ that will readily accept electrons. It should be noted that these ‘holes’ are non-physical concepts created for the convenience of solid state physicists. In reality the actual charge carriers are electrons travelling in the opposite direction to the ‘holes’. This concept often causes confusion.

\(^{51}\) When the dopant is phosphorus the resulting covalence produces a surplus of electrons.
limits of quantum determinancy. At such levels the devices become extremely sensitive to external influences, and it these external influences which are of direct interest to the security minded parapsychologist.

The external influences which effect semiconductors can take several forms. The main type that is of interest to this discussion are those that induce, or increase the amount of noise present in an electrical system. We will discuss these under the term of interference. Thermal noise is caused by the Brownian motion of sub-atomic particles and is the basic source of most random noise effects. There are also non-random noise effects, and these will be discussed later. As more energy is put into a system so the rate and degree of motion is increased in the particles that make up that substance. In a computer’s components such noise will be inherent, due to the electrical power being supplied to them; however any external system that supplies energy to the component can increase this noise to the point at which it will interfere with the component’s functioning. These include natural sources such as lightning, or sunspots, and artificial sources such as radiation, electrical discharges or sparks. Electric motors produce such sparking in the discharge between the motor brushes and the commutator. Internal combustion engine spark plugs and defective light switches produce them when the circuit is made or broken. These discharges can induce electrical noise in the computer through the mains power lines, or even (in the cases of effects like static, and magnetism) through the atmosphere.

The most well known noise effect that damages computers is the electromagnetic pulse (EMP), which is one of the by products of the detonation of a nuclear weapon. Fortunately this kind of disruption has been infrequent, and it was only discovered from the effects of test detonations. However the EMP shows several characteristics of more common noise...

---

52 Intel’s current microcomputer processing chip 80486 (launched April 1989) has modified the VVLSI CMOS (Complementary Metal Oxide Semiconductor) technology used in the previous processor (80386) and uses data pathways one micron across. Narrower data lines could result in the electrons being subject to Heisenberg’s uncertainty principle: “The product of the uncertainty in energy multiplied by the uncertainty in time must be greater than a particular small number” (Shallis, 1982 p78).

53 Named after Robert Brown who discovered it in microscopic pollen particles in 1827. Brownian motion is the random motion of atomic particles as they make contact with other moving atomic particles. The motion increases with the energy input to the system containing the particles, and is thought to only cease at absolute zero, or in a state of total system entropy (which are equivalent).
effects, and is therefore a good initial example of induced electrical noise. EMP is an induced electromagnetic pulse resulting from the massive ionization of the earth's atmosphere by the radiation produced by the atomic fission inside a nuclear weapon. The resulting pulse, which can be as large as 50,000 volts per meter, is induced into any sensitive electrical conductor within a 1,000 kilometers of the medium yield weapon. The effect of this EMP on a semiconductor is to overload every data path on the component, and this burns out (or shorts) the semiconductor, making it useless.

Although the EMP is an extreme example it does provide some help with regard to the effects produced and the precautions that have been found to be immune against induced electrical noise. The military have found that semiconductors, and other highly sensitive equipment are completely destroyed by an induced electrical signal of the strength of the EMP. Data held on magnetic media are also destroyed by the induced charge, which erases any stored bits. However the older types of electrical equipment (such as electron tubes) have been found to be less affected, and metal shielding around the device has been found to reduce the effect. Mechanical and optical devices have also been found to be immune to the pulse effect (although they need to be guarded against the blast).

The more normal effects induced noise can produce range from destroying or shortening of the working life of the equipment, to induced anomalies. It is at this point that we have two different computer security problems. The first is that the experimenter wishes to avoid the damage and loss of either recorded data or expensive equipment through induced noise, and the second is to avoid induced anomalies. Fortunately the precautions for both are similar, and can be covered at the same time.

3.4.2.6.2. Induced Anomalies.

These can range from inducing false signals, and inhibiting real signals, to altering the speed or function of a circuit. These effects can be permanent, if the physical properties of the

---

4 To guard against possible damage by magnetic fields, or other destructive influences, vital data should be duplicated at secure remote sites.

5 One of the major reasons that fibre optic cables have been so rapidly introduced into national telephone systems is that they are immune to the EMP, and so would allow the military to maintain their communications in the event of a tactical nuclear exchange.
system have been permanently altered (called a 'hard' fault), or temporary, if they are merely induced in the system for the duration of an external environmental factor (called a 'soft' fault).

3.4.2.6.2.1. The Associated Implicit Environment (AIE).

A computer system has what we have termed an associated implicit environment (AIE), which includes the total environment within which the system exists. During an experiment the AIE includes both the experimenter and the S (see figure), and anything or one which comes into contact with the system.

Some Of The More Basic Areas Of The AIE Are:

![Diagram of AIE](image)

Figure 68. The AIE.

Previous History Of The System.

As we have already discussed (under the UM problems), and will discuss further in later sections (see the section on Operating Systems), the length of time the system has been turned on, and the operations it has previously performed have a direct bearing on the operational performance of the system. For example if the system has been used to print some documents it will have installed memory resident printer drivers. These will alter the amount of free memory, and the start address of the available memory. If an experimental system was run once after the printer driver had been installed, and once when it had not, then the memory locations used by the program when it was operating during these two experimental sessions would be different. Such small differences would not be noticeable
for the subject, or experimenter, but would result in minute differences in the machine’s behaviour. This would mean that the experimenter was requiring the subject to complete a slightly different task in the two sessions.

**Interactions With Users Of The System.**

The way in which the user behaves, and the sequence of operations that they carry out are of importance in determining the functioning of the system. These cover all the people who come into contact with the system, making this one of the most difficult areas to control, but also the most important. For details on these kinds of interactions see the author’s AHCI paper in appendix 1 (Morgan, 1987).

**Atmospheric Pressure And Composition**

The components of the system will behave differently at different altitudes, and in different kinds of atmospheric environments. Computers have a limited tolerance to atmospheric pressure. For instance Apple computer states that their Apple Mac II series will not operate above 10,000 feet above sea level (Apple Computer Inc., 1988). Apart from obvious effects like cigarette smoke, atmospheric conditions such as gases can also influence the behaviour of computer systems. Dust and other substances are harmful, and must be excluded. They can obscure the magnetic surfaces of data recording equipment, and block up ventilation holes, thereby inducing overheating of the computer system.

**Humidity Or Moisture**

Obviously the humidity of the surrounding atmosphere is of vital importance to the computer system, since beyond certain levels it will cause damage and malfunctions. For example the IBM XT286 system can operate within the 8-80% humidity (IBM, 1984, & 1988), and the Apple Macintosh can tolerate between 20-95% humidity (Apple Computer Inc., 1988). In contrast the old DEC PDP-11 (often used in parapsychological research in the 1970s), could tolerate a range of 10-95% humidity (DEC, 1984).

It is also important to guard against the effects of liquids, even if this means Ss are not allowed to take cups of coffee near the experimental equipment.
Temperature

Most microcomputers have a wide ambient range of temperatures which they will operate within, but it must be remembered that saying a system will operate at both 10c and 40c does not mean that the performance of the system at both temperatures will be identical. As an extreme example, above a certain temperature (60c for the old DEC PDP-11 (DEC, 1984)) due to the heat induced expansion of the recording surface, read and write operations to backing storage become unreliable. As well as increased wear, temperature variations will result in minute changes in performance. These differences may be unnoticeable to the user, but can be quite large in electrical terms. Experimenters must ensure that if they want identical situations for each S, they must provide similar temperatures. That includes the temperatures of the components. Temperature differences can cause hardware random number generators to generate biased outputs. Very high temperatures or very low ones cause electrical components to react unpredictably. It is the experimenter’s responsibility to ensure that a fraudulent S could not disturb the AIE by temperature variations, such as by placing substances like dry ice on equipment during an experimental session.

Power Supply Artifacts (AC-line Transients).

There are at least two parts to the power supply problems; long term variations in voltage, and very short variations. The problems associated with the last group are relatively easy to solve. Power supplies can be filtered to remove what are called spikes and surges (transients). However the mains supply has longer term variations both in the rate (Hz), and voltage, due to such factors as variations in regional demand, and production. These variations actually follow a regular daily pattern and follow national electrical demands. Most power suppliers guarantee their supply to within certain parameters. The experimenter must ensure that his experimental results will be unaffected by minor changes or regular patterns in the mains. This can be done by noting the times of the experimental sessions, and by actually monitoring the mains frequencies. The use of an old electric clock which derives its timing from the mains frequencies (Hz) can allow the experimenter to determine mains frequency variations, and monitoring the voltage directly can help control for these potential artifacts.
Local Motor Frequency (Mechanical Vibrations).

The activity of electrical motors in the vicinity of the system can induce electrical artifacts, by sparking discharge (we will cover resonance effects separately). Various forms of shielding can be adopted to counter these, some as simple as metal cooking foil, or chickenwire. These effects are covered in greater depth in other sections.

Radio Frequency Transmissions.

Specific wavebands of radio signal (which are part of the electromagnetic spectrum) will interfere with electronic components (for details see the section on interference). These signals induce electrical noise in the same way as other electrical fields. Fortunately systems can be shielded against these effects by use of similar shielding as we mentioned in the previous section.

Electric And Magnetic Fields.

As we have already discussed, electromagnetic radiation can induce unwanted electric signals in electrical components. We will cover the actual methods that this employs in the section on interference or 'jamming'.

Shock And Vibration Susceptibility.

Computer components are connected via electrically conductive paths. On some older, or non mass produced circuit boards these connections are made by a technique known as soldering. Violent shocks, and vibrations can weaken or break these connections causing the signals using these connections to either fail, or become unpredictable. The mechanical parts of computer systems such as disc drives and cathode ray guns are of course susceptible to damage by violent shock.

Acoustic Susceptibility - Audible-Frequency Acoustic Artifacts, And Ultrasonic Radiation.

So far in the discussion we have ignored the effects of sounds and resonances on computer systems. For the sake of simplicity the effects can be split in to those that induce a sympathetic vibration in the large scale components, and those that induce an electrical
signal by a piezoelectric\textsuperscript{56} deformation resonance within the semiconductor material. The former effect will have similar results to the vibration effects we discussed above, while the latter will cause effects identical to those produced by electromagnetic radiation. They may range from 'soft' (temporary) errors to the devastating EMP phenomena discussed earlier. These sound signals can be either audible or in the ultrasonic ranges, and have the ability to be directionally targeted. Ultrasonics can travel through solids, gases and in the case of some specialised wave forms can travel along surfaces (Rayleigh wave). Sound is defined in terms of its pitch (frequency), quality (mixture of frequencies), and intensity (or amplitude). The intensity is the rate at which the sound wave transmits energy. Using these variables the sound wave could be specific enough to target a specific electronic component. The discussion of interference will be more specific as to the effects that can be produced. Obviously, ultrasonics could be used as an offensive, or defensive weapon, against an opponent who relied on high technology equipment\textsuperscript{57}.

**Pulsed Infrared Radiation**

Another effect which is often ignored is that of pulsed energy in the form of infrared radiation. The effects of this would be to increase the energy (not necessarily heat) of the target component. This could induce a typical electromagnetic noise effect.

**Ionizing Radiations.**

These induce an electromagnetic voltage in semiconductors, the power of which is determined by that of the charged particles (ions) that result from the process of radiation. Metal and earth shielding will help reduce this effect, however at present there is no known way to exclude these effects.

\textbf{3.4.2.6.2.2. Interference or Jamming of Signals.}

In order to know and be able to prevent the process of interference it is necessary to discuss the principles involved in disrupting signals. These principles have a wide applicability and

\textsuperscript{56} Piezoelectricity is a voltage induced by certain structures (like quartz) when they have stress, or pressure applied to them. It is caused by the displacement of charged atoms along the substance's structure.

\textsuperscript{57} Certain frequencies and intensities would also be destructive to biological systems.
cover not just semiconductor devices, but any electrical signals traveling through any medium (including radio waves).

First we must recall that from chapter one we discussed that all electrons have a wave-particle duality. This means that we can treat electrons as either, and for the processes of inducing interference, by whatever means, it is convenient to assume that electrons are waves, or have wavelike properties. Electrical currents and signals can therefore be assumed to be a wave of electrons. In its simplest terms, interference is when two wave forms of the same type combine. When this occurs it changes the original waveform's amplitude or frequency. Obviously if the wave is being used to carry information, as in an electrical field, this will change the nature of the signal. There are many types of interference or 'jamming'. Two of the most basic types which are of interest to us are called destructive and constructive, both of which are capable of disrupting the information carried in a wave form. Destructive interference is when two waves of the same frequency and type have 'crests' and 'troughs' which are exactly complementary. This results in both waves completely cancelling each other out\(^8\). Constructive interference is when two wave forms have 'crests' which coincide. This results in the new wave form with an amplitude equal to the sum of the combined wave amplitudes. In order to maintain these changes both wave forms must be steady. In the case of induced electrical noise this is very unlikely, so the resulting wave interference will not be coherent (steady or long lasting). The two main examples of constructive and destructive interference are the methods that can produce either false signals, or inhibit real signals, in electrical equipment. For example one could induce a false signal in an electrical communications line so a computer would interpret that false signal as valid information.

**Cooling Devices.**

These jamming effects may have potential beyond simple data corruption. Some sophisticated electrical scientific equipment (such as lasers, and high performance circuits) use something called the Peltier\(^9\) cooling system. This system is based upon an effect which

---

\(^8\) This is misleading, since the wave forms cannot 'destroy' one another (the 1st law of thermodynamics). The energy involved is either transmitted or absorbed by the medium. In the case of electrical components this could result in them burning out, or shorting.

\(^9\) Named after its discoverer French physicist Jean Peltier who discovered it in 1834.
occurs when two different substances are joined and an electrical current flows through the join. Depending upon the direction of the flow of electrons, heat is either emitted or absorbed at the junction of the two materials. By effecting this cooling process on could alter the system’s AIE so a false anomaly was created.

Solar Cells.

Devices which use the Seebeck\textsuperscript{60} effect are also likely to be affected by such induced electrical interference. These systems make use of the electrical current which is induced in an open circuit when two different temperatures exist at either end of conducting material. Some solar cells operate in this manner, and interference in this context would result in either a surge, or drop in expected current. This could easily induce any systems dependent upon this power source to malfunction.

3.4.2.6.2.3. Preventive Measures.

Some standard preventive measures that have been widely adopted within parapsychology, are to install the computer system within an electromagnetically shielded room, and place a commercial mains power filer on the electrical power that supplies the computer and noise generating systems. Many examples of this were given in chapter two. These measures will exclude specific wavelengths of the electromagnetic spectrum, and any alternating current power transients (spikes, and surges) external to the experimental area. These kinds of precautions are common, and we will therefore not discuss them, except to note that such precautions have the effect of creating a boundary around an AIE. This boundary only excludes induced effects from outside this defined boundary, and not from induced effects from within that boundary. For example the experimenter must ensure that the environment within the laboratory is free from the effects we have discussed. Power transients can still be induced between a mains filter and the target system by such equipment as electric kettles, light switches, or even the action of some parts of the experimental equipment.

If shielding is used it should be tested with all the electromagnetic frequencies that could induce an artifact. External interference from such sources as ultrasonic beams, are more

\textsuperscript{60} Named after its discoverer Thomas Seebeck who discovered it in 1822.
difficult to exclude. It may be more cost efficient to detect their presence than to try and exclude them.

Not all experimenters will have access to such sophisticated defences. Since detectors and shielding are expensive it may be beyond the financial resources, and expertise of the experimenter to set up such protection around his experiment. In an attempt to produce a cheap method of general purpose electromagnetic and ultrasonic detection the author developed a cheap, but reasonably sensitive methodology for the final series of KMDB studies. In this final series the Ss were aware of the psi component involved in the task. For this study the experimenter deliberately set up an experimental situation which would not have the benefit of either shielding, or mains filters. This was to try to demonstrate how an experimenter could still try to detect many artifacts by the use of some of the principles which have been outlined in this chapter.

The ideal electromagnetic and ultrasonic detector should be cheap, portable, unobtrusive, and be able to detect most influences likely to disrupt a computer’s AIE. A simple integrated circuit will respond to the same electromagnetic effects as the target computer system, and a quartz crystal will be highly sensitive to ultrasonic influences (piezoelectric effects). Modern quartz timepieces have both these components and are relatively cheap. However they are not sensitive to physical shocks, mains power spikes, vibrations or temperature changes. These still have to be excluded by the vigilance of the experimenter. A series of cheap quartz timepieces can be distributed around the experimental area (including the machine) after being synchronized to a mechanical chronometer. During experimental periods the mechanical chronometer should be calibrated to a reliable external source (such as an atomic clock) at regular periods. The times recorded by these devices can then be compared before and after each session. Any gross differences in time keeping or the functioning of the quartz watches should indicate that the AIE had been violated.

---

1 A unique method of detecting some mains variations was developed for the KMDB system. However it was not very efficient since it could be affected by other factors (see appendix 13).

2 The mechanical chronometer will be the most expensive part of this detection equipment. A Rolex Oyster 'GMT-Master' (Officially Certified Superlative Chronometer) was used for the KMDB project, and found to be more than adequate (see chapter ten for details).
3.4.2.7. Input and Output.

The moment a computer system starts communicating with the environment it loses its most precious aspect of security, isolation. The search for methods of secure communications is an important and growing area of computer science. Computer communications allow computer systems to exchange information with external sources. These external sources can be human or another computer. Such communications take the form either of a live connection (in real time to the target computer), or of an electronic message which will be received by another computer system. These electronic messages can either be directed to an individual (called electronic mail), or be intended as instructions for the computer system (examples are file transfer requests). Both live connection, and electronic mail have the potential for causing security breaches. We will not make a detailed attempt to overview computer communications, or 'hacking', both are beyond the scope of this dissertation. Instead we will concentrate on the threats which might be relevant during experimental work. This discussion will examine two aspects; first a review of the methods available for the secure coordination of experimental design via electronic mail (which is something parapsychologists should use more often). Second a review of the security risks involved in allowing Ss and remote experimenters to have access to a computer system. Both of these reviews will exclude a number of areas within communications security, but these are either outside the remit of a review on experimental security, or could endanger the national security of the United Kingdom.

3.4.2.7.1. Outgoing Communication.

This section will concentrate on two main types of these communication networks, which will be referred to as Local Area Networks (LANs), and Wide Area Networks (WANs) respectively. For the sake of simplicity we will include international communications networks under the latter heading. It should be much easier to maintain security for use of the LANs rather than the WANs, simply because the experimenter has access to all the

63 The term 'Hacking' originally referred to computer programming. However it has become associated with accessing computers remotely, and it is that usage which has been adopted throughout this chapter.

64 Interested readers are referred to the various 'Hacker's Handbooks' (Cornwell, 1988).
components involved. Data lines that have sensitive information transmitted through them should be clearly exposed, and be regularly checked for small listening devices. Having the lines clearly exposed allows them to be easily observed, and therefore makes them more secure. Sensitive messages and data should be encrypted using a cipher system that has a 'life' that exceeds the message's valid use. Alternatives to bugging the transmission line are the insertion of false messages, or the jamming of the signal (see interference). The use of encryption systems and a pre-designated identity number for each message between the legitimate group members can reduce the danger of any false messages being sent. The encryption keys should be pre-set. For example an agreed entry point in a table of random numbers, where each consecutive 11 digits becomes the next key, will be ideal, so long as it is only known to the people involved, and is not written down. This kind of system has the added benefit of having an implicit numbering system for each message. This will make an extra message, or a jammed (removed) message immediately obvious. The only danger is that if the system was revealed in any way it would no longer be usable. Passwords have a similar problem. They should be changed frequently, and should not be easily guessed. For instance typical passwords are: Help, Test, Tester, System, Sysman, Root, Engineer, Ops, Secret, Demo, Sex. First names are to be avoided, as are immediate family’s names, and words from central interests users may have. Finally repeating numbers are not recommended. H.R.H. Prince Phillip is known to have used the password '22222222' on the prestel system, as a result he got 'hacked' (Cornwall, 1988 p84). If the random number table method is not possible for passwords it is best to use a fixed repeating sequence system, so long as it is not written down.

Access to larger external systems face similar problems to those we have outlined above (and can use similar solutions), except messages are passing through many more sub-systems. There is a catch involved with security measures at this external level. If you do not take precautions, such as encryption, then your message will be accessible to a very large

---
5 Not relevant if fibre optic cables are used since these are secure against such methods. It is possible to intercept fibre optic cables, but such methods are beyond the scope of this dissertation.
6 Or a popular euphemisms for this activity.
number of people. However if you do use encryption you may attract attention from an unexpected source. The security forces of both the United Kingdom (via the General Communication Headquarters - GCHQ) and The United States (via the National Security Agency - NSA) monitor all international and satellite communication via a system called 'Echelon' (Cornwall, 1988). This system monitors all communications and the use of encryption would attract their attention. This could entail being forced to disclose the encryption key to these agencies.

### 3.4.2.7.2. Incoming Communications.

Under this section we will consider the methods of security against human and automated attempts to breach the target computer’s security. These will be termed 'Hackers' and Automated Processes (APs). The latter term will cover viruses, trojans, and other software threats. The use of an intermediate device to deal with incoming communications can help to reduce the risks considerably. This external device can check if the external user, or process has issued the correct password before the target computer system allows the communication link to be established. This system can be further improved by using a security modem which phones back to the number associated with that password, and then uses a secret password with that modem to establish that the correct modem is being used. Of course this will only ensure that the link is with an authorised user, it cannot guarantee that the user will not breach security once they have access to the computer system.

---

67 The British end of this monitoring operation is run under the authority of the 1985 'Interception of Communications Act' section 3(2).

68 Soon to be superseded by US Government project code name 'P415' (Cornwall, 1988).

69 Under the authority of the Interception of Communications Act (1985), and more recent legislation users of public computer communication networks who send encrypted messages can be required to supply the decryption key.

70 Trojans are programs which seem harmless or entertaining, but in fact hide a secondary purpose. The deception is to get the user to run the program on the target system. Games are very popular examples.
Hackers.

The 'hacker' is someone who infiltrates computer systems, usually as a hobby, but also professionally. There are various levels of security available which can be used on a computer system, and computer manufacturer’s claim these can make a system secure against an individual once they are inside a computer system. These are both in terms of hardware and software solutions (software based security will be discussed under operating systems). However these are mostly for large mainframe or mini computer systems, the problems faced by a micro, once a user has access, are currently unsolvable. Certainly many of the measures we have already described, such as encryption of files, off line duplication of data, and Write Once Read Many Times (WORMS)\(^1\) data records of the systems activities, can all help. However it must be stated that once a knowledgeable person has unsupervised access to a computer system, there is no effective protection against them committing a malicious act. The best safeguards in such situations would be those built around WORMS recording of the user’s every action (perhaps even right down to individual interrupts.). The WORMS files will still be readable by the 'hacker', and all magnetically stored material will be readable and easily modified by the 'hacker' (this includes all the software based security systems). 'Hacking' is becoming big business, and with the average 'hack' costing the affected business 390,000 pounds sterling (Hamilton, 1989), it is likely that hacking will become attractive for many highly intelligent people. A possible solution for the parapsychologist who wants to allow the Ss to participate in a computer based test from 'home' is to use intermediary hardware, and not to let the S actually log on to the target system. The S could communicate with the intermediary system, which is a slave to the master target system. This would reduce the security risks. Other alternatives are to use dumb terminals as the S's interface to the system, although this is merely an extension to the master/slave scenario outlined above. To give an idea of the difficulties involved in producing secure computer systems, the US Government’s former director of information systems at the Pentagon, Mr Stephen Walker has publicly admitted.

\(^1\) WORMS is a method of storing data on a compact disc, so once it is written to the disc it is uneraseable. The physical properties of such compact discs are such that the data is (for all practical purposes) permanent.
Foreign Intelligence services have gained access to classified information in US computers by remote means. (Peterzell, 1989).

If the military are unable to maintain secure computer systems, with huge budgets, it can be seen that the experimenter's task is going to be difficult.

**Automated Processes (Viruses, Trojans And Other Tools Of Disruption).**

Viruses are programs that 'hide' in the computer software, multiply themselves, and additionally can perform various tasks. These tasks can be as benevolent\(^{72}\) as a welcoming/warning message, or as malevolent as their creator's imagination. Both the US Government's NSA and CIA have experimented with the disruption of other nations computers by the means of viruses (Peterzell, 1989). The use of a virus in this circumstance is to infiltrate as many of other nation's computer systems as possible, and lie dormant until an activation signal is detected by the virus. This activation signal would probably be in the form of another software carrier (even as innocuous as a message with keywords), or be as subtle as a variation in communication line transmission rates (or mains frequencies). Once this activation signal was received the virus would then disrupt the host computer system. Of course if enough time has passed since the initial infection, all the generations of backup systems would also be infected. This would result in the opponent's computer system being completely immobilized. Stephen Walker, former director of information systems at the Pentagon, has publicly stated that the potential of offensive use of virus attack was rated as great as that of nuclear or chemical weapons (Peterzell, 1989).

Having described the effects of a computer virus we should discuss how they are transmitted, and prevented. A virus, can only infect a system by software or hardware contact with foreign (non target system) computers, or foreign media.\(^{73}\) However having said this we reach the limits of what can be stated definitely. The method used to hide the virus, actually copy itself to the host, and how it hides, are only limited by the virus writer's imagination. Crude preventative methods, like write protection of discs which come into

\(^{72}\) Even so called benevolent virus infections can cause data loss and damage, simply by the act of hiding within the system and data files.

\(^{73}\) Unless it is written 'in-house', a possibility which is likely, but beyond the scope of this chapter.
contact with foreign systems, can be avoided by using the non-volatile RAM (see real time clock) to hide until a chance arrives. Software protection systems are only effective against superficial attack, the dedicated virus writer directly accesses the hardware, bypassing any safety checks in the operating system. In summary the only guarantee of protection against software virus attack is to avoid contact with foreign computer systems, or use dumb (one way) connections with them\textsuperscript{74}. Viruses often use a technique called "Trojan-ing" (not to be confused with the 'Trojan Horse' program discussed later). This involves hiding the virus in pirate copies of games, or commercial software. The user copies the games onto the system, and infection is achieved by this route.

**Other Automated Processes.**

Apart from virus programs there are a whole host of types of software that a 'hacker' may use to help infiltrate a computer system. An obvious example is the automated password guesser, which uses a fixed algorithm based on the combined experience of hundreds of manual attempts to guess passwords. These programs simply repeatedly try to guess a password until they gain entry. Other programs such the 'trojan horse' record valid users log on procedures. This is done either by emulating the target computer system, or by recording all the log on activity on one line to a computer. The secure modem (see LANs) removes this possibility, since the user must have a special modem resident at a specific telephone number. The same precautions that apply to excluding virus attacks also apply to these threats.

3.4.2.8. **Communications With Devices Within The AIE.**

One of the prime dangers from disc drives and other peripherals is that of the devices themselves interfering with one another by their respective emissions (personal communication Macleod, 1988).

3.4.2.8.1. **Disc Controllers, And Disc Drives.**

Another of the artifacts which could arise if an experimenter assumed they were using an UM is produced by the disc storage sub-system. The time variability that it takes a system

\textsuperscript{74} Although even this is not totally secure (for examples see Cornwall, 1988 p87-91)
to access a file will depend upon where the file is on the disc, how much data is being written, how large the file already is, and the state of the disc. For example the layout (physical positioning) of a disc is determined by the disc drive index (called a File Allocation Table (FAT) in MSDOS). Two discs with identical amounts of free space could have very large access time differences due to the way that existing files held on the disc had taken up space on that disc. The time taken to write to an empty file will be different to that taken to append to the end of a large pre-existing file. These differences will result in performance variations for any user program. The potential for these file access time differences will be worse on a hard disc system that is used by other people outside of the experimental period. Each different user is likely to be adding and subtracting files, so the resulting FAT will be different for each S’s session. These problems can be minimized by using empty identical floppy discs, which have their identical contents written in an identical order by the same machine for each subject. This ensures that each S gets an identical performance system. This was the method developed and used in the KMDB system.

3.4.2.8.2.Hardware RNGS - Internal And External (Serial Port Or Plug In Board).

Parapsychology is one of the few academic areas which still makes active use of truly random sources of randomness. The majority of users are now content to use PRNGs (Personal communication, Royal Statistical Society, 1987). Chapter two of this dissertation covered the uses and types of random number generators in some depth, so they will not be duplicated here. This section will also assume that the RNG has been constructed to produce ‘good’, or ‘true’ random sequences. It is beyond the scope of this discussion to detail with the physical construction of ‘true’ RNGs to avoid bias. Interested readers are referred to standard electronics text books. Instead this section will concentrate on methods of compromising the integrity of hardware RNGs.

Parapsychological Hardware RNGs.

The problems involved in hardware random number generation involve more than excluding regularity from the system. There is a whole complex theoretical background to psi based

---

75 This problem would still apply for a floppy disc system that used the same disc for recording the Ss results, and loading the users program.
hardware random number generation which is not shared by traditional electronics in their hardware RNG construction. Some psi researchers feel that the regular noise processes of thermal or shot noise are less psychically influencable than irregularly produced noise from bulk defect and semiconductor surface leakage noise effects (Chevako, 1983). They also feel that the wider the signal processor bandwidth and the lower the frequency, the more likely PK actions are to be able to influence it. The problems often associated with hardware RNGs are high order non-randomness, which can produce repeating patterns. These are most often associated with the physical properties of the noise generation method. Semiconductor based noise generators have surface leakage currents, and arbitrary state-lifetimes ('traps') of the internal crystal structures (Chevako, 1983). In the older thermonic value noise generators it is the varying properties of spatial and temporal surface emissions that are the major factors. These regularities, caused by excess or 1/f noise, would be a particular concern if one were looking for 'psychic signatures'. Some researchers (non-OT) are concerned that within the RNG the resulting noise is only the result of the actual noise source and not any additional noise from the amplifying and measuring equipment. This is so they can experimentally adjust the characteristics of the noise generation system, to get more data on any physical processes involved. Obviously these are complex issues, and are beyond the scope of this discussion; interested readers are referred to the relevant parapsychological literature.

Compromising Hardware RNGs.

Obviously one of the easiest ways to compromise the results of a RNG based study is to invalidate the use of the RNG being used. In the case of the hardware based RNG this is done by either tampering with 1) the monitoring computer system such as to affect processing of the RNG output (already discussed above), 2) some physical part of the noise generating system or some part of its AIE (see section on AIE), 3) the interface between the RNG and the recording/using system, 4) the recordings or use made of the RNG by the system(1), 5) the experimenter, or 6) influencing the storage of the data once it has been recorded. These influences can be either to corrupt and spoil the data, or to change it to a

\[ \text{For example a null signal could be produced by turning the RNG off, by disconnecting it from the monitoring system, or disrupting its normal functions by turning it off and on during each trial. The experimenter should be able to guard against all these possibilities.}\]
'required' state. The latter is much harder, and the degree of difficulty is regulated by the precision of the influence required. The majority of parapsychology studies will have a weakest point somewhere in the areas one to 6. These factors should have their tightest controls during the 'live' part of the experiment, since it is generally assumed that will be the exclusive moment of subject/target interaction (Morris, 1986 p74). In most cases this is of course an invalid assumption, as the subject may have many opportunities to interact with areas one to 6, or their AIE. All noise generating systems have some AIE, and changes in this will produce some changes in the RNGs behaviour\(^77\) (see section on AIE, above, and the author's 1987 P.A. paper on simulating man machine anomalies for greater details about the methods behind this). This chapter already covers most of the methods available to influence areas one to 6, except perhaps the human experimenter part 5, and part three the interface between the RNG and the recording/using system. The human experimenter may be influenced in many ways, but these are beyond the scope of this chapter\(^78\), and have only been included for the sake of completeness.

This leaves only the interface between the RNG and the monitoring system to discuss. If the hardware RNG is separate from the computer system (called a serial connected RNG), then it is considerably easier to influence its output, especially if the S is left alone with the unit. Most hardware RNGs will be producing binary signals, that is 1's or 0's. It is also likely that the direct signal will be taken to be the resulting number. So if the S wanted to produce zeros all they would have to do is disconnect the RNG. All the serial connected experimental RNGs found in the review of chapter two use a communications system where the computer signals the RNG unit with a ready, and then gets the random digit back. A small serial port attachment that could simulate the RNG's response to this kind of communication would be trivial to construct. Alternatively the S could just induce an electric current in the connecting cable. The only danger with this latter method would be if the voltage was too much for the serial port and caused some damage to the computer system. These same effects

\(^77\) Even an atomic RNG will use electronic transducers to record the sub-atomic activity. These translation components will have the same AIE as all non-shielded electrical components.

\(^78\) It is unlikely that an experimenter will face influences beyond those produced by distraction, or sleight of hand. Plato's three primary motivators of Pleasure, Power, or Knowledge are rarely used against an experimenter just for the sake of altering experimental results. Interested readers are referred to Barber, (1976).
could be duplicated by software which redirected the serial port traffic and amended them to desired levels. In comparison internal plug-in board based hardware RNGs are much more secure, since the connection between system and RNG is inside the system unit. These systems can be influenced either through their AIE, or by influencing their behaviour via software similar to that described for the serial port (above). The user’s system software would be oblivious to such programs (see section on defenses below). By studying the various methods outlined in the sections given in this chapter it should be possible to work out a whole variation of techniques which would compromise any RNG system.

Defenses.

In the section on encryption of data, it was stated there were some applied uses of encryption which could enhance existing parapsychology’s experimental design. This is one of the examples which that statement was referring to. If the hardware RNG used encryption of its transmitted data it would become much more difficult to compromise. The hardware RNG spends most of its time idle, waiting for the system to ask for the next RNG output. This idle time could be used by the RNG to encipher the next result. If this included a counter of how many RNG digits had been produced in this trial, an extra safe guard would be in place. The advantages in increased security outweigh the time it would take the computer system to decipher the generated signal.

3.4.2.8.3.Computer Screens Or Monitors.

The computer screen can also be the cause of a breach of security. A system known as Radiation Intelligence (RINT) can be used to reconstruct the contents of the screen (and other features of the system), at a remote distance. These emissions can be shielded, or there is also a counter system called TEMPEST which can be used to check the emissions from a computer system. However access to Tempest is restricted, since it was developed by the US Government NSA in the late 1970’s to combat the KGB’s then widespread use of RINT against US computer targets.

3.4.2.8.4.Disc Files.

The dangers here are that the files used by the experiment may have been altered or deleted. The experimental system must be able to cope with either eventuality gracefully. The use of encryption, and unknown file formats are also a great help here. Files should not be written
to disc as ASCII, unless they are encrypted. Binary files with an unknown internal format, are much more secure. The degree of security can be improved if all numeric values are encoded in some preset, but secret way and written to the file as full 16 bit binary values. If 'dummy' 16 bit numbers can be inserted safely into the files, this will make the task of decoding and reading them more difficult. If data is being written to a disc, then it is a good policy to duplicate the data to several files (this was done for the KMDB project), each with different codes, and internal formats. The analysis program should read all the files and double check the contents of one against the other, marking any exceptions as probably fraud. This considerably enhances the security of the recorded data.

Disc Damage.

Magnetic media are highly sensitive to environmental affects. All the items listed under AIE can affect the integrity of magnetically stored data.

Disc Swapping

The experimental system should also be able to check against the possibility of the Ss (or confederates) removing or switching floppy discs in the middle of the session. Some methods of avoiding this problem were developed from the KMDB project. First the experimental program can leave some files open on the disc for the duration of the experiment, and write to them at unpredictable time intervals. If the disc is removed then the operating system will look for the open files on the replaced disc, and find the FAT (File Allocation Table) is not the same. This will cause the operating system to report a fatal error to the user program, unless the disc is identical to the experimental system. To guard against this unlikely situation the alternative (which was used for the final series of KMDB studies), was to have a hidden unique software readable mark on the disc which was read at unpredictable time intervals, and checked against a hard coded check mark (which includes the unique subject number) within each Ss program. These unique hard coded subject numbers should also be included in each of the seven independent data files generated by the system (see above).

3.4.2.8.5. Data Storage.

It is important that copies are made of all important computer based information, known as a 'backup'. These backups should be made regularly, and the resulting discs or tapes stored
at a remote secure site. If a rotation of several 'generations' of these backups are made then the system can be certain of being able to recover from most forms of breakdown. These procedures are well documented, and will therefore not be detailed in this chapter. Readers who are unfamiliar with these concepts should refer to standard computer management texts.

3.4.3. Software.

3.4.3.1. Operating System.

The operating system (OS) is the most important piece of software on the computer. It controls the basic hardware, and allows the user programs to access what seems to be an UM. However the operating system is a program just like any other, and you will recall that as such it must be expected to have bugs, and inconsistencies in it. In extensive tests with MSDOS these inconsistencies have been found to be quite extreme, and they aggravate the hardware UM problems we have discussed above. During the development and testing of the KMDB system in 1987, various time differences were found in the running of an identical piece of code. Independent confirmation of these effects were made a year later in a report published by the British Standards Institute (BSI) microcomputer language compiler testers. They reported the results from timing benchmarks they had been commissioned to complete on various BSI standard language compilers (including the compiler used in the KMDB project). The first effect they reported was that they had to perform many multiple timings for the same program, and report the average as the time. This was because the programs had an up to 10% variation in the times it took to complete a program (Souter & Davis, 1988). However above and beyond this quite large variability they noted that MSDOS had a tendency to take un-warrentedly long periods to complete tasks.

..in a few inexplicable cases. It seems that DOS occasionally sulks for longer than this (0.1 of a sec), although no pattern or predictability could be found for these rare cases (Souter & Davis, 1988 p141).

On the basis of the investigations conducted for the KMDB project, it seems certain that these effects are due to a combination of the actions that the operating system has previously completed (perhaps leaving bit flags set), and the date and time held by the operating system. To reduce this variability the KMDB system used identical disc formats (see discs), identical copies of the operating system, and set every Ss to have the same date and time (by overriding the defaults from the Battery based clock, and setting standard ones in the operating system
start up process). The order of programs run by the operating system was also controlled to be identical for each subject. Flaws in the operating system will still have caused minor differences, but these were beyond the experimenter's control\(^79\). Since the SSM task feedback to the S was directly in terms of the amount of time the system took to retrieve the record, even the smallest temporal variations were important.

### 3.4.3.1.1. Security Of Operating Systems.

The problem involved in the security of operating systems is that they are designed to provide the services of the hardware to the user. The more, and better service they provide the less secure they become. All the main computer companies supply security add-ons for their operating systems. For example IBM markets products called Resource Authorisation and Control Facility (RACF), and Access Control Facility 2 (ACF2). However even with these security features there is no such thing as a secure operating system. A well run single user micro system can easily be made more secure on the most expensive mainframe, simply by restricting access to only one user. The first levels of problems involve the apathy and inefficiency of the operating staff and users. To make a system secure makes the system unfriendly, and even 'user-hostile'. There are minimum standards in existence; for example, the US DOD has an orange book security standard for computer operating systems. Level 'C2' allows the system manager to set access and audit controls\(^80\), which are all too often left unset, because of the extra effort involved in maintaining security.

There are two types of security expert systems which are worth briefly mentioning. The first type allow the user to enter the specifications of the computer system, and will then highlight potential security risks. This system has been developed by the UK Government, and at the time of writing is available for commercial 'consultations'. The second type of expert system resides in the operating system and looks for potential security risks as they occur. Interested readers are referred to such specialist groups as the British Computer Society Security Special Interest Group to keep up to date with rapid developments in this field.

\(^79\) Unless the experimenter writes his own operating system.

\(^80\) Orange book standard 'B2' has these security measures set automatically.
Breaking Out Of The Experimental Program.

The experimenter must assume that the fraudulent subject will try to break out of the experimental program to the operating system. It is therefore important that all the various control codes that can normally achieve this are disabled during the experiment. Technical details of how to achieve this will be specific to individual computer systems. Interested readers should refer to the technical documentation for their system.

3.4.3.2. Software Random Number Generation.

We have already seen that currently available hardware RNGs are easily subject to fraudulent intervention (see hardware RNGs). It may be presumed then that software generated pseudo random sequences are a much better solution, but unfortunately this is not the case. Such PRNGs rely on the AIE of the computer, and the UM based emulation of the computer. In the section on problems associated with the computer’s emulation of an UM, we discussed that the faculties provided by the computer’s hardware and software may be deceptive. Parapsychologists have produced some good papers detailing the problems involved with the concept of randomness (Coffey, 198881), and comprehensive details of PRNGs used in parapsychological research (Radin, 1985).

Rather than duplicate the efforts of these two papers here, interested readers are referred to them. Instead this section will discuss computer based artifacts, and flaws that can compromise a PRNG’s use in an experimental setting, based upon the experience derived from this dissertation.

The actual requirements or use that an experiment will have for a randomness source should determine the method of generation adopted. If the task is to select one of two options on an occasional basis almost any method will be acceptable. However this discussion will concentrate on uses that require a more consistent, and reliable source. The general purpose

---

81 However it is possible to have reservations about Coffey’s conclusions that a set number of tests can define a random sequence. Defining the correct number of tests is equivalent to the famous ‘halting problem’ in computer science, and both can be mathematically proved to be unsolvable.
PRNGs provided by computer manufacturers are almost invariably poor\textsuperscript{82}. Experimenters should simply avoid the use of generic PRNGs. In an unofficial experiment conducted by the author of this dissertation, a generic PRNG (not ever used in the main studies) which had passed a Chi-Square randomness criteria, showed extensive regularity and sub-period patterns (see figure). These flaws were only shown by a random walk, and an amplitude spectrum from a fourier transform. These kinds of effects are dangerous for the parapsychologist, since they could be misleading in the generation of independent targets or RNG PK studies looking for 'PK signatures'. The largest problem with PRNGs is that every designer of a PRNG has created their PRNG to be ideal for a specific criteria, such as overall distribution shape, period length, or speed of generation. The very act of designing a PRNG for these uses can induce disastrous non-randomicity in other aspects of the PRNG, such as sub-periods, numerical interdependence, or just straight bias. What is acceptable for testing bandwidth tolerances, or simulating photon steps from a foggy atmosphere will be unsuitable for testing new and strange forces such as PK and ESP. The consequences of using a “bad” Pseudo Random Sequence (PRS) depend very much upon the application. Sometimes a study might find it useful to know that the sequence will produce wild glitches, and so long there is an equal probability of the “badness” being in one area as all others it

\textsuperscript{82} A notable exception being the PRNG provided by DEC on their Vax computers (Personal communication Ripley, 1987).
may not be important. Obviously the intended use will dictate the experimenter’s attitude, but any patterns found in data produced by a PRNG should be taken with extreme caution.

The experimenter who still wishes to use a PRNG is really only left with the option of implementing their own PRNG from a trusted source (see Radin, 1985) and testing it. Even when the experimenter has found reliable and efficient methods of pseudo random number generation there are still some traps for the unwariness. These problems can be summarized.

The Fixed Sequence Problem.

The sequence produced by a PRNG is determined by the mathematical functions performed, and more importantly the values fed into the PRNG at the start of the PRNG. The same initial seed will produce an identical sequence. This is known by all experimenters, and yet the UM aspects of the computer may mislead some of them into thinking that they have selected a ‘random’ starting point for each S’s session by using some aspect of the computer’s clock. Consider, an example where a S participates in a computer test that uses a PRNG in this manner. If the subject attends each session at a regular time every week, and some aspect of the computer clock is used to set the starting state of the PRNG the S is likely to start obtaining significant patterning in their results. This could mislead experimenter’s into thinking a ‘PK Signature’ was associated with that S, even though overall results may at chance.

Parapsychologists often attempt to increase the ‘lability’ of PRNGs by rapidly re-seeding the PRNG during its generation process. That is, the PRNG is not being used as it is intended with a single seed, and then being left to generate that seed’s respective sequence for the duration of required numbers. In this rapid re-seeding the output of the PRNG will be totally determined by the seed values. If the S could supply the same seeds to a PRNG they would get the same result. In many cases the computer clock is used to derive the seeds; this is called fast clock seeding (FCS). The statistical randomness properties of such a PRNG are totally defined by the randomness of the seeds that are fed to it. In effect the PRNG becomes

---

13 Or hardware RNG for that matter (see section hardware RNGs).
14 Being wary of such aspects as word size and rounding effects of computer arithmetic.
merely a transformation process, and the randomness is derived from whatever is producing the seeds. If the computer clock is being used there are two immediate problems. To simplify the process of illustrating these problems we will assume that the fast clock digits alone are being used, that is the computer's tenths and hundredths recorded time. First is the fact that we have already seen that the clock is not perfect, and in fact only records in 15-18 millisecond bursts. If the experimenter is counting on the seeds providing a 00 to 99 range of possible seeds they will be disappointed.

The second problem is that if the subject does not decide what clock digits are going to be selected, the computer does, or rather the computers interrupt scheduler does. This means that the factors outlined earlier under the UM problems will interfere with the seed values that actually get used in these kinds of processes.

Given the computer clock and interrupt timer problems we have discussed in the UM section we cannot assume a normal distribution of clock digits being handed to the PRNG. At least not so we could be happy for such a system to be used in parapsychological situations. Among the data collected from the KMDB project a large (N > 25%) number of the fast clock readings were highly biased towards even numbers. This was especially true of systems with weak real time clock batteries (see the problems section of the pilot experiment).

**Initial States And Possible Solutions.**

These discussions are leading to another problem, and that is who determines the initial seeds, and the resulting randomness. The scenarios we have discussed above would seem to suggest that the time that the experimenter turns the machine on or initiates the test would probably determine the psi effects to an undesirably large extent. Thankfully there are some ways of limiting the extent to which it is the experimenter's influence that is being measured. Consider if every subject in an experiment using a PRNG were given the same initial seeds! They would each generate an identical PRS. However if every action the S subsequently

---

85 If they are not, then it will probably be derived from a number which is either the date and time combined, or the number of channel one ticks since the machine has been turned on. It is hoped that it will be clear that neither of these solve the problems.
made supplied a number (derived from the fast digits of the computer clock), which advanced the PRS by the number of iterations determined by that digit, all the randomness produced will be derived from the S. It must be noted that the seeds are not replaced in this method. Rather the existing sequence is advanced by a number of iterations of the PRNG, and the last number to be generated is taken as the S’s generated number.

Regardless of the individual methods of PRNG used the initial seeds, and overall outputs should be recorded.

Control For Bias In The KMDB System.

The KMDB system’s use of binary tree placed targets, each with a 0.5 probability of being correct excluded any artifactual biases that might have existed in the REGs used in the smart noise system. The full details of which are discussed in appendix 13.

3.4.4. Other Factors Not Covered In The Previous Sections.

3.4.4.1. Subject Fraud.

We have not covered what the experimental system or experimenter should do if they encounter an artifact which would have produced misleading results, whatever its source. The reaction of the experimenter should of course be to invalidate that session’s results, and maybe the entire series. If a S is detected engaging in fraud then the actions to be taken will be determined by the protocols laid down by the experimenter for such situations. However the way in which the system behaves is also important. The system’s highest priority should be the integrity of the experiment. This means that if the system detects any attempts to compromise security it should destroy all the data, and system information it can. If this not done then it is likely that a sophisticated attempt will be able to determine a great deal about the system from its remains.

Each KMDB system is created uniquely for each subject by a special computer system. As well as ensuring that the disc file formats, operating system, and behaviour is identical for

---

\*\* Morris (personal communication 1989) has mentioned that this would be problematic if Ss saw their results, disliked them, and then created an artifact. \*\*
each subject it also adds some other features which can act as useful examples of methods to enhance security. Each subject has a unique number, and this is hard coded (as constants) by the system in the creation process (see later note on why hardcoded constants are used). In addition the disc dedicated to a subject is marked with hidden marker information that includes the subject’s number and the number of times that experimental system has been run. The latter value is used to mark the data files produced by the system to identify them uniquely for the analysis program afterwards. It also acts as a check to see how many times the system has been run. Within the KMDB project scenario it is possible that the system can be run more than once, in which case the system creates uniquely named archive files for each run, each identified by the run and subject number. This prevents the old data files of one session being overwritten by subsequent runs. It often happens that subjects inadvertently enter the exit or quit command in the middle of a session. Without this feature a restart of the system would overwrite all the recorded data of the last session. Instead at the end of a session (at close down time) all the data files are closed and then SEA’s public domain archiving ARC utility program is spawned as a sub-process. This collects all the recorded data files from the run into an archive based upon the subject number interface type and run number which is current for that system. For example, subject 1’s command line data in their first session would be called CID1.arc, a second session would be called CID2.arc and so on. This concept is explained in greater depth in the KMDB user guide.

One feature which is not mentioned in the user guide is the tidy which takes place if the KMDB system finds existing files which have the same names as those used as data files by the KMDB system, but which have not been archived into an appropriate '.ARC' file. This could only legally occur if the system had a fatal crash, or the power was interrupted during the working of the system. However this could also be caused by a subject trying to insert false data files. If the system does find such data files on start up, it ARC’s them into a file called FISHn.arc, where n is the current run number. This allows the system to continue and yet still separate the data files it would otherwise overwrite.

It is also appropriate to describe the activity of the system if it discovers some aspect of its Associated Implicit Software Environment (AISE) is missing. In most cases it can gracefully

87 Fish here denotes that the files are 'fishy'.
cope with such contingencies as missing system components, help files, and invalid seeds. The system can either generate temporary versions of the system components (for example the PRNG seeds), or can run equally well without them (for example the system data files). In the case of the help files it merely informs the user that it has been installed without the help files, but continues to run.

Security Marking Of The Discs.

The system has information stored on the disc, which it checks with its own internal values. Any mismatch causes the system to fail to boot up, drawing the experimenter's attention to the problem. For example the KMDB system hangs up with the message 'Dos error illegal op-code' if the run number, or subject number is invalid, or if the hidden markers have the wrong format. If the hidden files do not exist at all then the system generates the message 'Dos error invalid memory format'. These messages are both valid operating system error messages, and normally return specific codes in the computer program status word. The KMDB system duplicates the correct status word codes for each of these respective errors. This ensures that to all observers the effects look genuine. However in both cases the system will have permanently entered its security breach mode.

The KMDB experimental protocol was that in the case of a security error the subject would be removed from that subject pool, and this would stop any further attempts by that subject at using one of the systems. The system could then be examined to see exactly what had been happening. An important aspect of computer security is reviewing actual attempts at breaching security. It is by this review process that security measures can be improved. However it is also possible that a system could fall into the hands of some fraudulent subjects who would try to run the system again in the hope that they could generate false data, or learn about the system. To cover this eventuality security breach mode was developed. There were strong temptations to make the KMDB system actively aggressive to attempts to compromise it. This could have been by spawning destructive virus systems which would

---

88 In the Gem version both of the discs have these hidden markers, and both are checked at run time.
destroy the host interrogating system. Instead it was decided that the KMDB system would restrict its actions to its own components. As soon as security mode is entered a hidden marker called Hara-Kiri\(^\text{89}\) is created on the disc. The marker holds a number which is the number of times the system has been run since the attempted breach has been detected. This is held as a hard coded ProPascal numerical constant\(^\text{90}\) in the KMDB system, and is compared with the Hara-Kiri value. This process allows a set number of run attempts to be made before any action is taken. This feature was included for debugging purposes and during the experiments was set to the value 1 for immediate destruction. When the system does destroy itself it first disables all interrupts, and then deletes itself and all its data files from the disc. Its next act is to create a dummy version of itself on the disc, and fill its old file space with random numbers. This means that even if someone recovered (undeleted) the system files they would only recover a useless collection of random numbers and not the executable file which could have been analysed. If one wished to make sure that all traces of the system were destroyed one could, as a last effect, run a very small assembly routine to overwrite RAM as well.

**Experimenter Based Error.**

It would be nice to be able to put 'and fraud' in the title to this section, but if the experimenter is aware of all the methods that we have outlined above there is no realistic way to stop them from tampering with their results. However it is possible and desirable to have a system which creates each subject's experimental system so it prevents unauthorised, and mistaken duplication of such systems.

**Prevention Of Experimenters Creating Multiple Systems (Reduction Of Error).**

The special system which creates each subject’s experimental systems has a check to prevent a duplicate subject number being used. As we have already described every time the creation program is run it has to be supplied with a subject number. Every one of these are recorded

---

89 Named after the Japanese method of ritual suicide, exclusively used by the Samurai class in ancient Japan when they were disgraced (also known as Seppuku).

90 If the value was stored as a variable then it would be possible to examine it by using a disc editor. In contrast numeric constants are indistinguishable from any other binary value on the disc.
to a file, and when the creation program starts it checks that the subject number being requested has not been used before. It only continues the creation process if it is supplied with an unused subject number.

3.4.5. Post-Experimental Influence.

The final time period which has security risks attached to it is after the experiment has finished. It would be a waste of effort to guard against all the effects we have discussed if the data which has been generated was then left unprotected. The experimental system should have produced several independent copies of the data, each with different formats and encryption keys. This will protect the data from being amended only as long as it remains in that un-translated format. If the 'Beaufont Cypher' has been properly applied then the data held in the 7 files should be able to withstand a respectable amount of analysis. In the KMDB system the files were also combined by the experimental system into an archive using SEA’s archive program. This allowed the data to be stored using ARC’s encryption system as well as the system’s own secret file formats. It is possible (this was done for the final series of KMDB experiments), to hardcode unique encryption keys into each subject’s experimental system, so even the experimenter is blind to the keys being used. In this case only the creating program, and the analysis programs can then de-crypt the data files. This method would allow secure KMDB replications to be run by multiple experimenters, to test parapsychological experimenter psi effects.

Security problems only arise after the analysis has taken place, and raw data files are translated into processed data. This processed data is a vulnerable target for alteration or destruction. It is at this point that an alternative method of security is used, replication and distribution. The results should be duplicated and distributed to as many independent sources as possible, which can be subsequently used as reference values for the analysed data. In this process instead of restricting access and copies of the data, it should be made as freely

---

1 If the file formats are also unknown the task of decyphering them can be regarded as being practically impossible. This is why so much care was taken so the file formats were not revealed during the development phase.

2 It would however be unfortunate if there was a system error when this technique was being used, since it would result in the data being lost.
available to as many parties as possible. If this distribution can be securely automated in some way, so much the better. At regular periods a check should be made of the validity of the data being used to avoid the risk of it having been tampered with during the analysis stage.

3.5. Discussion.

Having discussed these possible artifacts it is appropriate to discuss the basic guidelines which determined the specifications of the KMDB system.

Principles of SNS design.

If a systems performance can be influenced by PK, or ESP then the concept of psi based FAA could be useful. Morris put forward this concept as 'Smart Noise' in his PF paper of 1984, but no research was conducted into this possibility until this dissertation. Given the complexity of the effects we have discussed it is perhaps little wonder. One of the major problems is defining exactly what is the function alteration 'task' the subject is being asked to attempt. This may seem a most simplistic question, since some would argue that PK and ESP computer based RNG studies have been doing this kind of research for nearly 50 years (see chapter two). However this is debatable. If you consider most computing based applications and the users of those tasks, the user's 'task/goal' may be many things other than the actual activity they are doing on the computer at that time. This may not be immediately obvious, so we will use an example. Let us suppose that someone is using a computer to perform the KMDB experimental task used in this dissertation. We can presume they may have a specific goal, but it may not be primarily to retrieve records. It may be to finish as quickly as they can, or to enjoy the feel and experience of the software which they are using. In these cases the actual computer task may be of secondary importance to them during the time they spend using the computer. Indeed the enjoyment of a specific 'feel' of a computer based task may be a major motivation for many users.

If this is the case then there is a complication in determining what the user 'wants' from the computer system. This risk will be increased in WIMP based environments, where the pleasure of using the interface may actually change the task Ss are assumed to be attempting.
The possibility of FAA being target specific is very important, and could produce many side effects. These could involve influences which would change aspects of the system beyond those required in the ostensible task. For example, our hypothetical S’s FAA might be directed at slowing system performance down, or concentrating the electron beam of the cathode ray tube (CRT) to improve the screen character definition, in order to increase his enjoyment of the graphical interface. Alternatively his FAA could be used to speed up disc access, but not influence the direct measuring devices. To try to minimize these possibilities we have to have some rather limiting directives placed upon whatever we use as our target application for a SNS (Smart Noise System). This why the so called ‘Ten Commandments for SNS’ were created.

3.5.1. The Ten Commandments For Smart Noise Systems.

1) All Taska FAP Must Be Limited (As Far As Possible) To A Small Monitored Domain.

If FAA can only occur within the FAP, then this rule will ensure that all the FAA is recorded by the system. If this rule is not followed then FAA might be limited to those parts of the system which are not being monitored, such as disc access, and screen definition.

2) The Task Must Have A Directly Observable Goal Which Is Easily Measured By The Recording/Monitoring Sub-System. The Monitoring Sub-System Must Have No FAP Of Its Own.

If the goal being used is not directly observeable, or easily measured it will be difficult getting objective criteria of success for statistical evaluation. It is an advantage if the goal has a statistical behaviour which can be easily determined. This reduces the work involved in evaluation, and makes the findings easier to explain to an audience.

3) The Process Initiated By The User To Achieve The Goal Must Be The Smart Noise Task.

If there are any sub-processes with FAP between the user and the SNS then the ability to investigate time determined process initiation will be lost.
4) The Task Must Be Unambiguous (It Must Not Have More Than One Possible Goal).

It is important that the user should only have one target (or none\textsuperscript{93}), for their task goal. If the SNS was allowed to work on a multi-goal task then the effects would be such that statistical evaluation would become much more complex, since the FAP would be much harder to define.

5) There Must Be Only Task Per Goal.

This is not quite the same as the rule stated above. This rule ensures there is only one set of activities of operations which allows a goal to be attempted. If this was not the case then it would be possible to have multiple routes towards a goal, and it is unlikely they would all have identical FAP. This would make inter subject trial comparisons problematic.

6) The User’s Feedback Should Be Directly Tied To The Success Rate Of The FAA Influencing The FAP.

This ensures that subjects get identical feedback for identical performance.

7) The System Must Record Everything (Possible) About The User’s Actions.

It is desirable to know what the user is doing on the system when they attempt a trial. This detail should include the commands they have entered, the errors they have made, and the respective speed of their actions on the system. This will allow an investigation into which aspects of general user behaviour correlates with the FAA.

\textsuperscript{93} It is legitimate to try to search for something that does not exist, and do it successfully. In the case of the KMDB it is possible to search for a record which is not held within the tree, and score above chance. This would occur if the SNS logic searched to the correct place in the b-tree so it could determine that the record did not exist, and that it did this in fewer trials than would be expected by chance.
8) The Experimenter Must Try To Avoid Creating Unlikely Experimental Situations.

The situation should be as near real life as possible. This will allow the experiment to investigate normal people under real world situations. If this is not done then one could be investigating 'lab psi' and not a real world phenomena.

9) The Experimenter Must Try To Exclude All Known Artifact.

The experimental setup should exclude all known methods of environmental interaction that could allow any form of information interchange.

10) The Experimenter Must Try To Duplicate All System Recordings, And Use The Maximum Security Against Every Eventuality (Especially System And Experimenter Error).

The security problems have been covered in this chapter. It is important to record everything it is feasible to record from temperature to weather (see session report details from experiments 1-4).

The result of all these rules was to force the study described in this dissertation to develop a database system which had all its record structures held in RAM, and which had no task based FAP outside the domain of its recording systems (see chapter five).

3.5.2. Conclusions.

This chapter has given a brief outline of the various problems involved in using a computer in parapsychological experiments, particularly those involving smart noise. It is not comprehensive, as that would be impossible, since security is only bound by the infiltrater's imagination, and technological innovation. However the use of covert studies or the avoidance of star subjects can allow higher degrees of security, since no one has any incentive to commit experimental fraud.
4. The HCI Aspects of the Experiments and their Design.

Within simplicity lies clarity. Within education the highest aim should be to convey discovered truth in its simplest form. This stops concepts being lost in language, and the changes that time, translation, and the audience’s attitude will have upon the facts presented in a discourse. Mastery of a subject is never shown by producing the most misunderstanding or reverence in an audience. It is shown by giving every listener the impression that the subject is both interesting, and composed of a simplicity that poses neither threat, or excessive complexity.


4.1. Introduction.

This chapter will give a review of HCI, including a summary of the major theories, interaction methods, attitude studies, and computer attitudes questionnaires respectively. This review will include details of the HCI aspects of the study described in this dissertation.


Computer systems form an integral part of many people’s lives. Many of these systems work to protect or help people, who may never know about the computer systems involved. Security and weather monitoring systems are two examples of these kinds of systems. However these are exceptional cases. For most people the computer means a complicated esoteric device that provides a nearly magical superiority upon those who take the time to master their use (Shneiderman, 1987). This indicates something implicit about the image that computer systems have had in the public’s eye. This image is derived from many sources, the most widespread of which are the unnatural (or perceived to be unnatural, or
de-humanising for the operator) methods required in controlling the computer's actions, and the 'mystique' created by the popular media (and perhaps some of the users).

Human Computer Interaction (HCI) is such a new discipline there is no universally recognised reason, or set of reasons, why some interface\(^1\) designs are so counter-intuitive for some user groups to use, and yet are preferred by others. This may be because the term HCI means different things to different groups of researchers. The academic discipline of HCI is very broad, and employs skills\(^2\) from areas such as computing (computing devices), computer science (software & interface design, construction, and computational methods), ergonomics (human physical capabilities), and psychology (human cognitive, and behavioral capabilities).

### 4.1.2. Types of Interaction.

Various types or methods have been developed to control computer systems. Since the following discussion will address detailed aspects of these methods, a brief overview of the major interaction method will be presented.

#### 4.1.2.1. The lowest level of interaction.

The power of a computer is directly derived from its ability to perform simple physical symbol manipulations at high speed. These symbols can only be from the set of binary symbols 0, and 1\(^3\). Any further attribution of meaning to these symbols only exists in the mind of the user. A device which performs such physical symbol manipulations can be termed a computational engine. This concept allows the user to adopt any method of interaction with the computer that results in sets of these symbols being presented for

---

1. The method which a human user uses to control a computer system is usually called the 'user interface', or often just the 'interface' within HCI.

2. The allocation of these major functions has been somewhat arbitrary, and probably reflects the author's own biases. In reality the reader should realise there is a large overlap. Within these different disciplines there is a diversity of opinion in defining what exactly are the most important aspects of HCI. This diversity usually reflects the 'mother' discipline of the researcher involved.

3. Soviet researchers have experimented with ternary state computation devices. These use the electrical states +1, 0, and -1 respectively. This has the advantage of increasing the capacity of the system, but the resulting technical problems have proved to be too large (Personal communication Mikov, 1988).
manipulation by the processor. For the duration of this discussion this will be our definition for the term interaction. The methods of human computer interaction which have been developed match the methods of sensory communication available to human beings in non-computational settings. Obvious examples are language, and environmental manipulation. These two divisions have been carried over into computer interface methods, and will be considered separately.

4.1.2.2. Conceptual models involved in interaction.

In computer systems the command constructs used in controlling the system are often called the syntactic definition. In contrast the underlying actions that are performed in response to the commands are termed a semantic definition (see figure). The semantic definition has been further split into divisions of computer concepts, and task concepts respectively (Shneiderman & Mayer, 1979; Soloway et al., 1982; Card et al, 1983). The task concepts are those actions undertaken by the system in completing a task. For example a user’s computer concept underlying the task of using a desktop publisher could involve the understanding that a program they control is making changes to the binary values held in a buffer in RAM memory. In contrast the same user could have a task concept of the same activity which was more abstract, involving the program creating a conceptual document. From these examples it can be seen that the computer concept is much more detailed about the actual operations being attempted by the system. The author believes that a mismatch between the user’s

<table>
<thead>
<tr>
<th>Syntactic Definition:</th>
<th>Command constructs used in controlling the computer system.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantic Definition:</td>
<td>Underlying actions performed in response to the commands.</td>
</tr>
<tr>
<td>Computer Concepts:</td>
<td>Behaviour of the computer system (hardware and software).</td>
</tr>
<tr>
<td>Computer Object Concept:</td>
<td>The object involved in an operation. Example = a memory buffer.</td>
</tr>
<tr>
<td>Computer Action Concept:</td>
<td>The action performed on an object in order to complete an operation. Example = put the character ‘A’ in the next position in the buffer.</td>
</tr>
<tr>
<td>Task Concepts:</td>
<td>The conceptual operations performed within the target task (regardless of their actual physical computer based operations). Example = format paragraph.</td>
</tr>
</tbody>
</table>


Figure 70. HCI Computer Models.
computer concept model, and the computer’s actual operations could result in many of the reported cases of computer anomalies. This would occur when the user assumed one set of operations were involved in a task, when in fact a completely separate set were involved. If the user then noticed a normal by product of these unknown operations they might assume that some anomaly had occurred.

4.1.2.2.1. Further subdivisions within computer concept models.

It is thought by some researchers (Shneiderman & Mayer, 1979; Soloway et al., 1982; Card et al, 1983) that the most powerful influence in computer use is the semantic computer concept. It is presumed that this knowledge differentiates novices from experts. The computer concept can be further split into computer object and action concepts respectively. Object concepts are the user’s conception of those things which are involved in some task, such as files, or buffers. In contrast action concepts involve the model held by the user of the operations carried out on these respective objects. Both of these can be decomposed to the lowest levels of computer operation. It is thought (Shneiderman, 1987 p49-50) that degrees of computer expertise are directly related to the depth to which a user’s object and action concepts cover computer operations, and are independent of the syntactic concepts. It is proposed this is why computer experts are so quick to learn the operational details of new systems. They already understand the set of objects and actions possible on a digital computer and can therefore rapidly infer the syntactic models.

These theories are important to good interface design, and easier computer interactions. The ideal interface, whose methods are easy to learn, use, and remember will have the closest match between syntactic and semantic concepts.
Computer syntax is based on a limited vocabulary of recognized commands which an input
device translates into a sequence of requested symbol manipulations. One of the earliest
examples was Charles Babbage’s proposed computational engine. He called this his ‘Ana-
lytic Engine’ (1833). This would have used such a language on punched cards which were
to be fed into a series of controlling devices. Although Babbage’s designs were not
implemented in his lifetime, the use of inputting commands and data on punched cards was
a highly successful interaction method during the early history of computing. It has had a
profound effect upon the method of interaction used ever since.

4.1.3.1. Command line.

Example from Microsoft’s MSDOS. This copies the file ch5.asc from a sub-directory on a
hard disc to a floppy disc.

`copy c:\qa\thesis\chapter5\ch5.asc a:v`

A punched card was derived from the cards developed by Jacquard for controlling the
weaving process on machines in the latter part of the industrial revolution. When these cards
were adopted by the computer industry they standardized on a card which could take up to
80 alphanumeric symbols on each card. These 80 symbols were represented by combina-
tions of physical holes punched into the surface of the card. These holes were sensed by
the machine, first by wire electric contacts, and later by optical detection methods. The
cryptic nature of commands in the commands language is derived from the habit of trying
to get as many instructions on each card. This was to minimize the time and effort involved
in data or command entry. When the first alternatives to card input arrived the operators
who had used the old card input methods wanted the same commands. This was how the
cryptic commands were carried on to the new medium. This was how the so called
command line environment was developed. In the same way the keyboard layout was taken
from the typewriter. The cryptic nature of the commands was perpetuated in the desire to
minimize the number of key strokes needed, since most technicians were not trained touch
typists. Most mainframe and mini computers have command line operating systems, the
most famous being the operating system called Unix. Obviously a command line system
is separated from the semantic model of the computer. Researchers have shown that
command line interfaces are comparatively easier to use and have less errors when command
names are consistently matched to the semantic models of the users (Ledgard et al., 1980; Black & Sebrechts, 1981; Scapin, 1981; Barnard, 1982; Kraut et al., 1983). The cognitive effort required to translate between the syntactic and semantic models makes command line systems harder to learn, use\(^4\), and remember than interaction methods which have a close link between semantic and syntactic models. These same features may make the command systems more attractive to the expert, or 'power' user (Shneiderman, 1987 p147).

4.1.3.2. Natural language.

Example humorously adapted from the natural language response from the US defense computer in the film 'War Games'.

Computer: Welcome Dr Macleod, shall we play chess, or would you rather play Global Thermonuclear War?
Macleod: Umm, Chess today I think...
Computer: Awwwh, but Hamish, its been ages since we played...

The dream of many HCI researchers has been the implementation of a computer that understood free-form English, either in typed form or as spoken language. These are termed natural language interfaces (NLI). There have been some difficulties in getting computers to understand NLI, due to the complexity of sentence construction. Problems with direct human voice input have been due to the natural variations involved in human speech. However there have been some successful systems, ranging from semi-free form typed database query languages, to a prototype voice driven word processing system. The early dreams of the advantages of using natural languages have not been borne out by research findings. Many studies have shown that alternative interactions systems, such as command line, or menu selection were preferable to natural language (Murray et al, 1983, Morrison et al 1984). This was in conflict with some of the results of earlier studies (Bevan et al, 1981). This switch could be due to a change in expectations of users over time. It is also the author’s view that the subjects in these studies may well have been used to either menu or command line interaction methods, but not to natural language. This could have resulted in a bias towards the interaction method the subjects already knew, and understood. Voice

\(^4\) Some recent research has reported that command line systems may be more suitable for controlling sequentially structured actions, while the superiority of the Direct Manipulation (DM) interfaces may only exist for manipulative actions (Rohr & Keppel, 1985).
driven systems have yet to reach levels where they are sensitive and accurate enough to be practical for anything but the most primitive of tasks, or where other interaction methods are impractical.

4.1.3.3. Menu Selection.

The so called menu system is often termed a halfway house between command line and some of the alternative interaction methods. It is ideally suited for systems with infrequent or novice users, since it presents all the available actions (requires minimal syntactic and semantic knowledge). The most recent graphical interfaces also make use of menus, in either their pull down, or pop down versions. The differences between these two variants are that the pull down menu requires the user to 'select (move the mouse over the menu bar) and drag (depress one of the mouse buttons and while keeping it depressed move the mouse down)' with the mouse in order to be presented with the menu, while the pop down menu merely requires selection. After the menu has been displayed the list of available functions are presented and by selecting one of the options (usually with a mouse controlled pointer), that particular action is selected. Some authors include form filling displays as a variant of menu systems (Shneiderman, 1987 p122), and even HyperText systems (Shneiderman, 1987 p120). These definitions are somewhat arbitrary, and it could easily be argued that these two methods have more in common with Direct Manipulation (DM) systems than they have with menus. In reality there is a large degree of overlap between interaction methods. It will be rare for any one interface to be purely dedicated to one interaction mode. The major question

Edinburgh University Computing Service.

Select From the following options:

1: Wordprocessing
2: Programming
3: Statistical Processing

Figure 71. Example Menu System.
to be answered is not 'what interaction method is best', but rather 'what interaction method is better, for which tasks, and by for which user group'.

4.1.3.4. Direct Manipulation.

Example:

Direct manipulation (DM) interfaces are where the user actually manipulates a computer generated graphical representation of the task. There is some disagreement about exactly where or who developed the first computerised DM interface. It is likely that video games were among the earliest ones. The earliest computer games emulated a ping pong game where the user controlled one of the bats with a game paddle. This interaction method is intuitive to use, unlike the command line systems, perhaps because it forms an automatic semantic model in the user's mind. The ease with which users learnt and used such DM systems was rapidly noticed by some
researchers, and research aimed at producing commercial DM applications were started. 'Virtuality' (Nelson, 1980) was among the first proposed computer generated representation of reality that could be manipulated by the user. Another early proposed system was 'PictureWorld', which IBM's Yorktown Heights laboratory reported as a future office system (Schild et al, 1980). However the most famous of these early DM interfaces is that of the Xerox Star (Smith et al, 1982; Foremski, 1989). From this pioneering interface work at the Xerox Parc laboratory was derived both the current standard DM interfaces, Apple's Macintosh, and IBM/MicroSoft's OS/2 interfaces called Presentation Manager, and Windows respectively. DM systems have as one of their aims 'Transparency' (Rutkowski, 1982), which is where the tool (in this case the computer) is no longer noticed by the user. The example which is often made is that of a hammer which is so intuitive to use that it becomes part of the user. The DM interface is still in its early evolution; in the next section we will discuss where that evolution may take it.

4.1.3.5. Future developments.

In a parapsychological dissertation it should be easy to include a heading, future developments, and be certain of what is discussed! Unfortunately it is not quite that simple. The future of computing is directed by consumer demand, unlike the developments in other academic fields. This demand is unpredictable, as the list of failed new products shows. However this section will speculate on what the future of DM interfaces may bring, since by doing so some of the reasons for conducting this research will be more easily understood. Ever since the concepts of 'virtuality' (Nelson, 1980) and 'Transparency' (Rutkowski, 1982) were first proposed computer scientists have dreamed of creating artificial realities graphi-

---

5 The Xerox Parc Laboratory is the founding influence behind much of modern computing, they invented: the first PC (Alto - 3 years before Apple), the first wordprocessor for novices (Bravo), ethernet, laser printers, smalltalk, RAM semiconductor memory, the Star interface, and with it the concept of WYSIWYG (what you see is what you get), and desktop publishing. This burst of development has been attributed to Xerox's policy of employing the best computer scientists and 'letting them do what they want' (Foremski, 1989).

6 At the time of writing there is a legal battle between Apple and Microsoft about the origin of the Microsoft product. This dispute has become heated since Microsoft Windows was recently adopted as the OSF (open software foundation) standard interface for future computer systems (Manchester, 1989).
cally represented by a computer. These realities would be directly controllable by the user, in the same way as current DM systems, only with increased functionality.

4.1.4. Adaptability.

The problems with such an interface are not purely technical, because as yet HCI researchers do not know what makes the proper mix between 'functionality', and 'usability' for different users. Functionality is where the design goal is to provide as many functions as possible to the user, whereas usability is the ease of using that interface. An interface which has a high degree of functionality will be hard to learn, and is unsuitable for the beginner. However as the user progresses through the learning curve the demands that they will make upon a system increase, and they demand increased functionality. Researchers have proposed that an adaptive interface may well be that solution to this problem. Such adaptive interfaces have been found to have their own problems. If the interface automatically changes its configuration as the user gains more experience this can be disturbing to the intermediate level user (Benyon, 1984), since it is felt to be obtrusive. Other research has shown that these feelings may only be applicable to certain user groups (Greenberg, 1985), and can be reduced if a more sophisticated adaptation algorithm is used. At the time of writing the most successful adaptive interface products (MicroEmacs, Borland Sprint) allow the user to change the interface themselves, but allow very little changes to the available functionality.

4.1.5. Method of manipulation.

To control a computer generated virtual environment specific physical methods of interacting with the system will still be needed. Current methods for this allow the user to control 2 dimensional environments using either a mouse (Lu, 1984), or an infra-red touch screen. Neither are ideal devices for gaining transparency (but the mouse often comes close). Unfortunately current touch screens are highly inaccurate (Beringer & Peterson, 1985), and tests between joystick, mouse, step keys, and text keys suggest that the mouse use conforms so closely with the Fitt’s Law slope\(^7\) that positioning times with mice are probably the minimum possible (Card, English, & Burr, 1978). Current mice are limited to 2 dimensional

\(^7\) The minimum hand eye placement ability of humans.
environments. Although some HyperText systems make strong claims about accessing 'depth', they can be ignored since that is not the kind of depth that is envisaged in future DM environments. Instead the use of some kind of hand model would be required for true DM virtual environments. Early prototypes of these hand sensing devices are already in use by NASA, for use in handling remote and complex computer systems (Foley, 1988). NASA's development system includes a head mounted monitor, but does not include tactile feedback (see footnote). There is also experimental evidence that brain waves can be directly used to control computer operations (Sem-Jacobsen, 1981). These systems would have the advantage of allowing the system to monitor the alertness of the operator (how much information can be effectively presented), and variables such as impending cardiac failure; obviously both of importance for military applications. The possibility of using a myoelectric channel for computer interaction has also been evaluated, and found to be highly effective, especially as regards the possible information transfer rates, but it was also found to be prone to large degrees of variability dependent on the operator's physical and mental states (Dockstader, 1987).

4.1.6. Summary.

In summary, the current state of the development of the computer has been likened to that of the aeroplane at the early part of the 20th century (Shneiderman, 1988). Computer systems are still evolving, and designers are probing the functions that such systems can fulfill, and how best to implement these applications. It is safe to assume that in the same way that aircraft of 100 years ago are regarded as interesting amusements, so will the functionality of the most advanced computer systems of today, as viewed by the users of tomorrow. In the overall history of the computer this current period of time is the switch from a designer/user to the mass market. The future could indeed be one in which the user 'plugs'

---

*In the write-up for experiment 2 (Morgan, 1988), the author proposed a possible computer environment where the user’s hands were in 'sensing-gloves' and accessed an array of 'virtual' tools in the computer generated environment. This would include a keyboard and mouse which would be generated on the computer screen and whose actions would be controlled by the actions of the user’s 'gloved' hands. Tactile and audio feedback would make such a virtual keyboard indistinguishable to its physical counterpart. Such keyboards could easily represent different language and symbol sets. This proposal predated Foley's (1988) publication, and the author's knowledge of NASA's prototype system, but still extends beyond any proposed use of the NASA glove.*
themselves into a computer system; this would be the ultimate in transparency. However all future systems, no matter what their interaction method, will still have to account for the variations in what the user expects and demands from a system. In the next section we will discuss the factors which HCI researchers have proposed to account for these differences.

4.2. Trying to know the user.

Many reasons have been put forward for the difficulties that some groups have when they interact with computers; these will be briefly summarized.

4.2.1. Individual differences.

These theories try to explain the preferences which different users have for interface types by some form of individual difference. These theories hope to isolate the factors responsible for these differences, allowing designers to implement some configuration in computer interfaces based on the results of a brief inventory of questions (Rich, 1983, Abler & Sedlacek, 1987, Benyon, Innocent & Muray, 1987). This research is still in its early development, and many of the findings presented here are controversial. The number of active researchers in the area of HCI is relatively small and this means that progress is slow.

4.2.1.1. Gender differences.

Many researchers in the area of HCI have noted with concern that computing is dominated by males (Reznikoff, Holland, & Stroebel, 1967). Many of the reasons that have been put forward to explain these differences are identical to the differences that have been proposed to explain the cognitive differences already found between the sexes. These are usually assumed to include aggression, verbal abilities\textsuperscript{9}, visio-spatial (Kaplin, & Sedney, 1980), and mathematical (Maccoby, & Jacklin, 1974, & 1980). These factors have been difficult to determine, and the exact differences are controversial (Block, 1976; Maccoby, & Jacklin, 1980; Demo, 1982; Caplan, Macpherson, & Tobin, 1985; Caplan, 1985; Elliot, 1986; Halpern, 1986a & 1986b; Hiscock, 1986). Explanations for these reported sex differences

\textsuperscript{9} Some feminist researchers have attacked the theories and research devoted to finding such explanations as being inherently sexist (Cameron, & Coates, 1985).
vary from genetic (Stafford, 1961; Hutt, 1972; Bouchard, & McGeo, 1977), hormonal (Broverman, Klaiber, Kobayashi, & Vogel, 1968); cerebral hemispheric lateralisations (Buffery, & Grey, 1972), through to social interaction and role stereotyping (Sharpe, 1976).

**Gender differences in HCI.**

It is only comparatively recently that HCI has tried to address this deficit in its understanding of a group that could and should make a larger contribution to computing. Within the past 35 years only 2% of HCI work has looked at gender differences (Hudgens, & Billingsley, 1978). Indeed a review of the literature conducted by Fowler and Murray (1987) found that many studies fail even to state the sex of the subject groups. Other studies have been limited to one gender purely by the nature of the task (for example Gomez et al, 1984; or Johnston et al, 1986). Other problems in researching gender problems are that the computer science departments, where most HCI research is conducted, are male dominated. Findings from such studies will obviously be biased towards any gender preferences from that subject population.

**Gender differences in computer task behaviour.**

Investigations in computer task behavioural differences between the sexes has been limited. One of the earliest such investigations looked at the effects of response time variation (between 1 and 6.5 sec) and list length (4, 6, or 8 digits) in a computer-paced addition task (Baker, Holding, & Loeb, 1984). Males were found to vary their speed of response, while women varied their level of accuracy. Subsequent studies confirmed this initial finding, which was not reduced by practice. The authors also investigated the effects of acoustic background noise, and time of day on the S’s performance. They reported that noise had some differential effects between the sexes. Females reduced their response rates with noise, while males made more errors. Some cognitive theories have predicted that such task based differences between the sexes would exist (Luchins, & Luchins, 1984). The apparent differences between cognitive strategies adopted by the sexes are discussed separately (see cognitive differences).
Another experiment which investigated the performance variation of males and females on computer tasks looked at the effect of ethanol \((0.4g/kg)\) on a computer video game task (Crow, & Hirdler, 1985). The ethanol reduced the males' 'uncertainty index of discrete-movement', while it seemed to increase the females. It was postulated that these differences were due to the higher levels of stress present in the females, because of their unfamiliarity with video games. Research into video game playing has found significant differences between the amount of game playing of males and females, with males playing such games significantly more than females (Harris, & Williams, 1985). However some researchers have pointed out that females have as much exposure to the skills required in video games (personal communication Macleod, 1989). Instead the sex differences found in video games may be due to the preference differences the sexes have in feedback. A study which varied the nature of the feedback from a computer based task to subject (Lewis, & Cooney, 1987) found that males performed best when they were given competitive feedback. Females performed best when they were given individualistic (non-competitive) feedback. In empirical research heavy video game playing has been found to correlate with developmental problems, particularly with regard to aggression (Kestenbaum, & Weinstein, 1985). Frequent players were found to have a lower frustration tolerance than the normal population, although perhaps surprisingly they did not have increased neuroticism, social withdrawal, or increased needs to escape into fantasy. Research has also shown that use of video games does not negatively correlate with overall academic performance (Braun, Goupil, Giroux, & Chagnon, 1986). It has been proposed that females dislike the implicit violence involved in many of the video games. Informal studies have found that games which have 'oral aggression', like Pac Man, are much more popular with females (Shneiderman, 1987 p28).

**Cultural or Social Class.**

Very few studies have really investigated the role of gender and cultural effects on computer use. One of these reported that Ss from a low socioeconomic class (SEC) had lower interest in computers and used them less than Ss from a high SEC (Miura, 1987). A strong gender difference was found, which suggests that SEC and gender may have to be considered together. However another study found that Ss from a lower SEC showed more interest in
computers, and wanted to use them more (Pulos, & Fisher, 1987). This latter study also found that most Ss had little interest in computers, and those Ss who did show an interest tended to have higher interests in intellectual activities. These Ss also had a lack of interest in typical adolescent activities. This trend applied equally to males and females, and over all levels of SEC.

**Gender differences in computer attitudes.**

Some studies have shown that there are no differences in the rates of computer anxiety between the sexes (Gressard, & Loyd, 1984). Instead they found that computer experience was a much better indicator of the level of computer anxiety. Most anxiety was shown by those of both sexes who had less experience of computer systems. Confirmation for this finding come from two other studies which used computer attitude scales (see later section on computer attitude questionnaires for more details). First is a controlled study in which students who had to use a computer in an assignment were found to develop more positive attitudes towards computer than those who did not (Eastman, & Krendl, 1987). Second is a study which measured attitudes to computers (Loyd, Loyd, & Gressard, 1987). However this study also reported that girls exhibited a more positive attitude towards computers than boys. This may reflect a change in attitudes (the study is more recent), or may be a biased sample. A survey showed that the careers advice given to school leavers was highly sex stereotyped where computing was concerned (Culley, 1988). Some research has also shown that males and females vary in their attribution of characteristics towards computers (Wise, Robinson-Staveley, & Nelson, 1987). Males tend to use more animate descriptions (less mechanical) as they become more experienced, while females tend to show the opposite trend. They become more mechanical in their descriptions as they become more experienced.

In a survey of 1,600 students gender attitudes towards computers and video games (Wilder, Mackie, & Cooper, 1985), both boys and girls rated computers and video games as being more appropriate for boys, and this trend showed as early as kindergarten. The boys in the

---

1 However there have been reports that levels of computer anxiety were not changed by attendance of a computer literacy course (Electric Learning, 1981). In this study males and females did not differ in their rates of computer anxiety either before, during or after the course. Morris (personal communication, 1989) has pointed out that many of these studies could have had their findings affected by expectancy effects.
survey rated their liking for the computer more strongly than did the girls. Both sexes showed a dramatic decline in their reported liking of computers after middle childhood. The second part of the same study compared the computer attitudes of 334 (141 Female, 193 Male) college freshmen. The sex typing found in the school population was found to be attenuated in this group. However one interesting factor did emerge, that being that males rated their computer expertise as being higher than those of comparable experienced females. This female underestimation of computing expertise has also been reported by others. Researchers who have looked at the subject preferences of the sexes have reported that females were under represented the sciences overall. This was especially bad in the so called 'hard' sciences, such as physics (Deboer, 1986). Both sexes were equally represented in biology and chemistry, and although females out performed males, the females rated their performance as lower than the males.

Many researchers attribute the gender differences to the differential opportunities given to the sexes in education (Chen, 1986). Research has shown that males receive more of the teacher's attention and more computing time than girls in computer based teaching situations (Omerod, 1981; Gerver, 1985). This may not be totally due to a selective gender role bias in teaching, since there is evidence that when girls are given equal access to computers they tend to use them less than males12 (Sheingold, Kane, & Endreweit, 1983; Lockheed, & Frakt, 1984, Miura 1986). It has being found that students (regardless of sex) who show more realistic and investigative personality types, have more positive attitudes towards computers than students showing artistic or enterprising types (Abler, & Sedlacek, 1987).

A study which looked at the sex ratio of 5,533 entries to computer courses at US summer camps found a significant sex difference favoring males (Hess, & Miura, 1985). This 3 to 1 difference became larger as the cost and course difficulty increased. The authors postulated that sex role differences in mathematics, the stereotype of computers, male-oriented software, and the symbolic and viso-spatial features of programming were responsible. The last two reasons are particularly unconvincing since some of the most able programmers have

12 These latter findings have been contradicted in a study which used a large sample size (3,085). In this experiment females given positive discrimination, were found to adopt the use of computers significantly more than boys (Fish, Gross, & Sanders, 1986). Although the use of positive discrimination would not make these results comparable.
often being women. Indeed research has shown that while females are less likely to enroll in programming classes, those who do are successful, and make up the majority in groups who are exceptionally talented (Linn, 1985a). Other research by the same author (Linn, 1985b), has found that females have different strategies for risk taking, help seeking, and more importantly attributions about success or failure. Other research has shown that when the computer based learning strategies of both sexes are compared males tended to adopt the more successful learning strategies spontaneously. This was termed the shifting method of cognitive engagement. Females tended to stick to one form of cognitive engagement throughout the task, regardless of its success (Mandinach & Corno, 1985).

4.2.1.1.1. Summary of the review on gender differences.

As yet there are no universally accepted explanations for the sex differences found in computing. A large (non-computing specific) meta analysis of 127 different studies with a sample sizes from 7 to 25,000, and ages from 2 months to 30 years found that males were more active than females (Eaton, & Enns, 1986). The authors concluded that social influences served merely to increase this natural difference. Certainly there is evidence that social and cultural effects play a very large role. Several researchers have surveyed the sex roles portrayed in the mass media (Braun, Goupil, Giroux, & Chagnon, 1986; Siann, Durndell, Macleod, & Glissov, 1987). These researchers have found that males are portrayed as being the predominant users, and being in a dominant role in any mixed sex portrayals. One other factor in the lack of females in computing may be the male harassment of females who work in traditional male occupations. Research has found that the rate of such harassment is 75%, compared to 50% in the general population (Lafontaine, & Tredeau, 1986). Sources of such harassment included supervisors, peers, subordinates, and clients. This reflects the existing sexist attitudes that too many men seem to hold, viewing females as 'sex-objects' rather than human beings, or fellow workers. The task of addressing such unfair practices will take considerable time, but it can only be hoped that a time will come when such unfair pressures are removed.

3 Although age and situational variables seemed to be equally important.
Changing females attitudes towards computers is one of the most important targets for HCI, and there is hope in this direction. While female attitudes are more negative towards computing than those of males, they have more positive attitudes towards writing (Collis, & Ollila, 1986). These authors have proposed that the use of computer based writing may therefore be an appropriate means of encouraging females to use computers. By trying such measures, and giving positive discrimination it can only be hoped that the male domination of computing can be reduced.

4.2.1.2. Cognitive Style.

Apart from gender, many HCI researchers have looked for solutions to the problems of individual differences in some kind of difference in the way users think. This is usually termed the user’s cognitive style, but is recognised to be the result not just of the inherent characteristics of the user. The cognitive style is influenced by age, education, and the experiential background of the person’s life. Messick (1976) defined the cognitive style as being the high level heuristic that organize behaviour. The possible importance of the concept for HCI is shown by evidence that users perform sub-optimally on interfaces that do not match their cognitive style (Pask & Scott, 1972; Fisher, 1986). It might be therefore surprising that very few studies have empirically tested the effect of cognitive style and computer use (Lusk & Kernick, 1979), though there are signs that this trend is ending (Austin, 1986). Such reluctance is due to the disagreement among psychologists whether cognitive style theories have any validity in HCI (Huber, 1983), and even if they do, factors within the task might outweigh the role of cognitive style anyway (Payne, 1982, Van der Veer, Tauber, Waem, & Van Muylwijk 1985). An example of the role which individual differences might have has been shown by research which showed that people with high self-esteem made fewer errors, and were more likely to give negative feedback to a computer system which gave simulated human responses when compared to usual computer responses (Resnik, & Lammers, 1985). However this effect was not reversed for those with low-self esteem.

14 Morris (personal communication 1989) has pointed out that even if task factors did outweigh the effect of cognitive style this would not be a reason not to study it. However HCI research has become highly pragmatic, and if an area does not seem to hold promise for future commercial application it will not be investigated.
Psychological Differentiation.

One of the most popular cognitive theories in HCI is called psychological differentiation (Witken, Dyke, Faterson, Goodenough, & Karp, 1974; Goodenough, 1976, Witken, Dyke, & Oltman, 1979; Witken & Goodenough, 1981). This proposes two main cognitive styles, field independence and dependence. Field independence is characterized by good analytic and restructuring skills. It is presumed to involve more participatory and hypothesis testing activity when faced with new situations. In contrast, field dependence takes a more holistic approach, which relies upon the inherent organisation of the material. In the face of a new situation, field dependent characteristics are likely to adopt a spectator type approach. The field dependent person would be more person-oriented, while the field independent would be more task-oriented. There is some empirical evidence that women are more field dependent than men. Some researchers have speculated that (within computer based tasks) field dependent users would prefer a system guided form of dialogue (Kupka, 1976), using a formal language context (question and answer systems). The field independent user would prefer a user guided interaction system which encouraged exploration, and flexibility (command line systems, or complex DM environments). There is some experimental evidence that users performed their tasks faster in environments which conformed to the requirements of their respective cognitive style (Fowler & Murray, 1987).

If these particular theories have any validity they should be carried over into computer manuals and on-line help systems. The field dependent user would find manuals which took a cooperative style better to use (Kennedy, 1974). In contrast, the field independent user would find a subordinate style preferable. However, it is probably dangerous to make the user models so hard and fast. In reality, it is likely that some form of individual configuration is the ideal method.

---

5 These tests have been done using an embedded figures test (EFT). It has already been noted (Sherman, 1967) that the EFT discriminates in favor of viso-spatial skills. This would mean that the EFT favored males anyway, and that any results that used the EFT to discriminate between the sexes would be suspect (Fowler et al., 1987).

6 These findings should be taken with caution, since they are based on post-hoc observations.
4.2.2. Poor design.

Many leading workers in the field of HCI believe that unfriendly computer systems are purely a matter of bad design (Norman, 1988). This links to some extent with the historical and sociological aspects (see below), so designers and users have tended to be computer scientists, or engineers. These groups have typically demanded systems that delivered the maximum amount of functionality at the minimum cost. Often this has been a recipe for producing systems which have only been useable by people with extensive training (or patience), thus creating a variant of the 'Real Programmer' philosophy. This 'Real Programmer' philosophy is from software design and is where programmers took the attitude that if the code was hard to write it should be hard to understand! The design of the user interface has often been relegated to the last and least important part of any software project. Thankfully this trend may now have changed, and 'usability' and 'user friendliness' have been found to be good commercial selling points.

4.2.2.1. Historical And Sociological Background To Poor Design.

What is 'bad design' for one user group may be perfectly acceptable for the group of users who have been historically involved with such methods. The early computers were designed and used by and for mathematicians and engineers. For the most part these computer systems were designed and built for the solution of specific mathematical problems. For instance the computer often quoted as being the first digital computer17, called ENIAC (Electrical Numerical Integrator And Calculator) was designed and constructed18 to calculate the trajectory tables for military projectiles. The user interface of ENIAC was a set of wire connections that had to be manually 'hard wired' into the required configuration. Such interfaces would only be suitable for the engineers who were familiar with the internal construction of the machine19.

17 This has been called into question by the discovery of the earlier work of John Atanasoff at Iowa State College (a decade earlier than ENIAC). In 1971 a US patent judgement went in favor of Atanasoff’s earlier machine.
18 Constructed at the Moore school of Engineering, University of Pennsylvania by John Eckert and John Mauchly in 1942.
19 Macleod (personal communication, 1989) has noted that the gender ratio of these early programmers was quite often highly biased towards females (due to war time job allocations).
For 20 years after ENIAC the methods of controlling computer systems continued to be designed and used by similar users, who were highly specialised people. Users tended to share a common technical background, and common concepts, which made the adoption of restricted interaction conventions acceptable. For example the earlier computer languages were all either written for business (COBOL - which is largely based on English language concepts), or for scientific users (FORTRAN & ALGOL - which use mathematical concepts). The differences between the two languages reflected the differences between these separate target user groups. These factors are all that are needed to build the important features that modern HCI has to battle against, sub-cultures, and established user communities who form, or actively train (indoctrinate) future users. The interface concepts that are taught first have the greatest impact upon the demands and expectations of future users. Future users in one user-community are likely to demand 'downward compatibility' in any future developments which a computer manufacturer makes. This restricts the amount of improvements available to a product once it has been launched, and the number of innovations launched by a single computer company. The user attributes of the target population involved in the original development of a concept often then plays a role in the continuation of the usability attributes of one type of computer interaction. A prime example of this is the split between users of so called graphically driven (direct manipulation) software, and those of command line systems. This particular group divergence is of particular interest, since it forms one of the reasons this dissertation took place. In the next section these two interaction methods will be discussed in greater depth.

4.2.2.2. WIMPS and Wizards²⁰.

An example of the functionality philosophy would be the text editor MicroEmacs, which has a function attached to almost every key combination. For example pressing the Escape key and then the Control and 'H' keys deletes a word backwards. Pressing the Escape key and then the control key and the 'R' key starts a prompted forward search and replacement of selected text items. Obviously MicroEmacs has a very high degree of functionality, but cannot be considered to be very easy to learn. Certain types of users demand as much

²⁰ These terms refer to the names commonly given to users of direct manipulation and command line operating systems respectively.
functionality as a system can give, and are simply not bothered by the amount of time it takes to learn the system, or the complexity involved. In fact many HCI researchers suspect that very complex systems which literally require years to master have an appeal of their own, which sets those who can master such systems in their own social elite. Complex systems have their own 'Command Line Macho' (Shneiderman, 1987 P147), which is an important aspect for products being aimed at specialised user markets. These power users traditionally dislike the DM interface, and will often state that command line interfaces are superior to their DM alternatives. This superiority is often expressed in terms of increased speed, and better 'feel', and more 'power'. Power and feel are relative terms, but for these users they probably refer to the faster perceived command rate, and ability to be initiating commands rather than responding to the computer. These power users have tended to be in positions of seniority within computer science, it is for this reason that the computing industry adopted the term 'WIMP' (Windows Icons Mice and Pull-down Menus\(^1\)) to refer to DM interfaces. It is hard to know why these command line users have such a low opinion of the DM interface, and its users. One might propose that DM interfaces permit non-members of a select group to have access to the same functionality, without the years of 'apprenticeship'. The reason could also be historical. If as was suggested in the review of DM interfaces they developed from computer games, it is possible that a prejudice remains from those days. It should be mentioned that the low regard which command line users feel for DM users is reciprocated by the typical DM user. The prejudices between the two user communities have tended to continue, although since this dissertation started\(^2\) (1986) more hard line command system users use a DM environment if the system provides a high degree of functionality, or is complex. Xerox Ventura Desktop Publisher is a case in point.

4.2.2.3. **Empirical Comparisons between Interfaces.**

Surprisingly little if any empirical research has been done to compare these two interaction methods. As the review on interface comparisons (see next section) shows most interface

---

1. Some of the problems within HCI are highlighted by the fact that even the derogatory acronym 'WIMP' has no one standard definition. Among its other (publishable) definitions it also stands for Windows Icons Menus and Pointers.

2. The hard line 'command line machos' the author comes into contact with may have had their perceptions changed due to the results he obtained in his research; or they might just be being polite to him!
research work has tended to concentrate on testing variants of one particular interface. Often these interface tests have been done conceptually, by Ss who have been asked to imagine they are using the interface being tested. The Ss’ actions in these tests are confined to paper and pencil responses (Carroll, 1982; Green & Payne, 1984), or selecting the correct commands from a list of possible answers (Barnard et al., 1981, & 1982). These somewhat unrealistic testing conditions have been used because of the large amount of work and expertise involved in implementing real versions of the interfaces to test them. Even those HCI laboratories who have the resources to conduct the highest levels of research have considered the amount of work involved in designing and implementing just one version of a system (without any experimental evaluation of that system) sufficient for the award of PhD (Iseki & Shneiderman, 1986). This may be the reason why so few direct WIMP and command line comparisons have appeared in the literature. To enable proper evaluation of the differences between command line and DM interfaces two functionally identical systems would have to be created, one with a command line interface, and the other with a DM interface. This system would then have to be used in controlled conditions by Ss from selected user groups, and the results compared. Before the study described in this dissertation, this kind of detailed evaluation between DM and command line systems had never been attempted in the field of HCI.

4.2.3. Past Interface Comparisons.

In the past, interface evaluation and comparisons have mainly concentrated on investigations into aspects of specific interaction methods. These investigations have included studies of UNIX based errors (Kraut, Hanson & Faber, 1983; Hanson, Kraut, & Farber, 1984), the investigation of different command line keyword styles (Ledgard, Whiteside, Singer, & Seymour, 1980), the comparison of sub-versions of the text editor EMACS (Green & Payne, 1984), the investigation of different versions of a robot control command line system (Carroll, 1982), and the study of the organisation of menu items (Liebelt, McDonald Stone & Karat, 1982). In comparison the field has spent very little effort in direct experimental comparison between systems that are functionally identical but have different user interfaces. This could be due to several factors, the most important of which have

23 To be functionally identical two interfaces must have identical operational capabilities.
already being discussed, such as the time and expense involved (Iseki & Shneiderman, 1986). The value of such comparisons can be seen by the relevance of one of the earliest such experimental comparisons (Roberts, 1980), which found that the full page editors were twice as fast as identical line editors. This finding paved the way for the whole WYSIWYG (What You See Is What You Get) concept, which is where the computer display shows the exact form of a document before it is printed.

4.2.3.1. Direct comparisons which have been conducted.

A large proportion of the direct comparisons that have been done are between Natural Language Interfaces (NLI) and their alternatives. Comparisons between NLI and the relational database command language SEQUEL\textsuperscript{24}, (Small & Weldon, 1983, Jarke et al., 1985) found that the command line SQL system showed a faster performance in data retrieval. Another database study (Hauptmann & Green, 1983) compared NLI, command and menu systems. They could find no difference between any of them in speed, errors, or user attitudes. A comparison between a form-filling interface and an identical command line system (Ogden & Boyle, 1982) found the form-filling interface to be faster, and preferred by 11 of the 12 subjects tested. Only such direct comparisons allow researchers to judge different interaction methods against one another in a way that is impossible with detailed single interface evaluations.

It was for this reason, and because of the lack of clear, direct comparisons between WIMP and command line systems in the current HCI literature that the HCI part of this study was

\textsuperscript{24} Usually known as SQL, and designed by IBM (Chamberlin et al., 1976)
undertaken. To address this question a computer system was created to compare computer users’ performance and preferences in both command line and DM interfaces, using a specifically designed database query system as the target application. Since this system had to be designed and constructed from scratch it meant that the parapsychological elements could be designed into the structure of the system. This gave a great degree of freedom to the study since the designer could choose which parapsychological AHCI questions it asked and how it asked them. It also freed the study from the constraints of copyright that would otherwise have limited the implementation and use of the system. The disadvantage to creating such a system was the amount of effort (so called ‘Man hours’) involved in the design and creation of a system which needed over 71,000 lines of high level program source code.

4.2.3.2. The KMDB system.

Having reviewed the aspects of computer interface design we can now review the design of the system used in this dissertation. The system had to be able to be used by students of both sexes, and from different subject areas, at a local polytechnic.

4.2.3.2.1. Target Subject Populations

The study was interested in looking at the interface preference differences between Ss in three groups: naive keyboard illiterate, naive keyboard literate, and sophisticated. Since these represent three very different populations, each with different abilities (for example,
keyboard skills or computer experience) it was felt they would best highlight the usability
criteria of the two interface types. We were interested in which interface would be preferred
by which subject populations, which interface would be most productive for which subject
groups, and what the error rates would be like over all subject populations. It was hoped that
the answers to these questions would give clues as to which interface was best, for which
tasks, and for which user groups. Given these considerations it was necessary to ensure that
the task was one which would be easy for the subjects to understand, and yet still be able to
be interesting for quite sophisticated computer users. The specific choices available were
constrained by the parapsychological element of the study. Part of the parapsychological
requirements were to evaluate smart noise, and function linkage in the same test environ-
ment. The rationale behind the parapsychological aspects of the task are fully described in
chapter five.

Since the study intended to use students from various academic subjects at Napier Polytech-
nic as Ss, it was important to choose a subject material which all students would find easy
to understand. After considerable thought a database concerned with the running of a college
was chosen. Ss from all academic subjects, races, SECs, and of both sexes would be equally
comfortable with such a task. The experiments would take the form of Ss performing
problem solving exercises as part of their course content. Another consideration was that
the system would be used by Ss with different levels of computing expertise. It was therefore
essential that the system be simple to use (each session only lasted 1 hour), and yet powerful
enough to still be interesting to the more experienced users within the same 1 hour period.
The two versions of the database system also had to be exactly matched for functionality
and as far as possible error opportunity. To ensure that these error and command oppor-
tunities were essentially equal did not just involve interface design.

A special tutorial which instructed the subjects about the database, and what actions to
perform also had to be created. It was important that all groups attempted the same sets of
commands in the same order. This was not only to enable the HCI study to compare
performance; it was also very important that the parapsychological tasks were identical for
each subject. This meant ensuring that subjects all attempted the same trials in the same
order. The author determined the sequence of parapsychological trials, and in collaboration
with SG determined the best order for the tutorial. However, since it was SG’s responsibility
to ensure the students received adequate academic benefit from the sessions, the implementation and layout of the tutorial sheet had to be her responsibility. Without these considerations it would be impossible to compare the differences in the various S groups over the experimental series.

4.2.3.2.2. Design of commands and syntax.

On the basis of an analysis of the requirements for classes in database use at Napier over various levels of expertise, and some informal testing sessions at the Edinburgh psychology department, a reduced set of simple commands were developed. These commands are both easily learned (target from the informal tests were 1 hour to competence, to reduce total session length), and enable sufficiently complex actions to be carried out so experienced users could still use the system. Competence was defined as the subject having gained an understanding of the database's various commands, so they were able to perform the problem solving exercises at the end of the tutorial.

Comparison of the interfaces.

The command line and Wimp interfaces had identical command syntax parsers, and functionality. This complete matching was implemented from the command syntax right

---

8 The particular WIMP standard adopted for the KMDB system was Digital Research Inc’s (DRI’s) Graphical Environment Manager (GEM).

9 A syntax parser is a software routine which translates the user’s command into the required series of operations.
down to the exact wording of error messages, and help facilities. The Wimp system comprises a graphical interface with the standard windowing layout of pull down menu bars (see figure). As an example of how closely the two systems matched each other the delete command will be illustrated. To delete in the command line system the following syntax would have to be typed by the user.

```
delete record-type key= key-number
```

An example of this would be 'delete student key= 1'. In direct comparison, in the Wimp system the user would pull down the options menu, and scroll down until they reached the 'delete...' entry and click the mouse to indicate they required that action. This is analogous to typing 'delete' in the command line version. The Wimp system will then present the user with a dialog box (see figure), which prompts for the record type and the key number of the record to be deleted. Clicking on the 'ok' radio button (when the user is satisfied with their entries to the dialog box) is equivalent to pressing the return key in the command line system. Both systems then prompt the user with identical wording which forces the user to confirm this delete action.
Two performance measures were used:

1) Errors. The precise part of the command that caused the rejection was recorded by the system (on each separate interface), with details of the length of time between each error, and the average time between each error. Further, the type of the error was noted (the type and sequence of commands to be issued are preset, and will be identical for all users).

2) Commands. The system recorded all of the commands actioned during each session (on each separate interface), with the length of time between each command, and the average time between each command.

4.3. Computer Attitudes.

The research area concerned with people’s attitudes to computers is rapidly growing. A comprehensive review of previous research and current trends exceeds the space available in this chapter. Therefore what follows is a brief overview of some relevant studies. Readers who are interested in this area should consult the bibliography for a more detailed list. For the sake of clarity this review will subdivide the area of computer attitudes into those studies which have looked at the public image of computers, which have looked at problems associated with their introduction, and which have looked at designing computer attitudes questionnaires. This last group will be reviewed with particular emphasis on the fear of computers.

4.3.1. Public Image of Computers.

The problems involved with conducting any kind of review of public views about anything is that they can never accurately reflect the current views. Such reviews can only hope to be of an historical nature. The earliest survey of public feelings toward computer as was conducted in 1963, but only reported several years later (Lee, 1970). In this American nationwide survey two main beliefs about computers emerged. First that computers were useful tools, and second they were autonomous and would eventually control society. The
latter view probably reflected the early 1960s science fiction portrayal of computers. However, one of the leading computer scientists (Minsky) working in the field of artificial intelligence (AI) has also stated that eventually computer systems would treat humans like pets. Given the current state of AI this stage is unlikely to be reached within this century. To get an idea of how the public’s feelings to wards computers had changed over a 4 year period Ahl (1975) compared the attitudes towards computer reflected by a ’Time’ Magazine survey in 1971, with those in ’Creative Computing’ in 1975. Obviously such a comparison is not ideal simply because the two target populations are not identical. However, ignoring these biases the main results which Ahl reported were that readers in both groups tended to be becoming more worried about the impact of computing upon jobs. Readers tended to be also concerned about the de-humanising effect which they perceived the computer had upon society. Overall the readers in both groups were optimistic about the use of computers in all areas of life, except credit card data banks (Ahl, 1975 p77). Ahl’s findings reflected the worry which many rights activists have voiced about the abilities of governments, and large corporations to gather huge amounts of data on individuals. These data can then be sold (unsolicited mailing lists), or used to monitor target individuals. Most countries have established acts to try to restrict these dangers, but it is debatable as to how effective these measures have been.

The use of computers expanded enormously in the 1980s. Many professions which had been considered traditional strongholds against technology were forced to adopt computerisation (Downs, 1985). This revealed that many of the people who had entered those occupations had done so explicitly to avoid dealing with technology. These people became the subject of academic study, in attempts to understand and help their condition. One of the earliest researchers in this field has been Weinberg. In a series of publications on ’techno-phobia’, he and some co-authors reviewed their findings from a 4 year study with 500 college students and business managers (Weinberg, 1982; Weinberg & English, 1983; Weinberg & Fuerst, 1984). The problems and research into computer phobia is a specialist area, and is covered separately (see below). The problems involved with having groups of the population that

---

31 The Time reader will tend to be rather business oriented while the creative computing reader will tend to be already interested in recreational computing.

32 The later publications by Weinberg are based upon his 1982 study.
avoid using computers are enormous. In a more recent Harris poll (1983), 90% of the population responded that they felt that computers helped to improve the quality of life. Those people who classified themselves as being 'computer literate' tended to be younger professional workers with college degrees. If his trend continues then computers will exaggerate the difference between the life style and opportunities for rich and poor families. This fear was reflected in the Harris poll by the percentages of people who reported feeling that computers had done more good in the past (83%), and those who felt this trend would continue (68%). Half the respondents felt that computers were a threat to individual privacy. These concerns must be addressed and corrected by governments and researchers quickly, before such worries become a reality. In contrast some researchers have speculated (personal communication Morris, 1989) that as computer based sharing of information becomes more widespread the computer could have the opposite effect, that is, it could cause a fairer redistribution of power among mankind. Unfortunately the author of this dissertation believes that it is more likely that mankind will probably repeat the circumstances which occurred in the first industrial revolution, where a few wealthy people owned the machinery, and the majority were reduced to extreme poverty. In the case of the computer revolution, mechanical power (which was the basis of the first industrial revolution), might be replaced by power derived from information. However, worrying though these possibilities are they do not fall within the remit of this dissertation.

4.3.2. Problems Associated with the Introduction of Computers.

As we have already discussed, the reactions of many people when they are confronted with using a computer system are diverse. They include the emotions of delight and fear. In this section we will discuss the latter problems areas. The recognition of computer phobia has come relatively recently, and has often being accompanied with the use or development of specific questionnaires to investigate computer attitudes. Those studies which have used such methods will be discussed in a following section. Instead this section will briefly overview the computer anxiety problems that have been reported in the literature.

The pioneers in this area were Rohner and Simonson (1981), Jordan and Stroup (1982), and Maurer (1983). These researchers found that workers can express fear when they are forced to use a computer. Researchers have found that common fears which new users have are that they will damage the computer, make costly mistakes, look stupid, work with a powerful
other, and that they will lose control over their working style (Bloom, 1985). These effects cover the entire work-force. Even managers have to accept changes in the way they make their decisions, which can result in problems if they feel the computer affects their 'locus of control decision-making authority' (Fleischer & Morell, 1985). It was Maurer (1983) who first introduced the term 'computer anxiety', as a recognised condition. Another important researcher in this area is Weinberg (1982); his study with 500 Ss has already been discussed (see above). These workers proposed some practical methods to reduced this phobia as well as addiction, both of which should\(^{33}\) be equally undesirable (Davidson & Walley, 1984). These measures include involving the user in the design decisions\(^{34}\), training courses, and individual one to one practical sessions with computer systems. Often computer games are very useful tools in breaking down the initial reluctance of people to use a computer system (Bloom, 1985).

Once people have seen that they do not break the machine when they use it many quickly overcome any reluctance to use the new equipment. The other important factor is the continual monitoring of the users satisfaction during the use of the system. Computer addiction is also a problem, and has received even less attention than computer phobia. One of the few researchers to address computer addiction have been Davidson and Walley (1984). They recommended the use of management skills, relaxation, stress management regimes, and as an extreme measure even the use of behavioral modification and Hypnosis programs (Matez, 1986). It has been found that the computer phobics can improve their computer learning if they can reduce their bodily stress, and distracting thought patterns. Such thought patterns often form part of a larger belief about the person’s self image. Fortunately positive (non-anxious) patterns can change that self image. However changing self image is a long process involving experience of success at the target tasks (Meier, 1985), and this can be achieved by education, and skill acquisition. These can begin after the reduction of stress, which restricts the attention and metal processes (Kahneman, 1973). Recent developments in applied industrial settings have included formal methods to help in the introduction of new systems (Hatcher, 1987). This includes questionnaires to

\(^{33}\) Unfortunately computer addition among computer workers may be regarded as an advantage by some employers.

\(^{34}\) See the work of Mumford (1970), Lang (1971), and Hall-Sheehy (1986).
highlight possible trouble spots, and some decision rules, which have been developed through experience. Training course which allow potential new users to experience computers in non-threatening situations have also being found to be very useful in reducing anxiety (O’Quin, Kinsey, & Beery, 1987).

4.3.3. **Computers in Education.**

The problems involved in the introduction of computers are not confined to industry. Within education severe problems have been encountered (Coleman, 1986). Many teachers have had little or no experience with new technology, and the problems involved with reluctance to be seen not to know how to use a computer is perhaps as acute as it is for the senior manager. For the students the problems have been found to be related to resources and the gender problems (Chen, 1986) discussed in the individual differences section (see above). Efforts are still being made to remove the gender difference that exists between the sexes in their acceptance of the computer (Gustafson, 1985). Research has shown that the negative attitudes of students is a major factor in deciding their preferred learning situation (Bird & Chung, 1986). Within educational settings computer anxiety has been found to correlate more strongly with self reported computer behavior than with any other personality or attitude measure (Morrow, Prell, & McElroy, 1986). This finding has been interpreted as suggesting that in educational settings familiarity may be the most effective single measure available to reduce computer anxiety. Courses which have computing content have been shown to have the highest levels of class number withdrawals than any other courses (Sproull, Zubrow, Kiesler, 1986). Theses comparisons were made at both teaching and research based universities. Both sets of students reported that courses with computing content caused more anger, reality shock, and confusion than other topics (Sproull, Zubrow, Kiesler, 1986). The withdrawal and psychological effects were more pronounced for
students from the social sciences and liberal arts. There was a significantly larger number of withdrawals from any course with a computing content by female liberal arts students, when compared to non-computing topics. The worst course judged by dropout rates was perhaps surprisingly computer science itself. This dropout rate was unlikely to be directly linked to high amount of mathematics, or complex subject material, since equally complex subjects such as pure mathematics and physics did not show such a high level of class withdrawal. Some studies have suggested that computer science as a subject requires too high a degree of critical thinking ability, mathematical and self awareness skills (Whitney-roper, 1986).

Some universities in the US have required their freshmen to acquire a computer. Research in these situations have shown that most resistance comes from members of the respective staff who knew least about the computers (McCord, 1985). Yet surprisingly these same staff members were the most likely to make an effort to enhance their computer skills. In non enforced situations the most important factors in deciding on the purchase of a personal computer has been found to be social interaction, demographic location, and group membership (McQuarrie, 1985). McQuarrie felt that so called macro-social factors might the determining forces behind all of these other factors. Further research into the factors which determine which members of a uniform group acquire a computer is still a topic of research (see the work of Scott, 1986, and Koohang, 1987).

35 Macleod (personal communication, 1989) has commented that students with no experience of what academic computing involves do not have a realistic expectation based upon their hobby based computer experience. Macleod further postulated that the reasons for dropouts from courses with a computing content might be very different to drop out reasons in pure computer science majors.

36 The author has a background in computer science and is therefore biased.

37 Further research into the gender differences revealed that males were more likely than females to associate negative attitudes towards mathematics with similar attitudes to computers (Collis, 1987). Collis felt that Negative attitudes towards mathematics was unlikely to explain the high dropout rate of females from computing subjects.

38 This is certainly not the personal experience of the author, however this is a controversial area, and certainly merits further research.

39 To exclude any educational level bias that is already known to exist (Dutton, Rogers, & Jun, 1987).
4.3.4. Computer Attitude Questionnaires.

This section will summarize the major studies that have investigated the computer attitudes, with an emphasis on users who are computer phobic, or have high degrees of anxiety towards computers. The pioneering work in the area of computer phobia was conducted by Jay (1981). He defined computer phobia as having 3 main symptoms. These were resistance to talking about computers, fear and or anxiety towards computers, or hostile and aggressive thoughts about computers. Studies which have used questionnaires to survey such aspects as intensive video users (Kestenbaum, & Weinstein, 1985; Wilder, Mackie, & Cooper, 1985; Braun, Goupil, Giroux, & Chagnon, 1986), or gender differences (Hess & Miura, 1985; Collis & Ollila, 1986; Deboer, 1986, & 1987 Miura) are treated elsewhere.

Some of the earliest research in the area of computer attitudes questionnaire development and use was conducted within the medical community (Reznikoff et al., 1967), and the medical community continues to be a source of computer attitude research (Shannon, 1987). Reznikoff et al. found that computer attitudes varied with sex, increased age, and level of education. These findings were replicated by Klonoff and Cambell (1975). They developed an 88 item computerization attitudes questionnaire. This was administered to the 42 staff members in a psychiatric unit of a hospital where a computer system was being installed. 28 of the staff were sent on a 21 day course, the remaining staff did not. Results showed that those staff members who went on the course had a more positive attitude towards the new system.

The first truly specialised questionnaire to look at computer attitudes was the ATtitudes Towards Computers Questionnaire (ATTCQ) (Raub, 1981). Other researchers before this time had used standard tests of satisfaction like the Minnesota Satisfaction Questionnaire (MSQ) to measure users responses to the introduction of a new computer system (Cheney, & Dickson, 1982). This was useful to the extent that it measured the user’s satisfaction. Its problems were that it did not record what aspects of the system caused the satisfaction, and dissatisfaction. Other typical measures used in early computer attitudes work were mood

---

The author of this dissertation believes that it is possible that the 14 people concerned resented that they had not gone on the course, and that this was responsible for this finding.
and arousal questionnaires (Cruickshank, 1982), or Rotter’s Internal-External Locus of Control Scale (RIELCS), and the Interpersonal Trust Scale (ITS) (Kerber, 1983). Kerber reported that no meaningful correlations emerged from either the RIELCS or ITS. The only useful data which tended to emerge from these early studies was that experience with computers played an important part in computer attitudes. These early workers knew that the introduction of computer systems was causing distress, but they had not yet developed the appropriate measures to enable them to investigate and start providing solutions.

The few researchers we have mentioned above had little experience in dealing with situations where working practices were being so radically altered (by what were often inappropriate solutions) in such a rapid manner. It can be argued, as we have previously, that the only equivalent situation to the rapid changes introduced by computerization had been during the industrial revolution. The second factor (of equal importance), was that these changes were affecting so many aspects of the working environment simultaneously that initially it was hard to determine which features of the changes were responsible for the problems. This made it difficult to start providing feedback to reduce anxiety, and improve the design of future computer systems. There was also a distinct lack of communication between computer system designers, and these early computer based human factors workers. Unfortunately this was almost totally on the side of the computer designers, who were content to continue producing the same kinds of systems they had been producing for engineering solutions (see the problems we mentioned under bad design, above).

Raub’s (1981) work with the ATTCQ enabled the first definition of computer anxiety to be made, as a form of state anxiety in which the computer formed the ‘personality threatening’ stimulus. Computer anxiety was found to be a multidimensional construct composed of computer appreciation, computer usage, and general anxiety concerning the possible nega-

---

41 A review of anxiety measures (Cambre, & Cook, 1985), found that it was hard to validate any measurement of anxiety, since it is hard to define in an empirical manner. The only way to investigate it (without physiological measures) is to correlate scores on an anxiety measure against variables assumed to reflect anxiety.
tive impact of computers on society. Raub’s research showed that academic major, experience, and gender were the major correlates of computer anxiety. Past experience and mathematics anxiety were found to be the strongest factors.\textsuperscript{42}

One other study which used Raub’s ATTCQ investigated the computer anxiety of managers from service and manufacturing organizations (Howard, & Smith, 1986). They reported finding that only a very few managers were computer phobic, that math anxiety was related to computer anxiety, that neither sex nor age were relevant factors, and neither was locus of control, trait anxiety or cognitive style. However they did find that computer anxiety was related with previous experience of computers, as opposed to level of knowledge which was unrelated. They also found that anxiety towards microcomputers was linked to anxiety about the impact of computers on society at large.

One of the other measures of computer anxiety has been the Computer Attitude Scale (CAS) (Gressard & Loyd, 1984a, & 1984b). The CAS was created on the basis of a comprehensive review of the use of questionnaires in computer attitudes. It has 3 sub-scales which assess computer liking, confidence, and anxiety using a Likert attitude measurement format. Research with the CAS has shown that computer anxiety is negatively correlated to computer confidence, and moderately (positively) correlated with mathematics anxiety (Gressard & Loyd, 1984a; Fisher, 1986). This suggests that the 3 sub-scales may not be independent, as these may all be measures of computer anxiety (Gressard & Loyd, 1984a). Subsequent research with the CAS (Gressard & Loyd, 1986) has shown that computer anxiety was especially likely in academic situations, were the S did not like, or have confidence with the computer. This study also found that mathematics anxiety was moderately associated with computer anxiety. Gressard and Loyd suggested that exposure to computers in a non-academic (non-threatening) way might be helpful in reducing some S’s anxiety. These findings with the CAS were followed up in the following year (Loyd, Loyd & Gressard, 1987). In this study the authors reported on a survey of 561 seventh to eighth grade students. Past experience with computers was reported to be the strongest correlates to computer anxiety,

\textsuperscript{2} Raub also reported that on a state-trait anxiety scale, males were likely to have more computer trait-anxiety than females.

\textsuperscript{3} The N involved in this study makes it too small to take Howard, & Smith’s findings as anything but anecdotal.
with less anxiety being associated with greater confidence, and liking of the computer themselves. Loyd et al. reported that in this sample females exhibited more positive computer attitudes than did males.

Another measure of computer attitudes which has been developed is the Computer Anxiety INdex (CAIN) (Maurer & Simonson, 1984). The CAIN was developed in conjunction with the Standardized Test of Computer Literacy (STCL). It has measures to investigate caution with, negative attitudes towards, and disinterest in computers.

Another specially developed computer questionnaire is the Cybernetics Attitude Scale\(^\text{44}\) (Wagman, 1983). Wagman’s use of this scale looked at S’s attitudes towards the introduction of computers in 10 sectors of society. Examples of these sectors include criminal justice, mathematics, statistics, counselling, and medicine. Findings showed strong gender differences and that computers are not generally welcomed in counselling and medicine.

Another questionnaire which was called the Computer Attitude Scale (CAS) was developed which evaluated positive and negative attitudes towards computers (Nickell, & Pinto, 1986). This was a 20 item Likert-type scale, and was used in 5 separate US universities and colleges. However the data from these 583 Ss only showed a positive overall attitude towards computers, and the usual male gender bias towards computers.

An investigation into 737 students asked them to complete a semantic differential test on their attitudes towards computers which revealed 6 major dimensions in a factor analysis (Arndt, Clevenger, & Meiskey, 1985). These dimensions were: Pleasing vs Unpleasing; Helpful Vs Ineffective; Good vs Bad (effect on society); Dominant vs Submissive; Easy vs Hard; and Warm vs Cold. The authors reported that those Ss who were experienced with computers saw the computer as warm, pleasing, submissive, effective, and easy to use.

The longitudinal effects of computer introduction in an office environment were tested by computer attitudes questionnaires by Gutek and Bikson (1985). The authors got 530 employees (45% male) in 26 different private sector businesses which had just introduced

\(^{44}\) Not to be confused with Gressard & Loyd’s CAS (1984).
computers to complete a computer attitudes questionnaire. A year later a follow up questionnaire was completed by 232 Ss (56% male). Although the combined results showed the traditional male domination of skills, authority, and computer based decision making, analyses that controlled for job category showed women were more satisfied with computerization. Males reported more problems with regard to computer access, software, and help.

Other lesser known measures include the 64 item Computers and Robots Attitude Questionnaire (CARAQ) (Moore, 1985), which was tested on 994 13-16 year olds. It was reported to be a valid measure of the attitudes of such a group, on the basis of the strength of reliability coefficients from a factor analysis. Another little known questionnaire was developed and used specifically for identifying computer anxious students (Campbell, & Dobson, 1987). This was found to be a particularly useful tool for pinpointing students’ who were experiencing problems with computers due to their high degrees of anxiety. Subsequent research which has looked at the factors in computer attitudes which predict students computing performance have found two major factors (Koslowsky, Lazar, & Hoffman, 1988). These were viewing the computer as a challenging, or controlling device. The view, of the computer as a challenge, was found to be highly predictive of the students future computer based activity, and attitudes.

Multi-racial studies of computer attitudes have been rare. One large Canadian study (Collis, & Williams, 1987), compared 2,105 Canadian and Chinese students’ attitudes towards computer studies, and other selected subjects. The Chinese children were significantly more positive towards science, computers, and writing than the Canadian students. The Chinese students had fewer age or sex differences with regard to computer attitudes. The authors reported that females from both cultures reported that females were equally competent in computer expertise. However males from both groups felt that males had superior computing abilities.

Another cross cultural computer attitudes study compared computer attitudes between the US and Canada (Gattiker, & Nelligan, 1988). In total 306 Canadian and 157 US employees were surveyed for their job satisfaction, work effectiveness, and locus of control. The experimenters postulated there might be some differences in these factors with Ss according
to gender, authority position, and type of computer system. Results showed some effects in these predicted directions, but they were not distinct enough to allow any hard conclusions to be drawn.

Very few questionnaires have been developed which investigated the cross relationships among mechanical competence, computer anxiety and competence. One of the exceptions is the 20 item Computer Anxiety Rating Scale (CARS) (Heinssen, Glass, & Knight, 1987). A trial experiment with the CARS with 270 college students revealed that computer anxiety was not only correlated with mathematics and test anxiety. It was also found to correlate with less interest in machines, and less experience with computers. The authors noted that during actual computer interaction, anxiety was associated with lower expectations, poorer task performance, greater state anxiety, debilitating thoughts and physiological arousal45. They concluded that these findings were 'consistent with a cognitive-attentional theory of computer anxiety' (Heinssen, Glass, & Knight, 1987).

A report on 5 studies of computer phobia with a specially designed computer phobia questionnaire revealed three nearly independent measures of computer phobia (Rosen, Sears, & Weil, 1987). These were called computer anxiety, computer cognitions and computer attitudes respectively. The authors reported that while older Ss tended to be more anxious about computers they did not have more negative feelings towards them. However they did find that females had more negative attitudes towards computers than men. Also they reported that computer phobia was related to other anxiety measures like mathematics and state-trait anxiety, but was a separate construct. Finally they reported that experience with computers did not reduce anxiety46, or improve computer attitudes47. This is in direct

45 These are all characteristics which would be assumed to be found with the MLP syndrome (Morris, 1986).
46 This is interesting and contradictory, because in a separate paper (commissioned by the US Department of Education to study computer phobics) the same authors used a computer skills module in their treatment schedule. They also claimed to have found a 'significant pre-post-treatment change in anxiety, attitudes, cognitions' towards computers due to this treatment (Weil, Rosen, & Sears, 1987).
47 Hypnosis and biofeedback have both been found to be effective in reducing computer anxiety (Farnill, 1985).
contradiction to most other work in the field, even the most recent (Dolan & Tziner, 1988).

4.3.5. Discussion.

One of the most striking things to emerge from this review of computer attitudes, and implementation is the range of contradictory findings. Some researchers found that females had negative attitudes towards computers, and others that they had positive, that attitudes towards computer systems were related to anxiety, and that it was not, that computer anxiety was linked to mathematical anxiety, and that it was not. The findings are confused. What is the researcher to make of them? The author of this dissertation believes that these contradictions (which also occurred within the attitudes, gender and individual differences reviews) show there are large geographic and social effects which are more powerful than any of the other considerations. Macleod (personal communication, 1989) has agreed with the author, that the research into these factors seems to have been an exercise in categorization, rather than something which can be applied to solving real world problems.

We can bring in parapsychology here for a moment. From the time of Rhine, researchers have often concentrated their efforts on getting only positive results. Indeed Rhine advocated a methodology in which negative results were to be ignored (Rhine, 1975). This attitude could be a mistake. It could be equally interesting for them to examine the much larger body of negative psi results, looking for consistent factors. Such an analysis could lead to new theoretical discoveries regarding the limitations of psi, often finding out what a phenomena cannot do leads to a fruitful discovery.

---

48 If the past data does not permit this, due to its lack of specific detail, then the professional bodies involved in parapsychology should introduce minimum standards of reporting within their membership. These enforced standards could include minimum security measures (see chapter three) and should at least include each study's (standardized) full details being recorded. These results would ideally be made available from a central administration point. This would ensure that at least future workers can conduct such analyses. The raw data, detailed write ups, session logs, and the associated suites of programs forming this dissertation will be freely available (from Sept 1989) to any interested party via JANET automated electronic file transfer protocol (FTP) from the author at Portsmouth Polytechnic's computer science department (Janet address: MorganK@Uk.Ac.Port.CsVax).
Returning to the computer attitudes work which we have selectively reviewed, the results do tell us that no one universal panacea exists for the reduction of negative computer attitudes. Instead the research has produced a whole battery of helpful measures which have been shown to work in selective cases. Those treatments which are most suited to the subject group will have to be found by the concerned party.

**Liaison between HCI researchers and the computer designers.**

We mentioned earlier that many of the problems which accompanied the introduction of early computing systems were based upon a lack of communication between the human factors experts and the computer system designers. Fortunately this situation is changing, and with it we should see an increase in the number of successful computing systems. This increased communication is taking place through the introduction of a new members in to the computer systems’s design team. These people have backgrounds from the social sciences, and provide urgently needed expertise in such areas as the interface design, and the systems introduction into the work place, both of which have been the cause of major problems in the past. Graduates from any subject with experience of human factors are now highly desirable in the computing career marketplace. It can only be hoped that the academic establishments concerned have trained these individuals thoroughly enough to warrant this interest from the computing industry. It could be extremely damaging to the future of HCI research if the designs produced by these new graduates were not commercially successful. The future acceptance of HCI depends on the results that its practitioners provide today in the market place, not the promises they make for tomorrow. Already criticisms have been made by the Governmental sponsors of HCI research of yesterday. In their 1988 statement at the end of the Alvey programme the head of the section of the Department of Trade and Industry responsible warned the HCI community against the danger he perceived of them gaining a reputation of only being able to give post-hoc advice. The computing world is highly applied, and it judges purely on current performance. It is on this basis that HCI’s improved interface designs will succeed or fail.

### 4.4. KMDB Questionnaires.

As well as the behavioral data gathered from the interface system the study hoped to take a measure of the S’s psychological, social and environmental differences, and how those
differences influenced the S’s attitude and behaviour with regard to new technology, and preferred interface types. A series of questionnaires were created for the study, each with a specific purpose. One was a Technology Attitudes Questionnaire (TAQ), which was jointly created with Robert Morris, and was based upon a series of similar questionnaires he had devised with the help of colleagues. The TAQ measures various aspects of the S’s experience and background with new technology. The other questionnaires are post session evaluation attitude measures, called the User Evaluation of Interactive Computer Systems questionnaire (UEICS). Versions of this were developed for both interface systems. The factors involved with these questionnaires are discussed in greater depth in subsequent chapters.

4.4.1. Technology Attitudes Questionnaire.

The technology attitude questionnaire was designed to take a measure of the user’s previous exposure and attitudes towards new technology, and in particular to computers. This also had sections which isolated self-reported machine interaction success. This hoped to isolate the MLP/FLP scale, for comparison with the S’s psi success rate in the database task. The TAQ contained 55 questions which investigated the following areas.


4.4.2. Post session questionnaires.

The post session questionnaires were designed to take a measure of the users attitudes towards the interface they had just used. The Wimp and Command line versions consisted of 27 and 25 questions respectively. Both versions had a similar format, which includes:

Copies of these questionnaires are provided in the appendix.

For both parapsychological and computer attitudes work in Syracuse.

Quote:

We want principles, not only developed, - the work of the closet, -but applied; which is the work of life.

Horace Mann, Thoughts, 1867

5.1. Introduction.

For the purpose of this dissertation we will attempt to conduct a selected overview of the various attempts, and advances that have been made by academic parapsychology to increase the effectiveness or use of the phenomena they study. Although the researchers of the late nineteenth century, and early twentieth century could been said to have been applying the phenomena they studied (either as attempts to contact the dead, or to provide solace to bereaved people, depending upon your view point), this review will concentrate more upon the experimental work that has been conducted since modern parapsychology came into existence. For purposes of convenience this date will be assumed to the 1930s, since that marked the beginning of Louisa and Joseph Rhine’s work at Duke University, Durham (1934). Of primary concern to us in this review will be studies which have contributed something to the methodology that was eventually adopted in some way, or had some influence upon, the study that this thesis describes. No attempt will be made to define the terms ESP, PK, Psi, or accepted scientific methodology. The problems associated with these definitions have been briefly outlined in the first chapter, and will also be dealt with in greater length in subsequent chapters.

1 Although there had been some earlier experiments it can be argued that they did not form a systematic effort.

2 The fundamental concepts and background to a subject are assumed to be known by the reader of a PhD dissertation. Readers who unfamiliar with the field are referred to an introductory work such as 'Foundations of Parapsychology' (Edge, Morris, Palmer, & Rush, 1986).
The problem involved in trying to use psi in any useful way is that it is much too unreliable and weak. It is also notoriously difficult for experimenters or subjects to differentiate hitting from chance guesses (Rhine, 1958; West, 1962 p142; Delanoy, 1986; Milton, 1986), although there is some evidence that this is not always the case (Kanthamani & Kelly, 1974). Indeed some theorists believe that one of the consequences of the way psi works makes it an unreliable source of information (Von Lucadou, 1987). Most attempts at applying it have tried to increase the signal to noise ratio using techniques from a branch of the computational sciences known as information theory.

5.2. Signal Enhancement Techniques.

In traditional engineering terms signal enhancement techniques include several options. The first of these is changing the transmission rate or strength of the transmitted signal. This is to try to avoid possible amplitude or frequency clashes between interfering signal(s) and the target signal (see section in chapter three on interference). The next option is to increase the receiver's sensitivity. In mechanical terms this means increasing the accuracy of the physical receiving device. Often this involves modifying the characteristics of the physical receiver, or reducing the noise within the receiving system which is causing the system to corrupt the target signal once it has been received. The alternatives to these physical techniques involve altering the target signal’s content. These measures include duplicating the transmitted information, and including confirmatory check data within the signal. This check data allows the receiver to check that the received data is correct. Duplication can either involve multiple transmission of the same data, or massive amounts of duplicated redundancy within the transmitted signal. Internal message checks range from simple check sums (a parity check being among the simplest examples), to highly complex checks which involve encoded versions of large parts of the data. When dealing with known signal transmission methods these techniques can combat very high degrees of signal interference. The problems with their use in parapsychology are that they assume that something is

---

3 The IDS models (May et al, 1986; May, 1986) propose that even in some forms of PK the S uses GESP to determine the best moment to initiate a trial. It is as yet uncertain if IDS proponents will try to expand their model to account for Macro-PK by these means. If they did it would be interesting to see how they proposed to explain such Macro-phenomena as levitation (Batcheldor, 1979, & 1982), or piezo-crystal deformation (Isaacs, 1984).
transmitted from a source to a receiver. Many researchers (such as the observational theorists) are not sure this assumption is justified.

5.3. Signal Enhancement Techniques in Parapsychology.

5.3.1. The Multiple Guessing Technique.

One of the earliest examples of an attempt to enhance the signal is a war time study by Foster (1943) which used 50 Plains Indian children at a primary school. Foster was interested in the use of a then novel method for enhancing ESP, by using multiple guesses made by the children to an unknown 'yes' 'no' target answer. This multiple guessing technique was based upon the assumption that ESP would respond to standard signal enhancement techniques.

The use of multiple guesses, or as some authors have called it, the majority vote technique, assumes that when a subject repeatedly guesses the nature of a target, each guess is independent, and the accuracy of those guesses can be enhanced by the use of simple mathematics. Thus if there were two possible targets, say red and blue, by getting a subject to guess the same target on several occasions, the total cumulative number of guesses made for each target should be able to give a more accurate indication of the target. Obviously such a simple example as having two possible targets, red or blue would make the subject's natural biases toward the target more problematic. Fortunately this can be overcome by either making the subject blind to the number or nature of the targets, or by ensuring that the target order controls for any bias. In this first use of the multiple guessing technique Foster got the children's non-indian (Caucasian) teacher to conduct two types of ESP tests which compared the old technique (single call) with the new multiple guess technique. Even though this was a novel approach it had an unfortunate start since Foster only found chance results with multiple guessing, and significance with the old method.

Assuming that the S is blind to the previous guesses made on that same target.
Due to the intervention of the Second World War there was little work carried out in applied parapsychology. Foster's ideas were not followed up until Fisk and West's (1957) attempt to use the majority vote guessing technique. Although Fisk and West got an impressive number of repeated guesses for some targets, they were unfortunately not the correct targets. The repeated guessing technique was also investigated by Thouless (1960), but with little success. However Thouless did introduce the 'index of preference' which he tried to use to compensate for the inherent preferences of subjects for certain targets. Although Thouless's 'index of preference' was a helpful advance it also created some statistical problems, which were corrected by Scott (1960).

5.3.2. Internal Message Checks.

As well as duplicating transmissions, the reader will recall that an alternative method of signal enhancement is to use some form of redundant accuracy checks within the transmission. It was not until the early 1950's that the first attempts to use a form of accuracy check were made within an ESP task. This task was finding lost objects, by the use of dowsing. Although scientific evaluation of dowsing had also been made by Dalton (1952), an investigation of more importance to this study is the work of Cadoret (1955). Cadoret tried to improve the reliability of dowsing success by getting some measure of the accuracy of the predictions, before they were used on a more major task. This was done by using a sample of trials to predict performance in the main series of trials. This use of a sample to predict performance was similar to the idea proposed by Stuart (1941), except that Stuart was interested in ESP card calls and not in any form of applied use. In his work Cadoret used a local test target and real remote target in the same study. The test targets were hidden coins. The Ss had to predict which row and column of 25 (5x5) tiles concealed the coin. On the basis of Ss success with this first guess predictions were made for the real target which was a hidden object on a plot of land in the author's residence. This crude method appeared to enable S's guesses to be more accurately applied to finding the target. Good sample

5 Although Rhine (1945) commented upon the applied use of parapsychology, and its problems, his comments were directed more towards the personal use of psi, and not towards the mechanical enhancement of the phenomena.

6 Cadoret's findings were later replicated by Osis (1960). However Osis did find the technique less effective for targets that were at a very large distance from the S (Osis, 1960; West, 1962 p179).
guesses on the test target tended to be followed by good guesses on the real target as well. These results seem to be in accordance with information theory’s predictions since 'bad' guesses were eliminated before they got applied to the real target.

The work of Clark is also of methodological importance (1958), since he proposed the use of ESP cards to predict the temperature. However of more interest is the work of Taetzsch (1958), who proposed using ESP targets with two sets of characteristics. These dual characteristics allowed one set to be used as a testing index, and the other as the actual target. However work by Schmeidler (1960) which used such dual aspect targets did not show a consequential increase in significance. Taetzsch (1962) further improved his proposed method by the suggestion of using a machine or computer to record the repeated guesses. This idea has formed the basis for many of the developments that have taken place in the applied use of parapsychological phenomena.

The next developments of major importance were the series of casino oriented trials conducted by Brier and Tyminski (1970). In the first of a pair of papers the authors described how they used the methods we have already described to try to enhance the ESP success at predicting the outcome of various gambling targets. The authors used repeated guessing at the same target, combined with a predictive sampling technique. This worked so that Ss produced ten 25 trial predictions which were to be matched for 50 spins of a roulette wheel. Five of the 25 predictions were then combined to be a (1 x 25 trial) run using a majority vote system. If the first predictive sample run provided an above chance number of hits then the second series of 25 trials was reduced using the same method. The authors conducted four pilots, of which the first and second gave positive results ($p < 0.01$) but during the third and fourth series the predictive indexes were at chance so the subsequent trials were abandoned. This was as well since the 2nd half of the later series were at chance. The authors also reported having successfully used the same method to predict the outcome of casino thrown dice (20 trials, $p < 0.05$); and cards at baccarat ($p < 0.05$). The authors reported that they felt one of the most important features for applying psi was the need to determine whether the psi based deviation was going to be positive or negative, to make the percentage of hitting as high as possible.
In their second paper, based upon the data from the same study, (Brier & Tyminski, 1970b) asked two questions which had been raised by the research outlined in their previous paper. The first of these was how successful was the majority vote technique. The second was did clear cut majorities, those in which the vote was based on five correct guesses out of five guesses, produce more hitting than slim majorities, where there were three or four correct guesses out of 5? In investigating the first question they found that the majority vote technique had raised the percentage of correct trials in six out of every seven runs. While surprisingly they found that the answer to the second question was that the most successful majority was when there were three out of five correct, and this occurred for both predictive and actual uses of the data.

The use of psi to help predict the outcome in gambling was also adopted, but with less success by Dean and Taetzsch (1970). The fact that one group of researchers finds success, while another does not, while both are using the same, or similar techniques, is a persistent paradox within parapsychology.

One of the most important applied studies was that conducted by Carpenter (1982). He used a repeated guessing technique, and a mood predictor to try to enhance ESP to the point at which a five letter word 'PEACE' could be sent via psychic means. He used the repeated guessing method of improving the efficiency of psi, but noted that this technique was not useful unless some means of predicting psi hitting and missing was available. Another problem faced by Carpenter was that index sampling of performance as a run proceeds is only of help if a prediction of the size of the overall deviation was available. Otherwise the experimenter cannot tell if the signal is hitting or missing. It would be undesirable to enhance a signal which was in the wrong direction, since it would give the opposite result to that required. To try to overcome these problems Carpenter used a mood scale to predict the psi direction and deviation. He found that using this mood indication to 'swap directions' when appropriate enabled him to predict the direction to within 76 percent. A second series of studies applied this technique and the word PEACE was embedded (with sample targets) via morse code into a 24 trial target list. This was successfully transmitted via 12 correct

---

7 Leading researchers have remarked that inventors of techniques tend to get the strongest results with those techniques (Personal communication Morris, 1989).
calls, using a repeated guessing technique. This was an enormous improvement in efficiency over that achieved by Ryzl in the early sixties. In a later paper Carpenter (1983) showed that the development of a so-called V scale enabled him to predict the size of the departure of the total score from chance. This V scale recorded the subject’s mood state and was developed as a predictor of the runscore variance. Carpenter showed that with this enhancement the single lists of majority vote calls from his previous series could be made significantly more accurate than the overall body of calls from which they were derived.

Over twenty years after the initial suggestion by Taetzsch (1962), Hal Puthoff (1984) reported the details of a system he had developed to enhance ESP guesses. He designed and wrote a small program for a Hewlett Packard (HP-41c) pocket calculator, which he claimed allowed him to enhance the predictions for the outcome of the spin of a miniature roulette wheel. Puthoff had used coding theory to allow him to enter the red or black calls as ones and zeros to the programmable calculator. These calls of one or zero were then recorded using the calculator’s pseudo random number generator (PRNG) into specific register locations that were associated with the two states. These inputs were then recorded using a pseudo RNG to specific memory locations that were associated with the two target states. This process, called by Puthoff ‘the subtrial process’ continued until three entries had been allocated to one of the two registers, thereby using a so-called five-bit majority vote technique to try to amplify the signal. To generate the target for this process Puthoff spun a miniature roulette wheel, in his desk drawer. Since Puthoff was using a single subject experimenter protocol he closed the desk drawer so he could not determine the roulette wheel outcome until he was ready; then all he had to do was pull open the drawer. Puthoff reported extremely high levels of success using this technique ($p < 0.01$ for the actual trials, and $p < 0.05$ for the subtrials).

In a follow-up study to his previous work with a calculator Puthoff (1984) enhanced the design of his system by making the subject’s button presses interrupt a fast internal logic flip-flop of 100 counts per second. A discussion on the advantages and disadvantages of using pseudo random number generators is given in chapter 3 of this dissertation. Suffice to say that the use of manufacturer’s built in PRNG cannot be recommended.

The 100 counts a second was used since Puthoff felt it exceeded normal biological motor responses.
the states of one or zero to generate the multiple guessing. Puthoff used two subjects (E.M. and M.T.), and this new improved process to guess at the same roulette wheel concealed inside his desk drawer. One subject got chance results, but the other obtained significant hitting (p < 0.01).

5.4. Objections to Signal Enhancement Techniques.

Kennedy (1979) has proposed that if psi was independent of task complexity, as he claimed he had shown in a previous review (Kennedy, 1978), then methods employed to enhance the signal would be irrelevant to the overall psi effect size. This would be in direct conflict with the results found by those studies that had successfully used the 'majority-vote' technique. Kennedy reported having reanalysed the only detailed majority-vote data available (Brier and Tyminski, 1970), and found these results to be consistent with his proposed goal oriented, or complexity independent psi. Kennedy reported that when he removed the effect of the majority vote technique on the raw data the effect size remained constant. He proposed that this was consistent with the goal oriented models of psi.

5.5. The Use Of These Signal Enhancement Techniques Within This Dissertation’s Study.

The SSM (skeptical software monitor) system was an attempt to use logic to detect and enhance computer-based smart noise in a problem solving situation. In Morris’s original proposal of smart noise (Morris, 1985 p17), it was suggested that the computer system could check if the FLP component was actually contributing to the bit stream by checking for the PK-Signature\(^\text{10}\) of the FLP before proceeding. Unfortunately the author’s investigation and evaluation of around 20 different RNGs failed to reveal one which was pure enough for

---

10 Such a “PK-signature” would be likely to vary from period to period, and any system would have to account for this variation. Given a reliable enough signal such fluctuations might eventually also be useful in the diagnosis of certain diseases, which might produce a certain common effect on PK-signature patterns. It would also be interesting to see if stress changed the pattern in a consistent and recognizable manner. Such a stress detector when built into the system could dynamically change the user interface, to find a presentation which it detected caused less stress in the user.
these kinds of pattern searching. This meant that the author had to develop an alternative method of detecting any psi hitting, and reducing the effects of psi missing. His solution was called the Skeptical Software Monitor (SSM). However before we can discuss the SSM we will have to review the background to the task selected for the smart noise application.

5.5.1. The Nature of the Task.

Before the processes involved in trying to enhance any FAA can be discussed we will have to explain the nature of the task being attempted.

From the very beginning of the three years devoted to this study one of the primary aims has been to learn as much as possible about the possibility of using Function Alteration Activity (FAA) to perform useful and intelligent tasks. The definition of what defines an intelligent action is highly complex, and would have involved a PhD in its own right. Fortunately a large body of work in artificial intelligence (AI) and computer science (CS) has already been devoted to this problem, and some basic definitions of the prerequisites of intelligence within a task can be presented. The primary one of these is called the Heuristic Search Hypothesis, which was defined in Newell & Simon’s 10th Turing Lecture of 1976 (Newell & Simon, 1976)

A physical-symbol system exercises its intelligence in problem-solving by search—that is by generating and progressively modifying symbol structures until it produces a solution structure (Newell & Simon, 1976).

Such a Heuristic Search is the formal method by which a system extracts information from its problem space\(^\text{11}\). The efficiency with which a system can conduct this search is a direct measure of its intelligence. Since the study described in this dissertation was interested in testing the possible intelligence in a subject’s FAA, it was important that the task be one in

\(^{11}\) Whether neural networks extract information from problem space using a heuristic search is a controversial point within computational science. Fortunately it is beyond the scope of this dissertation to define if such connective computational engines uses heuristic or not. It is important to highlight that it may be possible for devices to be capable of problem solving without a formal heuristic. Although many computer scientists would regard such a suggestion as heresy, it in no way invalidates the use of the heuristic search to detect intelligent behaviour.
which the psi directly reduced a specific problem space. The HCI interface study has separate requirements, but it was found that both sets of experimental requirements could be matched within an information retrieval system. This was eventually implemented as a college database system. The problem space was to find and retrieve the requests for information made by subjects, in the form of database records.

A highly complex part of applied computer science is the design and construction of database systems. Within such database systems there are currently several different record storage and retrieval methods. An evaluation of each of these methods had to be conducted to determine which was most suitable as a potential smart noise task. The author reviewed the possible use of a hash function, binary tree, linear search, step search, and binary search methods. The binary tree (Bayer, 1974, Knuth, 1973) was considered to have some practical advantages over the other possible methods. The binary tree is conceptually easy to understand, and allows the simplest statistical evaluation technique of all the other possible methods.

5.5.2. Database Record Searching Method.

The structure of the b-tree used in this study is best described by the use of an example. We will assume the simplified b-tree shown in this diagram. We can see that the values 3,5,6,7 and 8 are held within the tree structure. The study used a b-tree of order two, known as a binary tree. This is where there are only two possible branches from any one node. The

---

12 See chapter four for a discussion of the HCI requirements within this study.
13 It is beyond the scope of this dissertation to describe database design. Interested readers are referred to texts on database techniques such as Bradley (1982).
14 Readers who are interested in the evaluations conducted on the various alternative searching methods should refer to the appendix of the author's end of year study report May 1987. This is held by the psychology department.
ordering within the tree is such that values to the lower left of a box (or node) are less, and those to the right are greater. There are no fixed rules for this ordering direction, but the left to right sequence was adopted for the KMDB system. During normal logical searches of the tree the process enters the node at the top of the tree (called the root node). It then checks if the value it is searching for is the same as that held in the current node. If the value it is searching for is less than the current node’s value then it uses a pointer in the left hand side of the node to progress to the next lower left hand node. If the value it is searching for is larger than the current node’s value then it uses a pointer in the right hand side of the node to progress to the next lower hand node. This lower node now becomes the current node, and the process is repeated. This continues until either the value is found, or the process comes across a null or empty pointer. In this latter case it can assume that the value it is searching for does not exist. The example given shows what is called an unbalanced b-tree. This is so called because the three left hand sub-nodes with the values 3, 6, and 5, are not stored in the most search efficient form. To access the value five the process would have to descend to three levels of recursion15. This would be wasteful of computing resources. It is much more efficient to use what is termed a balanced b-tree. A balanced version of the example is shown. In this one can see that the worst case retrieval now only takes two levels of recursion. To maintain the maximum efficiency a b-tree should be rebalanced after each deletion, or insertion of a record. From these descriptions it should be obvious that the

---

15 In computing, recursion is where an algorithm’s solution is defined in terms of itself. It is solved by repeatedly calling itself until it reaches a terminating condition. The algorithm then ascends the levels of recursion it previously descended providing each level with a solution.
b-tree is a recursive structure, which lends itself to recursive programming. In the b-trees used in the project the nodes contained information associated with the node number. This number was known as the key value and was unique.

5.5.3. The Psi Task.

Having described the processes involved in normal logical searching we can now compare that with the process involved in the smart noise task. To do this we will have to discuss the role of the Skeptical Software Monitor (SSM).

The workings of the SSM.

The skeptical software monitor, or SSM is the basis of the entire parapsychological experimental work conducted for this thesis. The idea of using a real application as a psi task may well attract some attention, but it is the SSM which really makes the project unique. The SSM uses several sources for the decision making, spread around the entire functionality of the system, from the clock to the actual hardware design of the Amstrad target machine. The idea behind the SSM system was to have an experiment which evaluated smart noise, and AHCI in general. It also had many other agendas, such as the exclusion of as much artifact as possible, and the attempt to gather as much data as possible. Even if some data was not used by the author, it would be available for future evaluation. For this reason the system included many features which were never evaluated. An example of these unused features is the ability to record the user's mouse movements in the WIMP system, and replay them.

5.5.4. SSM

This section will not assume any knowledge of the KMDB SSM system, but due to space constraints has to assume some knowledge of computer science. We will conceptually
summarize the events in a typical SSM trial. The full technical details of all the processes involved in a trial are documented in the KMDB technical reference manual. They are too lengthy and technically complex to be reproduced here. The first point to make before we start this description is that trials can only occur during 'eq' database searches on a record's key field. This means the search has to be for a unique record. Searching for multiple records makes the statistical evaluation too complex. For this example we will work through the processes involved if the subject issues the command 'select dept key eq 100'.

Once the user has entered this command, either on the command line or Gem system it will pass through the parsers, and eventually on into the highest levels of the SSM. Since this is the point we are most interested in we will ignore the rest of the process. The highest level of the SSM determines which of the several possible record types this search is for and calls the relevant procedure (in this case procedure dept_psi_search). It passes control over to that module with parameters of the relevant b-tree, the requested target department key number, and a boolean flag 'found' (which is used to determine if the search has been successful).

Smart Noise Programming.

The logic involved in using psi to solve problems is rather strange.

In its purest form smart noise programming has no GUARANTEED TERMINATION. It also has no logical method behind its inference decisions. For the author the most noticeable effect of these differences was a constant uncomfortable 'subconscious' feeling. This was similar to the feeling experienced programmers have when they have made an unnoticed logical error. It took the author some time to become accustomed to this
sensation. Given that there is no assumed terminating condition, the KMDB system has to ensure that it will eventually return to the user with something useful. It would be quite possible for the SSM to spend a very long time in the searching process, if it made a lot of poor choices trying to find where the record is held in the b-tree. The early versions of the software did indeed do just this, and had to have a time out flag so it could return to the user. Unfortunately under these circumstances not only did the software fail to retrieve the record, it did not even know if the record existed. Complete failure of a 'psi' search is not only infuriating for the user, it is also useless for the investigator, since we do not learn anything useful about the processes involved. Since these negative results were unacceptable in a real world situation the search had to get feedback and to use logic to act in response to that feedback. The SSM plays a major part in this active response to feedback. It limits the decision making sources which are doing very badly, and allows those sources which are doing well to have more control over the decision making process.

The original design of the SSM was such that when the process entered the tree, it checked if the current node was the target node. If it was not then the process called the decision making sources to determine whether to go down the left or right sub-trees. The process then moved down the selected sub-tree and checked if the decision it had just made was in the correct direction. If it was incorrect then it backtracked to the previous position before starting the next trial. The result of each trial was fed back to the SSM to adjust the noise source weightings for the subsequent trials. This meant that all the correct sources had their weightings increased, and all the incorrect sources had their weightings decreased. The net result of these weighting adjustments was that the successful sources had more influence than the incorrect ones. However as we have already mentioned the early models of the SSM could remain searching in the tree forever, and still be no nearer a result. This would have been impractical in a covert study such as the one at Napier. So, to avoid this, after each incorrect trial the process was backtracked up the tree, and then 'shoved' down the first level of the alternative sub-tree it had previously just rejected. The incorrect trial was, of

---

17 When there were a large number of records the SSM could take over 15 seconds to return with a record. Since there is experimental HCI evidence (Shneiderman 1987) that delays greater than 2 seconds reduce usability, and the study was covert, it was felt that delays greater than seven seconds should be avoided.
course, still recorded. Once the process had been shoved down the first level of the correct sub-tree it was again left to make the next decision. This was the only acceptable method which allowed the system to have a finite number of attempts to guess at the records location.

A brief example, using the diagram of the balanced b-tree we used previously, will suffice to demonstrate that this modification does not fundamentally alter the task. For this example we will assume that we have entered the tree with the task of retrieving the record identified by the key 6. The process enters the root node and randomly decides to go left (which is correct). This takes it to the node containing the value 5. The process checks if it has made the correct decision, records that outcome as the result of the trial, and gives that feedback to the decision making software. The process then randomly decides to go to the left (which is incorrect). This results in the process descending into the left most node with the value 3. The process checks and finds that it has gone in the wrong direction. It records this as the result of that trial, and gives this feedback to the decision making software. The logic then backtracks the process up to the previous node (containing the value 5), and moves it down into the right hand sub-tree (containing the value 6). The process checks, and finding that to be the target node returns to the user with that record’s details. The whole process took two trials, one of which was a hit the second of which was a miss. For details on the various systems used within the SSM see appendix 13.

5.6. Increasing the Receiver’s Sensitivity.

As already discussed the alternative to reducing the noise strength is to increase the sensitivity of the receiver. Within the communication paradigm of Psi that involves altering the some aspect of the S. In a famous study from behind the 'Iron Curtain' a Czechoslovakian parapsychologist successfully used hypnosis to increase the sensitivity of individuals. This method was apparently so successfully that Ss reached the point at which they could use these new abilities at will (Ryzl, 1962 & 1963). Ryzl reported that from a total subject pool of 226 persons (73 male, and 153 female), aged from 16-33 years, 3 men and 24 women
displayed relatively good clairvoyant ability, and the ability of 13 women became ‘very good’. Ryzl’s method was illustrated by a detailed case description of a Ms J.K., who over a period of some months of intensive deep hypnosis developed extraordinary abilities. Ryzl also attempted the use of repeated guessing techniques with Ms J.K., but due to pressure from Ms J.K.’s family this research had to be interrupted (Ryzl, 1966). Unfortunately some attempts to replicate this most promising work have failed to achieve the dramatic success reported by Ryzl (Beloff, 1966).

Throughout the early 1960’s Ryzl continued to be a major influence in the development of applied parapsychology, and with a male subject called P.S., who seemed to be reliably producing ESP he tried a series of experiments in which he used repeated guessing, and predictive index trials to successfully send five three digit numbers (Ryzl, 1964). However this took some 19,350 ESP calls, and took over 50 hours, so it could not be considered an efficient method of information transfer, it did however demonstrate that such information transfer was possible.

Jacobs (JJ) and Breederveld (HB) (1979) published the results from a study which had been an investigation into applying the observational theories to influencing the outcome of the national German Lotto game. They used HB as the S, with an objective of trying to force the lotto machine to chose certain numbers. This study had two sets of conditions, the first where HB was hypnotised in his home (either by his wife or by JJ). Once hypnotised he was instructed to observe the numbers chosen by the Lotto game as being the same as the target numbers. The second condition was similar, except the hypnosis was omitted. This study reported finding that with a Mean Chance Expectation (MCE) of 3.7 hits, they got five hits using hypnosis, as opposed to only three when they did not (P < 0.01 corrected for optional stopping). However, there were some unavoidable minor methodological weaknesses in this study, such that the sequence of the experiments (under hypnosis or not under hypnosis) could not be planned beforehand (Breederveld & Jacobs, 1979 p4). Such a protocol of course has a weakness in that the hypnotiser chose when to select the conditions for the Ss observation after the lotto numbers had been published.

---

18 Apparently including precognitive clairvoyance of the outcome of a horse race, and locating lost objects and persons.
5.6.1. The Use Of Receiver Enhancement Techniques Within This Dissertation's Study.

The purpose of any PSI enhancement strategies within this dissertation were to enhance the functioning of the Smart Noise System. As the first three experiments were covert no explicit instructions could be given to the subjects regarding their mental attitudes during the experiment. However one of Morris's (1986) hypotheses for function linkage was explicit in proposing that subjects who enjoyed working with computer systems would be expected to produce more function linkage. This would mean that they were expected to PSI hit. Malfunction linked subjects would be expected to dislike the system with which they were working and degrade the performance of the computer system. This would involve the production of PSI missing. To this extent the questionnaires completed by the subjects and the recorded SNS results could give an indication of the validity of these hypothesis.

In the final experiment (see chapter 10) subjects were asked to try taking a wholeistic view of their interactions with the computer system. It was suggested that the subjects imagine themselves to be part of a whole system which included the computer. They were told to imagine that the single objective of this 'team' was to retrieve the records within the database system. Subjects were also told not to try too hard, and to avoid focused mental attention on the physical actions they were performing. They were told to recall highly successful and satisfying moments of their past computer use, and to have confidence that they could produce function linkage with machines. If they were skeptical of the existence of Psi, and many of them were, then they were told to hold an open mind and positive attitude to the existence of function linkage during the experiment. Finally, subjects were asked to try to forget that they were in an experiment and to concentrate instead on 'playing' with a new and different piece of software. It was hoped that these strategies would enhance any positive FAA which the subject might possess.

5.7. Applied Uses.

Rubin and Honorton (1971) investigated the properties of the ancient Oriental oracle the I'Ching. They got groups of Ss to throw three pennies from a plastic cup to generate a hexagram while thinking of their forthcoming week, and of personal question. As a control
Rubin and Honorton used the RAND tables to generate another control hexagram. They presented the commentaries from both hexagrams to the Ss and asked them to rate how accurate, or appropriate the commentary was to the question they had asked. Although a blind judging procedure found no significant difference between the two ratings, a post-hoc analysis revealed that believers in psi rated the passage generated by the coin throws as being more accurate (p < 0.05).

5.7.1. Remote Influence Over A Subject's State.

William Braud and Jan Jackson (1983) reported upon the results of an experiment which indicated that the mental activity of an individual could influence the vividness of a subject's reported mental imagery. In the experiment Braud used 24 Ss (12 of whom were control and 12 of whom were actual trial subjects), and administered the Wilson-Barber Creative Imagination Scale (CIS). The results showed that influenced Ss had a significantly higher CIS score, which Braud and Jackson interpreted as indicating that an experimenter could exert a conscious psychokinetic influence over the mental processes of a distant target subject. In the same year William Braud also conducted an experiment with Marilyn Schlitz, (1983) which investigated a similar remote psychokinetic influence, but this time over a subjects arousal (electrodermal activity) state. Thirty two Ss were used, 16 of whom were placed in a state of high sympathetic nervous system (SNS) activity, and 16 of whom were placed in a state of low SNS activity. These states were induced by special induction procedures, which included REG controlled stimulating noise, and suggestion. Braud and Schlitz found that the highly active Ss showed a significant calming effect, but that the other already calm group were not affected. A second experiment investigating the size of the effect S's could induce upon themselves by biofeedback showed a similar magnitude. However it has been noted that the effects described in both these studies could have been due to GESP of the experimenter's wishes by the subjects (Personal communication, Morris, 1989).

5.7.2. Increased Temperature.

Evidence that psychic intervention could increase the temperature of a target was confirmed in a paper by Gertrude Schmeidler (1984), which defended her claim to this effect, which had been criticized by Brian Millar (1976 & 1978). Schmeidler concluded that her earlier
experiment (Schmeidler, 1973) was adequately controlled, and reported the results of a posthoc non-parametric analysis of her data to support her case. These findings, if true, are of some interest to this study, since heating of computer components is one of the major causes of malfunction and breakdowns.

5.7.3. Machine Anomalies.

Of more direct interest to the study outlined in this dissertation was the humorous article that appeared in 'Psychic' magazine in May of 1975 (Greenhouse, 1975). Entitled 'Pk-Missing People, or Mind undermining matter', it outlines a negative form of applied parapsychology, which would later form the basis of part of the study being described in this dissertation. Greenhouse described an informal 'experiment' in which he compared the breakdown and successful repair rates of household appliances. He reported the rate achieved by him, and by 'the average householder', and reported finding that he had statistically significantly worse repair rates than the average19.

At a symposium for the application of anomalous phenomena Robert Morris outlined the basis of a research programme into human equipment anomalies (Morris, 1983). Morris reviewed the literatures of accident proneness, and general human factors, concluding there was much that parapsychology could help computer science and psychology clarify in this problem area. Since the background and features of Morris's proposals have been discussed in the introduction to this dissertation, and there will be a further discussion in later chapters, the details will be omitted for this review. In a follow up paper to this work Robert Morris reported on the work that he planned to conduct in the Communication Studies lab at Syracuse University, New York (Morris, 1985). He outlined that the work was in its initial phase, but that some preliminary data had been found as a byproduct of a computer system failure during one of his laboratory's recent psi experiments. A hard disc was found to have crashed more frequently when it was used to test subjects who had negative attitude towards the existence of psi (p < 0.05) and also those who had rated themselves as being more prone

19 It is important to maintain caution towards these kinds of non-experimental reports since they could be artificially inflated.
to performance anxiety on a sports performance attitude questionnaire (p < 0.05). This area of research will be described in much greater detail in other chapters.

Uri Geller and Guy Playfair (1986) wrote and account of Uri Geller's apparently paranormal interactions with machines and particularly computers. These applied phenomena with computers are dealt with in the author's paper 'AHCI towards an understanding of what constitutes an anomaly', which is reproduced in appendix 1.

5.7.4. Remote Perception Of Some Aspect Of The Environment.

In the US another area of applied parapsychology was conducted by Braud, when he replicated Johnson's experiment (Johnson, 1973) which demonstrated clairvoyant knowledge of exam answers was a possibility (Braud, 1975). In Braud's study 46 undergraduates were administered an examination in which some of the answers to the exam questions were concealed in an envelope that contained the examination questions. In both the pilot and confirmatory study Ss scored significantly better on questions for which hidden correct answers were supplied without their knowledge than for questions with no concealed answers. Braud also found a significant positive correlation between 'psi' hitting and the number of errors on the exam. Braud and Johnson felt that these findings supported Stanford's (1974) Psi-Mediated Instrumental Response (PMIR) model. However there are some problems with these kinds of covert examination studies. There is no reason to presume that the Ss were excluded from psychically accessing each others examination answers, or the master answer sheet held by the experimenter, or indeed the knowledge base held within the experimenter's mind. With this kind of possible psi-sensory knowledge contamination inside the examination room it is very hard to tie down the source of any possible psi effect, at least until some form of psi-shielding has been established.

One of the major developments within the field of applied parapsychology was the Remote Viewing (RV) work of Targ and Puthoff at Stanford Research Institute (SRI) (Puthoff &

---

20 These examination findings could equally be due to some kind of experimenter effect.
21 Now called the Stanford Research International (SRI).
Targ, 1976). Although various methodological criticisms and counters have been made about this work (Marks & Kammann, 1980, Morris, 1980, Marks, 1981), it formed the basis for one the most successful research methodologies in parapsychology; and was certainly one of the major advances in applied parapsychology. Targ and Puthoff’s work lead to the development of the Associated Remote Viewing (ARV) protocol (Puthoff & Targ, 1979; Targ, Puthoff & May, 1979; Spotteswoode, 1983), which has been widely used in some parapsychological applications. In the standard RV the S’s try to describe a target which has been randomly selected from pool of possible targets. The Ss produce a free response verbal transcript and produce a drawing of they had experienced in the session. The differences between the RV and the ARV allow this protocol to be applied to real world situations. To achieve this the targets in the target pool are assigned to particular real world events, and the one target selected by the Ss is taken to denote which of the associated real world events will take place. Often to make judging of targets easier the descriptors associated with each target in the target pool are recorded on a computer. These are called descriptor lists. It is beyond the scope of this review to describe the ARV procedure in any greater length, interested readers are referred to the bibliography.

By the 1980s the RV protocols had become much more sophisticated than the work of Targ and Puthoff at SRI in the mid seventies. One such example is the project run by Schlitz & Gruber (1980), which investigated transcontinental remote viewing between Rome (Italy) and Detroit (USA). The agent visited 10 of 40 sites in and around Rome, and the percipient, at the same time recorded her impressions. Analysis of the results by blind judges produced highly significant matches (p < 0.00001). Although this impressively high score may have reflected a minor artifact in the judging process. This artifact was later noted, and corrected (Schlitz & Gruber, 1981).

Marilyn Schlitz and JoMarie Haight (1984) collaborated on a replication of the previously successful RV experiments that Schlitz had been involved in (Schlitz & Gruber, 1980). In

---

22 Marks and Kammann (1980) reported finding that the unedited transcripts used for judging frequently included the date or other information indicating when the trial had taken place (Palmer, 1986 p177).
this 1984 study 10 trials were completed between Durham, North Carolina, and Coca Beach, Florida. Schlitz acted as percipient, and Haight as agent. The authors reported that independent judges using a count-of-permutations\textsuperscript{23} method found a highly significant effect. They concluded that applied parapsychology could have much to gain from the use of both empirical and phenomenological disposition toward the investigation of psi.

In a novel adaptation of the RV technique Vallee used a computer conferencing system to coordinate and control an RV experiment which used geological samples as targets (Vallee, 1985). Vallee reviewed the use of the RV process from the work of Targ and Puthoff and Jahn (Jahn, 1982), and described the potential of using a computer system, since it recorded the date, time and mentation automatically. This gave an additional degree of security, and also enhanced the comfort for the subjects. Geological targets were used since Vallee felt there was an Undeniable economic value of the accurate recognition of minerals by psychic means.

Vallee reported a significant result (p < 0.05), but sounded a note of caution to those who would use this technique in applied settings when he said Attempts to specify location were usually in error as were descriptions of the size of the samples and their exact substance.

Finally Russell Targ (1989) outlined the results from two series of ARV studies which had tried to use the ARV protocol to predict the outcome of real world events. The first series tried to predict the silver bullion stock market fluctuations, and was remarkably successful, with nine successes out of nine trials. However an attempted replication failed, with nine misses out of nine trials. Targ concluded that the Data is trying to tell us something, and it is not that there is no ESP.

\textbf{5.7.5. Crime Prevention.}

Another investigation which attempted to evaluate the use of psi in the real world was the study conducted by Reiser, Ludwig, Saxe, and Wagner (1979). They tried to evaluate the

\textsuperscript{23} Count of permutations is a method by which the number of distinct descriptions made by a subject when describing a target are compared against the number of control descriptors used for non-targets.
usefulness of psychic information for the solution of crimes investigated by the Los Angeles Police Department. They used 12 reputable psychics as Ss and presented each of them with evidence from two solved and two unsolved crimes selected by an investigator not involved with the study. However, little if any of the information elicited from the 12 psychics provided material useful to the investigation of the crimes. There was a low rate of interpsychic accuracy and congruence among the responses, although other studies have claimed to show cases where psychic intervention had helped to solve crime (Vilenskaya, 1984).

In the same year Duncan and Russell (1979) conducted a study with 20 moderately successful managers and normal group members to see if they displayed any difference in precognitive powers. The previous studies of Dean and Mihalasky in the early 1960s, had revealed a strong correlation between business success and GESP ability. Unfortunately Duncan and Russell reported to be unable to find any evidence for paranormal abilities from any of these Ss in simple card guessing experiments.

5.8. Non-Experimental Speculations.

Separate from the experimental work parapsychologists have produced some philosophical and methodological speculations about the use of applied parapsychology (for example Eisenberg, 1977). This section will give brief examples of these speculations. Readers should be aware that the amount of non-experimental speculations vastly exceed the amount of available space in a thesis.


Reichbart (1981) noted that there might be a huge potential role of self-destructive and malevolent psi faculties, and that these forces are completely ignored by western law24. Reichbart noted that psychic aggression is not detected by the victim, or society; and that a

---

24 Many countries or states still have old statutes which prohibit the practice of ostensibly paranormal phenomena.
psychic crime would be undetectable, since the legal concepts of guilt and innocence are inadequate to accommodate psychic causality.

These pessimistic views were compensated for when, in the same year William Braud published a book on the practical application of psi (Braud, 1982), which outlined his thoughts and hopes for the applied use of psi to enhance the quality of human life. Postulations about the potential use of psi to predict the weather, medical diagnosis, and to avert disasters was proposed by Kautz (1983) in a brief monograph. Unfortunately such projects are still not possible, since psi is such an inherently unreliable phenomena.

Gambling.

Helmut Schmidt (1983) outlined some ideas in a brief paper that postulated that there were some inherent random processes in the human mind, and that these could be used as a natural random number generator in games of chance. He proposed that using his mathematical model of psi (Schmidt, 1975), one could produce extra chance success in the traditional gambling paradigm.

Increasing the sensitivity of the receiver.

Another review, this time on the techniques used by 'professional psychics' throughout the US (Galanti, 1984), found that they used similar techniques to facilitate the use of psi. This involved entering some kind of altered state of consciousness, through relaxation, or mediation, and then 'clearing the mind' of all other material. Galanti noted that this was similar to the highly successful Ganzfeld Technique (Bertini, Lewis, & Witkin, 1969; Honorton & Harper, 1974; and Braud, Wood & Braud, 1975). Galanti also reported that the psychics reported having access to information while in the altered state that was not preset during normal consciousness. He noted the close correspondence between the characteristics of the psi experience and those of classic psychosis.

Kelly and Locke (1981) produced a review of the historical use of scrying. Scrying is the use of reflecting devices to induce visual hallucinations, projected by the subject on the reflecting surface. Kelly and Locke suggested that the technique might be of great relevance to applied ESP research.
5.9. Discussion Of The Findings From This Literature Review.

The gradual development of the elements discussed in this selected overview, were of vital importance to the study. The computer system which was developed for this study used the majority vote technique, and some other techniques which assumed, as have many of the studies outlined above, that psi (if it exists), conforms to the same methods used by every other known sensory modality. However, as we have pointed out in the review of Kennedy’s work (1978 & 1979), not all parapsychologists agree that it can be assumed that psi operates in this manner. Many of those parapsychologists that favor observational theories, feel that psi may well be totally goal oriented, to the point that any effect can only take place at the moment of ‘feedback’, and such effects are thus independent of the techniques that are involved in the actual process (Klip, 1967; Schmidt, 1975; Walker, 1973; Millar, 1986; and Von Lucadou, 1987). At the time of writing there is no answer to such a debate, and if an answer ever does emerge it will only be by further empirical research. However a lot of the work that has been outlined above (Ryzl, 1964; and Carpenter, 1982) shows some very strong and more importantly, useful affects, which have been achieved using signal enhancement techniques. It was for this reason that many of the ideas developed in the work outlined in this review were incorporated in the computer system developed for this study, along with some innovations that had not been tried before in parapsychology.

5.10. Possible Applied Uses Of Smart Noise.

Finally, since the author is one of the few people in the world who has implemented some smart noise systems, it was felt to be appropriate to devote some space to the problems and advantages he can see in its use. This section will include some applications the author sees as the potential uses of successful smart noise systems.

25 Among the prototype systems created by the author are remote viewing systems, and a football pool prediction system, neither of which were more successful than a simple guess.
The first and most important point is that smart noise cannot be recommended for applications where normal computational alternatives exist. The effect is so weak and unreliable\textsuperscript{26} that it would be only suitable for problem domains which have no data, or known method of solution. The computation effort involved in the SSM, or any responsible use of smart noise is intense because in providing feedback to the system the solution has to be determined through normal methods.

Bearing this consideration in mind precognitive (or logically unsolvable) tasks are the only ones which make any real sense as Smart Noise Applications. Fortunately many highly complex computational tasks are in some aspect precognitive. Task domains where the waste in both time and effort of a total failure are of low importance are ideal, however as Morris (personal communication, 1989) has pointed out this may reduce the users motivation.

The potential of using smart noise in applications where there was no current method of solution led to the author proposing a computer system with Both Random And Inductive Networked-processors\textsuperscript{27} (BRAIN).

This would be the next logical development of the SSM, the difference being that BRAIN would have the ability to function in situations where there was no mathematical model to allow the normal "SSM" feedback process to be adopted. To achieve this the system would run a normal computational engine and a SNS in the same computing resource environment. The two processes would be in effect a twin processor system. Each process would be allowed access to both shared and exclusive resources. In effect the two processes would collaborate in the final decision as to what action to perform, but would be able to benefit from the intuitive action of the SNS. Disagreements in deciding which of two recommended actions to perform would have to be made either by a smaller master processor, or the human operator. This higher level (be it human or machine) would only need to be aware of the higher level processing of these two processors. Possible applications for these kinds of systems would include robot vision, pre-fetching instructions/data for a traditional bus-

\textsuperscript{26} This presumes that such effects actually exist, which (in the author's opinion) remains unproven.

\textsuperscript{27} It should be obvious that the author tries to exercise his sense of humor in the creation of these acronyms.
bound processor or speech recognition. In these applications the smart noise processor would act as 'hunches' while the inductive processor could be dedicated to logical processes. This type of system could be simulated in software before the expense of hardware construction.

An example of an exploratory test for these concepts would be where two robots had to traverse an unknown maze. One robot would use a smart noise system, and the other could use normal logic. If the robots were made to resemble small furry animals then children could be used as subjects and be assigned a robot to support in a 'race'. If the children were blind to which robots were using which system then this could allow an evaluation of the various combinations of observer effects upon the performance of these systems. If smart noise was ever successfully applied it would vital to determine the effects of various types of human observation upon the system’s performance. Again this study could be simulated in a computer game before it was actually implemented in hardware.
6. Evaluation of the covert parapsychological smart noise system, and the interface comparison software.

If a man will begin with certainties, he shall end in doubts; but if he will be content to begin with doubts, he shall end in certainties.

Francis Bacon.

6.1. Introduction.

Previous chapters have already outlined the aims of both the parapsychological, security and HCI parts of this thesis. For this reason their content will only be briefly summarised.

6.1.1. HCI aspects.

These series of three studies were designed to investigate the interface performance and preference differences between Ss in three groups: naive keyboard illiterate, naive keyboard literate, and sophisticated. Of specific interest were; interface preferences, productivity, and error rates across these three subject populations. These questions would be answered by the use of extensive recording and monitoring of each subject's actions. As well as the behavioral data gathered by the system the study involved the development of a series of questionnaires, called the TAQ & UEICS respectively. These were designed to measure the S's psychological, social & environmental differences, and how those differences influenced the S's attitudes and behaviour towards new technology, and their preferred interface type.

---

1 An example of data collected from each subject's session is shown in appendix 2. An example of how detailed these recordings could be is given by the fact that the system can record and play back the user's mouse movements made within the application. This playback can be at twice, normal, or half actual speed. This feature exists within the KMDB WIMP interface, but was not used within these series of experiments, since recording the mouse movements would have required more disc space than was available.

The aims of this part of the dissertation were to try to develop a comprehensive guide to the security of the parapsychological use of computer systems. Details of this are given in chapter three.

6.1.3. Parapsychological aspects.

The parapsychological aspects of the study revolve around the evaluation of a Smart Noise System (SNS), implemented within the Skeptical Software Monitor (SSM). Both these concepts are discussed in depth in chapter five.

6.1.3.1. Control Trials.

Control trials are notoriously difficult in the traditional RNG-PK paradigm (see chapter two). The KMDB system was developed to allow what would be the keyboard input to be redirected from a file. This control input file was created to contain the same series of commands contained in the subjects tutorial sheet\(^2\).

Other than the existence of this one file, the control and experimental systems were identical. To reduce the potential timing problems all control runs were started by inserting the KMDB system disc into the machines drives, and powering up the machine so it booted\(^3\) up directly after its power up tests. This reduced the potential amount of experimenter determined machine time differences. In chapter three we discussed that all the experimental versions of the operating system were set to the same date and time\(^4\). A control run, such as this was conducted for each of the experiments, and will be indicated in the results as 'control'.

---

2 This included a repetition of some of the trials to simulate the displacement artifact in the control runs. See appendix 13 for full details of this artifact.

3 Popular contraction of the technical term 'boot-strapping'. This is the process by which the operating system is loaded from a storage media.

4 This date and time was 14:01:00 HRS on the 10th of March 1988 (10:03:88). This was the date and time of an experimental session in the pilot experiment. It was chosen from all the potential ones by the hardware random number generator in the experimenter's Casio\(^{TM}\) FX-450 scientific calculator.
6.2. Hypotheses.

6.2.1. HCI Hypotheses:

1) That WIMP interfaces would be preferred by naive users, but would have a slower throughput of work and a longer overall session time.

2) That once competence\(^5\) was reached in the command line interface, it would be preferred by experienced users and naive users who are 'keyboard literate', due to the interface's greater information interchange rate, and the S's familiarity with the keyboard.

3) That in all users, over both interfaces, the error rates would be greatest in the command line interface.

6.2.2. Parapsychological Hypotheses:

These were basically developments from the hypothesis (particularly Hypothesis 9, 10, 11 and 12 & 13 respectively) as set forth by Morris (1985). These major hypotheses are reproduced below (they will be reproduced in their original numbering sequence for ease of identification).

**Hypothesis 9.**

Psi input factors are involved in Human equipment interaction, such as to enhance and degrade system performance. Equipment operators may psychically access information relevant to equipment functioning such as to affect which equipment they select to operate, when and where they decide to operate it.

Within the KMDB system experiments these factors would be assumed to be determining which sessions Ss attended, and which machine/system pairing Ss used. In the present study the experimenter determined which machine/system pairing was issued to which subject, but the subjects still had the option whether or not to attend the experimental session. Subjects could either not attend any sessions, attend the wrong session, or only attend one

---

\(^5\) Competence in this case was judged to be if the subject was able to complete the series of questions at the end of the tutorial.
of the two sessions. These are only some of the problems involved in conducting experiments in ‘real-world’ conditions. The experimenter has to adapt to the subjects behaviour as best he or she can, without offending the subjects or breaching experimental protocol.

**Hypothesis 10**

Psi output factors (PK) also affect human equipment interactions and consequent system performance. Equipment operators and indirect operators alike may psychically influence labile components of the equipment.

As far as possible the Function Alteration Potential (FAP) of the systems were contained within the monitored smart noise task. However it is still possible that the subject’s influence was directed to enhancing un-monitored aspects of the task, such as the character definition produced by the electron guns in the screen, or adjusting background lighting conditions. Since it would be problematic to exclude all possible un-monitored aspects of a task, this study decided to only exclude those factors which could be controlled without intruding upon the subjects normal cognitive and biological activity.

**Hypothesis 11**

The degree and direction of psychic influence upon systems is affected by both state and trait personal factors, by physical and social environmental factors and lability factors in the equipment.

The personal factors were recorded by the various questionnaires used in the study. This data was then correlated against the recorded behavioral and parapsychological data. The lability factors (FAP) within the task was controlled within the constraints outlined in the comments on Hypothesis 10 (above).

**Hypothesis 12**

Noise plus FLP (Function Linked Person) and FLE (Function Linked Environment) equals a smart noise system. If the bit stream from such a system is interfaced with information processing and decision making components of a computer system, it can bias the computer system to function more intelligently to execute the goals of the FLP.
Hypothesis 13

Noise plus MLP (Malfunction Linked Person) plus MLE (Malfunction Linked Environment) equals a dumb noise system. It’s bit stream when interfaced with a computer will tend to act against the goals of the MLP.

In his 1985 paper Morris proposed two classes of MLP and FLP, the direct and hypothetical. The direct types would only be defined in terms of their actual performance on a SNS. In contrast the hypothetical types could be identified by various psychological predictive measures. For example Morris proposed that hypothetical MLPs would have several characteristics, these were: never having used a computer before, disliking or placing a low a value on technology, disliking learning new things, having low self-confidence and low self-esteem, being inclined to performance anxiety and having an external locus of control. Morris further predicted that hypothetical FLPs would tend to have the opposite characteristics. The experimental system was able to discriminate between both types of users, and investigate the empirical validity of these hypothesis. It was hoped that hypothetical FLP and MLPs could be identified by using information from the questionnaire data, particularly those aspects concerned with past success with high technology. Once these subjects were identified their SNS performance could be predicted, and then checked against the behavioral data. The direct FLPs and MLPs would be identified by the behavioral data recorded by the SSM. At the start of these experiments the author proposed some limited extensions to Morris’s hypothetical human function linkage types, these being the proposed existence of individuals who only influence a system’s performance. It is proposed to call these people ‘Performance Altering’ persons, PA-FLP and PA-MLP respectively. In contrast and perhaps in addition to these people are people who stop the functioning of that target equipment altogether. It is proposed to call these ‘Hardware Influencing’ persons, HI-FLP and HI-MLP respectively. It is also possible that these divisions overlap so PA&HI-FLP, and PA&HI-MLP might exist, or that may be divisions of the same scale 6. Obviously it will be preferable for experimenters to work with subjects who predominately produced effects within the performance alteration (PA) range. For a

6 These divisions may reflect the boundary between micro and macro psi effects.
full description of the classification system developed by the author for describing experimental anomalies see appendix 15.

6.3. Planned analysis.

The proposed analysis for each of the 3 studies were the same, since they were all linked. These proposed analysis are provided below. Note that throughout these analyses parametric tests were applied to data assumed to have the appropriate population distribution, and non-parametric tests were applied to those data sets which could not be assumed to have the appropriate population distribution (for example the parapsychological data was tested with non-parametric statistics).

6.3.1. HCI data.

First the HCI in-session recordings (session length, number of errors, number of commands, and the mean times for errors and commands respectively) would be correlated against one another by a simple parametric correlation of means. This allows a check to be made on the internal consistency of the data. Taking each of the hypothesis outlined above the planned analysis were as follows:

6.3.1.1. Hypothesis 1 Part I.

Wimp interfaces will have a slower throughput of 'work' than the command line system.

'Work' was defined as the number of commands the user issued, and the mean time between those commands. These figures are automatically produced from the users session logs, which are automatically produced by the system. These can be compared using a simple parametric t-test.

In addition to these in-session totals the average time between each command issued on each interface can be converted into a graphical representation. This allows any trends in the subjects inter-command times to be noted.
6.3.1.2. **Hypothesis 1 Part II**

Wimp interfaces have a longer session time than the command line system.

This is easily determined by use of the behavioral data in the session logs. The average time taken by all the subjects on each interface was taken to give an accurate indication of this factor. The totals from each interface can then be compared using a simple parametric t-test.

6.3.1.3. **Hypothesis 1 Part III**

Wimp interfaces will be preferred by naive users.

The study has a number of points at which the preference of the users can be detected. The last 5 questions of the post session questionnaires assess the users attitudes about the system used in that session. These questions came under the heading “What were your overall reactions to the system?”. The questions were the same on both the GEM and Command line post session questionnaires. These questions were 8 point Likert scales, asking subjects to rate their feelings on each of the following dicotomies: Good/Bad, Frustrating/Satisfying, Dull/Stimulating, Easy/Hard, and how ’powerful’ subjects (adequate power/inadequate power) rated the system as being. It was proposed to compare these ratings by using a non-parametric Wilcoxon matched pairs test.

6.3.1.3.1. **A Second measure of interface preference**

A second measure was to be from a post experimental essay that the subjects composed on the subject of which of the two interfaces they preferred and which they would recommend for use within their academic subject. However there were insurmountable problems with this measure, which will be discussed at length in the relevant chapters.

6.3.1.4. **Hypothesis 2 Part I.**

Once competence is reached in the command line interface, it will be preferred by experienced users and naive users who are 'keyboard literate’, due to the interface’s greater information interchange rate, and the familiarity with the keyboard.

It was proposed to determine this by comparing the data from the two preference measures outlined above, with the subject group’s background. It should be noted that this experiment
was unable to get access to subjects for any length of time. The author was therefore limited in what he could use a measure of 'competence'. It would require extensive experimentation to investigate this hypothesis fully, this was beyond the scope of this dissertation.

6.3.1.5. Hypothesis 3

In all users, over both interfaces, the error rates will be greater in the command line interface.

This hypothesis was to be evaluated in various ways. The primary measure was to look at the major error totals from the interface. These are mean number of errors, and the mean time between errors in seconds. These can be compared by a simple parametric t-test. It is also possible for the system to produce a graph of the time between errors. This allows an evaluation of any major trends within the error making behaviour of the two interfaces. Obviously both these are very crude measures of the overall error making behaviour made by subjects on the systems. However there are more subtle measures which are available. The interface recording system also records each error made and the exact nature of that error. Every possible error has a number and category associated with it. This detailed breakdown of error types allows the further discrimination into the categories of mistakes made by the subjects into semantic, system, or syntax errors.

6.3.1.5.1. Syntax errors.

Syntax errors occur where the user has mistyped either a verb, command or condition, or has incorrectly formed the command structure. For example if the user typed 'ADDD' instead of 'ADD', this would be a syntax error. It may indicate a lack of understanding of the required command syntax, or a mistake in command entry. To compare error rates on both systems we considered only those errors which are possible on both systems. This is done by removing those command line errors which are not possible in the Wimp system.

---

7 An example of the descriptions available for each error type is reproduced in appendix 5.
6.3.1.5.2. Semantic errors.

Semantic errors are those where a user has attempted to carry out an action on the system which did not work. This was either because the user had misunderstood the method required by the system to achieve that act (called an action error within this evaluation system), or the user did not understand what an object, for example a name, could be or do (called an object error within this evaluation system).

Many semantic error messages also overlap into syntax errors. It is possible to produce a category for errors called a 'probable category'. This allows us to remove errors messages which were not likely to be semantic errors, such as 'invalid student sex type'. In these cases it is highly probable that the user mistyped the gender field (syntax error), rather than the possibility that the user did not understand the concept of gender (male or female). By using this process we can arrive at a direct comparison between the types of errors made on the two interfaces.

6.3.1.6. Order effects.

Finally to test if any presentation effects had occurred we segregated the data into presentation order, and carried out a Kruskal Wallis Anova by ranks on that data.

6.3.2. Parapsychological data.

The system was designed to produce one overall psi measure. This became known as the 'Hit-Rate', and is the ratio of correct and incorrect guesses made by the SSM. All other measures which were created are purely exploratory, and do not supersede the 'Hit-Rate'. The TAQ questionnaire provides a series of questions (self report anomaly measure) which tried to record the frequency with which various high technology items acted as though they required repair or replacement. The subject was asked to state this malfunction rate on the following 5 point scale, daily, weekly, monthly, 6 monthly, or yearly, respectively. The items on covered in this scale were:

- a. Car, Motor Cycle, or Bicycle.
- b. Personal electrical appliances. {Electric Shaver, Hair Drier, Wrist watch etc.}
- c. Camera & Film.
• d. T.V., Video Recorder, or Personal Computer.
• e. Hi-Fi, Compact Disc Player, Radio, Tapes, Compact discs, and Records.

The subjects ratings on these questions were felt to give an accurate indication of their self perceived function linkage.

6.4. Planned Overall Analysis for combined results from experiments 1 to 3.

It was planned that when all three experiments were completed (and separately analysed) the results would be combined and various analysis performed. These analysis can be summarized:

6.4.1. Combined in-session recordings.

The Command, GEM and Overall in-session data (session length, number of errors, number of commands, and the mean times for errors and commands respectively. This excludes the psi measures\(^8\)) would be separately analysed using a parametric correlation.

6.4.2. Overall Investigation of the 3 main hypothesis.

Note that throughout these analyses parametric tests were applied to data assumed to have the appropriate population distribution, and non-parametric tests were applied to those data sets which could not be assumed to have the appropriate population distribution (for example the parapsychological data was tested with non-parametric statistics).

6.4.2.1. Work differences.

The overall work differences (as per hypothesis 1) would be determined by use of a parametric single t-test on the session details (session length, number of errors, number of commands, and the mean times for errors and commands respectively).

---

\(^8\) As with the previous in-session correlations (in the main experiments) these psi measures were included separately in the questionnaire sections, using non-parametric analysis methods.
6.4.2.2. Interface preferences.

The subject's preferences between interfaces would be compared by a non-parametric Wilcoxon Matched Pairs test on the last five questions of the UEICS2 post-session questionnaires, and by analysis of the post experimental essays.

6.4.2.3. Error Rates.

The interface differences in both the number of errors, and the respective error rates were to be distinguished by use of a parametric single t-test.

6.4.2.4. Presentation Effects.

The presentation order effects would be determined by inserting a presentation order code for each session and performing a 2 way Kruskal Wallis anova by ranks on the data.

6.4.3. Overall Parapsychological effects.

6.4.3.1. SSM (Skeptical Software Monitor) Performance Measure.

The main overall measure of performance is the overall SSM figure (hit rate), and a since this is a binary trial to hit ratio a two tailed Z score would be used to determine its significance.

6.4.3.2. Difference between hit rates over both interfaces.

It was possible that the command line or Gem system would prove to be more function or malfunction linked. It was thought this might prove to be associated with how satisfied each user population was with the interfaces. This would initially be investigated by comparing the hit rates (for all subjects) on both interfaces using a non-parametric Wilcoxon matched pairs test.

---

9 However the preference measure provided by the post session essays proved to be useless, due to problems in its administration (the final subject group would not have completed it), and in its content (there was strong evidence of subject collaboration)
6.4.3.3. Looking for any specific in-session data correlations with function or malfunction linkage (FLPs, and MLPs).

It was possible that some aspects of the subjects' normal' behaviour on the system, for example error rates, might correlate with their hitting or missing on the psi component. It was planned for this to be determined by the insertion of identifying codes within the data for subjects who consistently hit, missed, or varied direction, called hitters, missers, and MCE respectively. The data files would then be subjected to a Kruskal Wallis 3 way anova.

6.4.3.4. SSM Decision making sources.

It was hoped that if the smart noise concept showed any significance, then it might also be possible to determine which, if any, of the decision making sources was more successful (psi sensitive). To try and investigate this the data from the experiments would be processed into success rates and graphically presented to see if any source demonstrated an outstanding amount of positive SSM enhancement.

6.4.4. Investigation of overall questionnaire data and the recorded Command, GEM and Overall hit rates.

This would be conducted to see if there were any factors within the questionnaires which correlated with the SSM data produced by the subjects. Specifically the overall hit rates would be correlated against the TAQ data, and the Command and GEM hit rates would be correlated against their respective post session questionnaire data, using a non-parametric spearman rank order correlation on the respective hit rates.

6.4.4.1. Inter-Group Comparisons.

One of the major reasons for conducting this study was to investigate and compare the differences between respective groups of users. It was therefore important that the data sets be combined with an identifying group code. This allowed a 3 way Kruskal Wallis anova by ranks to be conducted to see what differences emerged from the three groups data. This was planned for the session, TAQ and UEICS2 data respectively.
6.5. Discussion on the number of planned analyses.

The nature of the study described in this dissertation was exploratory in the parapsychological, HCI, and human factors content. As we have already discussed in previous chapters, the study is unusual both in the effort it expended on creating the experimental system, and in the subject areas it addressed. In the case of the parapsychological content of this dissertation (computer anomalies and smart noise) there had been no previous systematic experimental work\textsuperscript{10} conducted. There had also been very little previous empirical work done in the area of controlled Command, and Wimp interface comparisons\textsuperscript{11}, parapsychological computer security, or function linked technology attitudes\textsuperscript{12}. Given these factors, and how unlikely it was that another individual would have access to a similar range of related data in the near future, the author felt justified in trying to gain as much useful information on these difficult problem areas as possible. The author extensively consulted statisticians during both the planning and analysis stages, and the nature and number of tests conducted on the data owe much to their helpful advice.

However in the context of the number of analyses conducted upon the data it is important to realise that the results of any one analysis cannot be taken as evidence for the existence of any effect. Rather the overall trends within the data should be taken as suggestive evidence that certain areas merit future investigation. It was never the purpose of the author to provide definitive answers to any of his experimental hypothesis.

\textsuperscript{10} Morris (1986) reported some correlations with unplanned experimental computer crashes.
\textsuperscript{11} The only remotely similar WIMPVs. Command studies the author is aware of have been privately conducted, and their findings are commercially restricted.
\textsuperscript{12} There has been some parapsychological investigations into machine breakdowns both by Bob Morris and researchers such as Dr. Michael Shallis. However none of these were able to correlate empirical measures of machine influence to measures of reported function linkage or technology attitudes.
6.6. Method

The method used for the pilot and 3 major experiments was basically the same and will therefore be described in detail here, and any variations which occurred will be noted in the description of each experiment.

6.6.1. Background.

SG and the author (K) agreed to an arrangement in which SG would provide the experimental facilities and subject populations. In return SG would be given co-authorship of any HCI publications produced by the study, and would be acknowledged in any parapsychological publications.

SG negotiated with the respective heads of department (HOD) responsible for access to students who matched the required subject populations. If the HOD was in agreement for the experimenters to have access to the students for 3 hours, then the most convenient time was scheduled for access to the subject population. The potential subject group attended a one hour lecture given by SG. These lectures were conducted within one of the lecture

---

13 As it transpired SG played a much larger role in the study, and the author will never be able to repay her kindness, or tireless support.

14 At date of writing only the in-session HCI details of experiment 1 have been written up and submitted for publication. The paper 'When does a mouse become a rat? or... Comparing performance and preferences in Direct Manipulation and Command line environments', has been accepted for publication by the British Computer Society's international journal of computer science 'The Computer Journal'.

252
theaters within the Napier Main building. The lecture introduced HCI and the study’s database system (which was described in chapters 4 and 5). Each person attending was provided with two handouts, one on human interface design, and the other on the database application. The material in both of these cases were created by S.G. Towards the end of the lecture subjects were asked if they would be willing to participate in this HCI experiment. Those who agreed to participate were split into groups of 12, the determining factor being the subject’s lecture timetables. These groups of 12 Ss were then assigned to a particular session, and told when and where to attend their session. These sessions were usually within a day of the lecture. The necessary number of experimental systems (software & questionnaires) were created by the author, and SG ran off the necessary tutorial handouts. The experimental discs were given unique identifying numbers, and the subject number to be associated with that disc was recorded on a session sheet. Figure 78. Plate 1. Napier Polytechnic.

In the cases of experiment 1 and the pilot attendance formed part of the subject’s course work. In the other cases attendance was entirely at the discretion of the subjects. This resulted in some unpredictability in subject attendance.

Identical “SCOTMEDIA” educational floppy discs were used. These are double sided/density 96 t.p.i. soft sectored discs. Each disc was formatted freshly for each session and the system files copied in an identical order for each system/disc created, using the same 360k drive on the development machine. This means that every disc was guaranteed to have an identical file/track layout, and by implication would behave identically for each subject.

This is in addition to the hidden security markers placed on the disc (these were discussed in chapters three and five).

The order of interface presentation experienced by the Ss was determined by a coin flip at the start of each experiment. So for example odd subject numbers experienced the GEM interface first, and even subject numbers experienced the Command line system first if a flipped 10p coin showed heads, and vice versa for tails.
perimenter to record the unique machine identifying number which the subject used.

The experimental sessions were always carried out in the same location. This was within a large (circa 1968), 8 floor iron and concrete structure. The building is shown in plate 1. The experimental room was located on the first floor at the extreme right hand side of the building.

6.6.1.1. The Experimental Room.

Figure 79. Plate 2. View of Computers. Figure 80. Plate 3. Experimental Room.

Figure 81. Plate 4. Amstrad Machines. Figure 82. Plate 5. View From Room.

This was 30ft long and 25ft wide. The walls and ceilings were white. Lighting was from strip fluorescent tubes. Plates 2, 3, 4, and 5 shows the experimental room, the layout of the Amstrad machines, and the view from the windows in the experimental room. This view

19 The full colour plates from which these digitised images were taken are reproduced in appendix 4.
consists of the student’s car park, some tower blocks (extreme right hand side), and Arthur’s seat (an extinct volcano in the distance). During the experiments the author tried to ensure that each session was held in the same environmental conditions. The lighting was always on, and the temperature was kept at 68F. Ventilation was both via two top 'bar' windows and via air conditioning vents in the false floor (the room had once housed the polytechnic’s mainframe computer). Environmental variations were noted in the session reports. These session logs required the experimenter to record the following details for each session: the date, the start and finish times of the experiment, the location of the experiment, the experimenter’s mood state, the expected number of subjects, the actual number who attended, the respective numbers of Command line and GEM systems used, the barometric pressure, the weather conditions, the phase and position of the Sun and Moon, and finally comments on anything unusual which transpired during the session. A sample session log is provided in appendix 2.

At the time of the construction of the session log the experimenter had developed a very informal theory of psychic functioning (Morgan, 1987). This theory involved the sensing of the changing levels of various forms of electromagnetic radiation and gravitation by the human endocrine and nervous systems. It was proposed that this 'sense' would be via the minute changes in neural electroconductivity and chemical levels associated with variations in electromagnetic radiation, and gravitational fields. This theory relied upon such environmental effects as solar & terrestrial radiation causing behavioral changes in humans, and their environment. Examples would be the increased likelihood of accidents (or 'breakdowns') in certain geographical areas due to reduced human performance associated with increased solar activity. It was proposed this increased likelihood of an 'accident' would be unconsciously interpreted by the human organism as a 'feeling' of an accident befalling persons (including themselves) known to be in that 'accident prone' geographical area. This theory was never intended to cover all 'psi' phenomena.
6.6.1.2. Procedure.

SG ensured that the 12 experimental machines were reserved for each session, and usually collected the author from the Edinburgh psychology department before each experiment. The author brought the experimental interface systems with him. These systems were prepared on the Koestler computer systems, usually the night and early morning before the day of the experiment. About 5 minutes before each session the author went to the experimental room and prepared all the discs, tutorial sheets, and questionnaires. If it was the first session for that group of 12 Ss then when each subject arrived the author took their name and issued them with a subject number. Subject numbers were issued on a first come, first served basis, since it was impossible to tell which Ss would turn up at any particular time. The author arbitrarily selected which of the 12 machines to be used by that subject.

Before starting to use the system subjects completed the technology attitudes questionnaire (TAQ2). This was completed by subjects as they sat in front of the computer. Both the TAQ and UEICS were uniquely identified by the subject numbers. Subjects were not allowed to start their session without having completed the TAQ questionnaire first. The subjects experienced both versions of the system using a standard counterbalanced order within subjects design, so that half the group in each session were using either the Gem or command system. On the second session subjects names were matched with their unique subject numbers from the session log created in their first session. The author or SG then assigned

---

21 The subjects were also provided with the record details of each of the lecturer and subject files held on the database. This information was necessary to complete some of the tutorial exercises.

22 One of the machines was an IBM PC without mouse capabilities. This made it only possible to run the command line system on that machine.
which machine they would use for the second session. Many of the subjects did not complete the second session, so to get the necessary numbers to allow statistical comparison as many subjects as possible had to be run through each experiment. During the session the subjects worked through the tutorial sheet\textsuperscript{23} created by SG, with the author and SG providing help and clarification where necessary. At no point did the author or SG actually issue a command for a subject, so all recorded behaviour was exclusively determined by the subject’s actions. 8 minutes before the end of the hour session the author issued each subject with his or her respective post-session questionnaire; to record the users attitudes towards the interface they had just used. Subjects were not supposed to be allowed to leave until they had completed their questionnaire\textsuperscript{24}. At the end of the whole experiment subjects were asked to comment on their impressions and feelings about the interface either in an essay or verbally\textsuperscript{25}.

At the end of the experiment the questionnaires, session log, and discs were collected and transported with the author back to the Edinburgh psychology department. The behavioral data which had been gathered were then run through the suite of analysis programs in the same evening on which the experiment was run. This procedure was completed regardless of the time it took (often resulting in the author finishing at 2 am). The resulting raw and processed files were then encrypted and duplicated to separate secure sites. The master copies were stored at the experimenter’s flat. This site is extremely secure, since the author’s landlord is a retired senior army officer with an extensive firearm collection. The firearms are located within the author’s basement flat with iron bars on all skylights, full alarm systems, and guard dogs. The master copies were periodically checked with the copies held elsewhere (more details of the security arrangements are given in chapter three).

\textsuperscript{23} These tutorial sheets are too large to be included as an appendix so, as in the case of the 71,000 lines of program source code, they are available on request from the author.

\textsuperscript{24} On 3 occasions subjects failed to complete their post session questionnaire within the 1 hour session, and would have been late for their next lecture had they been required to stay and complete the questionnaire. It would have been beyond the authority of the experimenters to enforce the subject to be late for their next lecture. In these cases the subject was allowed to take the questionnaire away with them, and return it to SG as soon as possible. The subjects concerned were noted in the session log. In all cases they returned the questionnaire to SG.

\textsuperscript{25} It was found that these informal post experimental measures (especially the essay) were not valid measures of the subject’s attitudes. The full reasons behind this are discussed in the section on experiment 1.
A series of specially designed and written analysis packages were created to analyze the data from the experiment. These results were then correlated against the questionnaire data (both TAQ2 and UEICS2), from a specifically designed interactive questionnaire database system, using Symantec’s Q&A™ Version 2.0. The statistical analysis was conducted using Statsoft’s Stats+™ package. All data entry was double checked by SG and the author before analysis began.

6.6.1.3. Task selection procedures.

These are fully covered in the discussion of the SSM workings in chapter 5, and the HCI design in chapter 4.

6.6.1.4. Apparatus used at Napier.

The target machines were 12 IBM compatibles made up as follows:-

- 7 Amstrad 1512 Pcs (Intel 8086-2@8MHz : circa 1986), with two 360K floppy drives, 640K RAM memory (expanded from the original 512K), standard Amstrad black and white screen, standard Amstrad mouse supplied.

- 4 Amstrad 1512 Pcs (Intel 8086-2@8MHz : circa 1988), with two 360K floppy drives, 640K RAM memory (expanded from original 512K), colour screen (CGA standard), standard Amstrad mouse supplied.

- 1 IBM PC (Intel 8088 @4.77MHz : circa 1985), with two 360K floppy drives, 640K RAM memory, amber screen, no mouse.

These were numbered and positioned in the same places in each of the experimental sessions within an experiment, but these did vary from experiment to experiment. These changes are shown in the diagram 'Machine positions', provided in the description of each experiment.

6.6.1.5. Operating systems

Each system was supplied with an identical copy of the operating system PC/MSDOS 3.20. The system automatically set all systems to have the same date and time (14:01:00 HRS on the 10th of March 1988).
6.6.1.6. Possible artifacts in or around the experimental location.

- Major multi-lane road 600yrs away - environmental effects from vibratory and atmospheric contamination are unknown.

- The experimental room was located above the boiler room for the building, the resulting electromagnetic and heat pollution factors are unknown.

- The experimental room was situated next to a large printer room, and subjects sometimes shared the experimental room with other students who were logged into the other college services. When this happened it was noted in the detail for each session. These factors might well have had an adverse affect upon the experimental performance of the subjects in all the tasks, implicit and explicit, but were beyond the control of the experimenters.

6.7. Pilot Experiment.

6.7.1. Reasons for pilot.

The purpose of this initial study was run to test the experimental system, and the associated methods. Having developed around 20,000 lines of software on a non-target system (IBM XT286 machines) it was essential to test the system before it was used within an experiential setting. A series of tests had been conducted by the author at the parapsychology lab to evaluate the system, but these tests had been on the Intel 80286 processor under the IBM version of the GEM operating system, using the author’s supervisors, or friends as subjects, and therefore could not be regarded as realistic tests of the software. A pilot study was also necessary to allow the testing of the methodology in the actual settings in which they would be used.

The computer based implementation problems encountered before the pilot experiment are described in appendix 16.
6.7.2. Questionnaires.

The three questionnaires which were issued to the subjects needed to be exposed to use by actual target populations before they were used as measures in the main experiments. Although they had been given to various dummy subject populations, these had an un-representative background, either in their level of education, or computer experience. The subjects in the pilot experiment were encouraged to give their detailed comments on version 1 of both the TAQ and UEICS after they had used them. These comments were noted on the respective questionnaire databases, and detailed reports of the comments and suggestions were produced.

6.7.3. Notes on the Pilot Study’s Data Analysis.

6.7.3.1. HCI Data.

Since the pilot study was only designed to test the experimental methodology, it was decided that it was not worth the effort to produce the graphical representation of the data. Obviously none of the hypothesis which involved subject group comparisons could be evaluated in the pilot study.

6.7.3.2. Parapsychology Data.

Since no empirical weight would be attached to its parapsychological findings, it was proposed to adopt an exploratory attitude to the gathered data. This approach allowed the experimenter to investigate any promising avenues of investigation, in the hope of being able to form more concrete hypothesis for the following experiments.

6.7.4. Method and Design.

The major aspects of method and design have been described in the proceeding sections. This section will therefore only describe the differences that occurred within this pilot experiment.
6.7.4.1. Subjects

These were 46 students from various courses which required a computing content at Napier Polytechnic of Edinburgh. Of these 46 only 19 provided useful data. The subject group was made up of 9 females, and 12 males, of whom the average age was 20.71 years (minimum 18, maximum 33). Within these groups 17 were right handed, 3 left handed and 1 claimed to be ambidextrous. These subjects proved to have a low motivation towards the task, and especially the questionnaires.

6.7.4.2. System - Targets, Selection processes

The test (beta) version of the database system was used, and versions 1 of the TAQ and UEICS questionnaires respectively (copies of the questionnaires are reproduced in the appendix).

6.7.4.3. Apparatus

These were the 12 machines described in detail in the proceeding sections (See diagram, machine positions).

6.7.5. Procedure

The pilot experiment took place between the dates of the 7th and 14th of March 1988. Within this time 6 one hour sessions took place, the full details of which are held in the session log for this experiment. An example of the level of detail recorded for each experimental session is given in appendix 2.

26 This dramatic loss in numbers is due to problems with the experimental systems, questionnaires, and the subjects.

27 Investigation into those subjects who claimed to be ambidextrous invariably revealed the ability was either highly limited, or un-demonstratable.
6.7.6. Results

There were some major problems in the experimental systems which were revealed by this pilot study. Unfortunately these problems affected the data from both the experimental WIMP system, and the questionnaires. These problems were due to the students being dis-interested in participating in the experiment, software incompatibility problems, and the inappropriate target education levels of the questionnaires. Since the software problems invalidated any comparisons between the two interfaces, and the inappropriate wording of the questionnaires made the subject’s responses unreliable, little empirical weight should be placed upon the results from the pilot study. These problems and their solutions will be discussed in depth in the discussion section.

6.7.6.1. HCI

6.7.6.1.1. Taking our experimental hypotheses one at a time and looking at the data from the pilot we find:

1) WIMP interfaces will be preferred by naive users, but will have a slower throughput of work and a longer overall session time.

Due to the problems encountered running the GEM version of the software on the Amstrad computers, there is no meaningful way of comparing the recorded data from the two interfaces. The available behavioral data will therefore just be provided, without any attempt at statistical evaluation.

---

28 The questionnaires had been targeted at populations with higher literacy levels than those of the subject group used for this pilot experiment. In fairness it had never been intended to run this particular subject population (who were taking diplomas in various subjects) in the main experiment. The actual experimental populations used in the subsequent experimental series were undergraduates, postgraduates, and postdoctoral researchers respectively. However the subject group used in the pilot gave the experimental methodology the most thorough testing.
<table>
<thead>
<tr>
<th>Mean session time</th>
<th>Command Line</th>
<th>2228.05</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gem</td>
<td>890.1</td>
</tr>
<tr>
<td>Mean Number of Commands</td>
<td>Command Line</td>
<td>31.63</td>
</tr>
<tr>
<td></td>
<td>Gem</td>
<td>15.0</td>
</tr>
<tr>
<td>Mean Time between Commands</td>
<td>Command Line</td>
<td>70.43</td>
</tr>
<tr>
<td></td>
<td>Gem</td>
<td>59.34</td>
</tr>
</tbody>
</table>

The available data seems to show that the command line system was not only faster than the GEM interface, but also had a larger throughput of work. If this data was truly representative it would confirm part 1 of hypothesis 1. The surviving questionnaire data seemed to confirm Part two of this hypothesis, since the Gem interface was found to be preferred to the command line interface. This had to be evaluated by informal verbal accounts, written comments provided by the subjects on the post-session questionnaires, and rough comparisons of the data from those few subjects who had completed (and seemed to have understood) the post-session questionnaires. The number of subjects who completed both sets of questionnaires was below the minimum requirement for meaningful analysis by non-parametric statistics. The results must therefore be taken with extreme caution. These results showed that subjects felt there was little difference between how good the two interfaces were (N=8, T=6.5, Z=1.61, P = 0.1). but seemed to find the Gem interface more satisfying (N=8, T=3.0, Z=2.10, P < 0.05), easier (N=8, T=3.0, Z=2.10, P < 0.05), and more stimulating (N=8, T=3.0, Z=2.10, P < 0.05) to use. There was no difference reported between what the subject regarded as the interfaces 'power' (N=8, T=17.0, Z=0.14, P = 0.86), but as we have already mentioned this is a difficult concept for even the most expert users to comprehend.

2) Once competence is reached in the command line interface, it will be preferred by experienced users and naive users who are 'keyboard literate', due to the interface's greater information interchange rate, and the familiarity with the keyboard.

Since our pilot subject group was not among any of the experimental target populations no evaluation of this hypothesis was possible.
3) In all users, over both interfaces, the error rates will be greater in the command line interface throughout the training.

<table>
<thead>
<tr>
<th>Mean Number of Errors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Command Line</td>
<td>25.05</td>
</tr>
<tr>
<td>Gem</td>
<td>6.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean Time Between Errors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Command Line</td>
<td>88.93</td>
</tr>
<tr>
<td>Gem</td>
<td>132.85</td>
</tr>
</tbody>
</table>

These figures would seem to indicate that the command line system had a higher number of errors, and error rate than the GEM interface. However these must be weighed against the problems which we have already mentioned (it is hard to make many mistakes when the system you are using keeps crashing). Unfortunately the problems created by the fragile nature of the software was increased by subjects deliberately crashing the system.

6.7.6.2. Parapsychological findings.

Since most of the subjects in the pilot did not want to take part it would be surprising to find any strong psi effects. Very little parapsychological experimental work has been done with subjects who did not wish to be participating in the experiment. Most researchers would regard such a situation as being non-psi conducive, and predict either psi-missing, or chance results. Since the computer system could be made to crash by certain actions it is possible that MLPs would turn their attention to discovering or issuing these fatal commands, rather than influencing the SSM. This could have been investigated by an analysis on which subjects crashed the system most often, however many of the subjects were intentionally replicating actions that others had already found to crash the system, so such an analysis was not regarded as being useful. The main SSM FAA measures showed no significance, with 421 attempted trials, and 203 'hits' (Z = - 0.7). Since this was the pilot experiment, and the results were suspect due to the system faults, and the reluctance of the subjects to participate it was not felt to be worthwhile to produce the graphical representation of these results.
6.7.6.3. Questionnaires.

As we have already discussed (above) the questionnaire data was contaminated by so much artifact from the subjects misunderstanding, and deliberately spoiling questionnaires that the resulting data is highly suspect. The number of subjects who completed both systems was so small that it was felt to be inappropriate to use the post-session data in any exploratory analysis.

6.7.6.4. Results of the analysis:

The data produced by the system was checked to see if the results produced were internally consistent with what should have been expected. Using a Pearsons R test (N = 19, one tailed) on the session data it was found that the session length correlated with the number of commands issued (p < 0.001), number of trials attempted (p < 0.01), and the number of errors (p < 0.05). The number of errors correlated with the number of commands issued (p < 0.01), and the speed with which the subjects issued those commands (p < 0.01). These and other internal audit checks confirmed the validity of the data recording systems.

6.7.6.5. Results from main analysis using a Pearsons R test (N = 19, one tailed).

The TAQ data was cross correlated with itself and an interesting factor which persisted throughout the entire series of experiments emerged. This was that the subject number correlated with fear of damaging the computer, and other computer attitude questions. It would appear that the later subjects attended the experimental sessions the more afraid of computers they were. It could be that such subjects were actively avoiding contact with computers.

6.7.6.5.1. Age.

There was quite a range of ages in the pilot subject group, and some age effects emerged. The younger subjects enjoyed working out problems on their own more (p < 0.05), but tended to be less motivated to be good at everything they did (p < 0.01). Younger subjects also preferred to use the computer themselves (p < 0.05), and were less afraid of damaging the computer if they used it (p < 0.01), or of looking silly if they made a mistake while using it (p < 0.01).
6.7.6.5.2. Gender

This subject group was also quite balanced between the sexes, and some interesting gender differences emerged. Females reported finding the computer was harder for them to understand (p < 0.05), and also reported playing less video games (p < 0.05). Males were more likely to have a micro computer at home (p < 0.01), while females were more likely to own a typewriter (p < 0.01). Males were more likely to have attended a course to learn about computers (p < 0.05), but females had longer session lengths in the experiment (p < 0.05).

6.7.6.5.3. Early experiences.

Part of the TAQ investigated the subject’s childhood interactions with machines. However of all the aspects of the TAQ this section was taken the least seriously by the subjects, so little weight can be given to those correlations which emerged. The strongest trend was that subjects who had being raised in families which repaired equipment, and who encouraged exploration of equipment developed a more positive attitude towards machines, and computers. Although these findings must be taken with caution for reasons we have already outlined.

6.7.7. Discussion.

The results we have discussed were severely affected by software faults, inappropriate questionnaire literacy levels, and disinterested subjects. The subject populations in the main experiments would be more motivated, and have higher literacy levels. However some substantial redesign of the experimental measures were still undertaken. Since the subjects were disinterested, and the systems proved to be unreliable (see below), it was felt to be understandable that the parapsychological measures were within MCE.

6.7.7.1. Questionnaires

Many of the subjects could not understand the words used on the questionnaires, and the layout of the questions caused many of them severe problems. The first versions of the questionnaires were designed so the scale direction alternated frequently to avoid a potential response bias on one side of the scale. However the pilot study revealed that subjects
encountered difficulty in adjusting to these scale direction switches. Apart from simply not being able to understand the questions, problems associated with the scale direction were the most frequent criticism subjects made of the questionnaires. Examples of the kinds of words or concepts which subjects experienced difficulty with were: terminology, interface, task, appropriate, feedback, corrective-action, ambiguous, programmable calculator, library catalogue, and malfunctioning. Obviously these kinds of misunderstandings would be unacceptable in the main experiment so all 3 questionnaires were redesigned to avoid complex words, and to simplify the concepts used. The other major change was to reduce the amount of scale direction changes used, and ensure scales remained constant for a minimum of one page.

6.7.7.2. Instruction Sheets.

Subjects had problems with the instruction or tutorial sheets prepared by SG, both with regard to the words used, and the concepts expressed. The original tutorial sheets had been identical for both systems, and merely instructed the subject which actions to perform by explicitly naming the command name. For example the 'ADD' command. The subjects in the pilot experiment had difficulty using the tutorial sheets, especially on the GEM system. It was decided that the presentation of the tutorial would have to be re-worked, and two versions, one for each interface, be produced.

6.7.7.3. Hardware

It has already being noted that one machine (the IBM pc), had no mouse attached. This meant only allocating command line systems to that machine.

6.7.7.3.1. Clock problems

Some of the recorded results from the SSM suggested there was a possible computer clock malfunction on some machines. Detailed investigation showed that the real time clock back up batteries had been exhausted on some of the Amstrads. This results in a loss of the highest resolution of the system clock. SG ensured that the maintenance staff at Napier installed new batteries before the next session. However there were problems with this solution since the computer clock batteries are identical to those used in personal stereo systems, and some students exchanged exhausted personal stereo batteries for the new clock system batteries.
Fortunately the SSM can cope with any amount of disruption (non-randomness) in the random sequences feeding it’s decision making systems without losing its MCE integrity. This is due to the nature of the task, as explained in chapter 5.

6.7.7.4. Software

6.7.7.4.1. Gem crashes: “Gem radio buttons”.

It was found that when the Gem radio buttons in the 'SELECT' card were left un-selected, the Amstrad 1512 detected a fatal memory parity fault. An example of the 'SELECT' radio buttons are given in this diagram. This error could not be duplicated on the development machine (IBM XT286 - AT type). Both Amstrad and DRI were unable to give any advice as to why this fault occurred, and the reason remained a mystery for some considerable time. Since this was the major cause of data loss, other than subjects deliberately crashing the systems and the 'accident' (see below), the chance of blank (un-selected) radio buttons occurring in the system had to be removed. It was decided to change the 'SELECT' card to have pre-selected options which were syntactically illegal, but that which did not cause the Amstrad memory fault. This meant that subjects who did not select any options in the 'SELECT' card generated a recorded error to the system, as opposed to crashing the system. Later investigations into this problem showed that the version of the operating system used for the experiment had been implicitly designed for true IBM PC/XT/AT computers. Inside the code of the ProPascal GEM bindings Propascal reassigns several of the devices, including 'STERR' (the standard error output in the Pascal language), to the NUL: device. When the blank radio buttons generate an error message they write to devices such as STDERR, which

![Figure 85. The Select Card (showing 'radio buttons').](image)
in this case had been redirected to this NUL: device\textsuperscript{29}. On true IBM systems many of the peripherals demand the operating system addresses extra device cards. In contrast the Amstrad directly controls its peripherals without the need of extra cards, and memory addresses. It is proposed that when the redirection of output takes place within GEM, during the un-selected radio button operation on the Amstrad, the operating system writes to a non-existing location, and thus generates a memory parity error. If this explanation is correct these problems could have been solved by using a version of MSDOS which had been specifically written for the Amstrad. However at the time of creating the experimental system GEM was only supposed to run under DOS+ on the Amstrad 1512. Amstrad later released an MSDOS version for these machines.

6.7.7.4.2. The 'Accident'.

In preparation for the Thursday 10th March session (which took place late on the Wednesday evening through to early on Thursday morning: 20HRS to 02:00HRS respectively), the author removed several of the safety checks that had been built into the GEM system creation suite of programs. These safety checks double checked that everything had been written correctly to the experimental discs. This double checking process doubled the length of time involved in each system’s creation. The author removed these checks in an attempt to complete the preparation before 4am. The author was very tired, since he had been running experiments all that day, and had finished completion of the previous evenings work at 03:00 HRS the previous morning. Removing the safety checks allowed the author to complete by 02:30 HRS, but caused all the Gem systems to be used on the Thursday March 10th sessions to be inoperative. This would not have occurred if the safety systems were in place, and was therefore totally the author’s fault. It was resolved that to ensure such a waste of time, and loss of data should never happen again, the safety checks would never be removed again.

\textsuperscript{29} The NUL: device is a conceptual convenience to which unrequired output from a process can be directed. It acts like a black hole, and is usually implemented as a memory address beyond the limits of physical memory. For obvious reasons it is usually the fastest device in the operating system.
Summary

These problems highlighted the weaknesses of the experiment, and gave the experimenter a series of tasks to complete before the main data collection could begin. The questionnaires and tutorial sheets were redesigned, and the software 'bugs' were removed. The author was dissappointed at the way the pilot study had been run (due to the problems he encountered), but he felt that the experience allowed him to remove the major problems from the experimental method, and to demonstrate that the overall experimental design was sound.
7. Experiment 1.

Quote:
Nature will tell you a direct lie if she can.
Darwin’s Observation, Charles Darwin.

7.1. Introduction.

7.1.1. Time of the Experiment.

This took place between the 11th of May 1988 and the 30th of May 1988.


These were as outlined in chapter 6.

7.2.1. Subjects

These were 72 second year B.A. commerce students at Napier Polytechnic of Edinburgh (Sighthill court). The sessions were part of their coursework, and attendance was compulsory. This group was computer naive and keyboard literate. The statistics about this subject group were: Of the 55 who took part, 33 attended both sessions, 42 did the command, 46 did the Wimp. 10 of the subjects were male, 45 of whom were female. 3 of the subjects were non-uk nationals (European exchange students). The youngest in the subject group was 18, the oldest 24, 4 subjects were left handed.

7.2.2. System - Apparatus, Targets, Selection Processes

As outlined in the initial section.
Machine Positions are shown in the figure.

7.2.3. Procedure

This also as outlined in chapter 6, except for the following minor variations. There were 14 one hour sessions, preceded by a one hour lecture on human computer interaction given by S.G, where the subjects were informed on the nature of the study, how to fill in the questionnaires, and how to use the mouse and pull down menus.

7.3. Results

7.3.1. In-session recordings.

7.3.1.1. Experiment 1 Command Line Session Details, Parametric Correlation (n = 42)

Longer sessions were found to correlate with more errors (p < 0.01), more commands (p < 0.01), attempting more trials (p < 0.05), and a slower rate of issuing commands (p < 0.077). More errors were found to correlate with more commands (p < 0.01), and a faster rate of making errors (p < 0.01). Issuing more commands was found to be correlated with attempting more trials (p < 0.01), and issuing commands faster (p < 0.01).

7.3.1.2. Experiment 1 Gem Interface Session Details, Parametric Correlation (n = 46)

Longer sessions were found to be correlated with making more errors (p < 0.05), more commands (p < 0.01), and attempting more trials (p < 0.01). Making more errors was found to be correlated with issuing more commands (p < 0.001), attempting more trials (p < 0.054),
making errors more quickly (p < 0.01), and issuing commands more quickly (p < 0.01). Issuing more commands was found to be correlated with attempting more trials (p < 0.01), and issuing commands more quickly (p < 0.01). A fast error rate was found to be correlated with a fast rate of issuing commands (p < 0.066).

7.3.1.3. Experiment 1 Overall Session Details, Parametric Correlation (n = 55)

Longer sessions were found to be correlated with more errors (p < 0.01), issuing more commands (p < 0.01), attempting more trials (p < 0.01), and making errors more quickly (p < 0.59). Making more errors was found to be correlated with issuing more commands (p < 0.01), attempting more trials (p < 0.01), making errors more quickly (p < 0.01), and issuing commands more quickly (p < 0.09). Issuing more commands was found to be correlated with attempting more trials (p < 0.01), making errors more quickly (p < 0.065), and issuing commands more quickly (p < 0.01).

These correlations were taken as evidence that the in-session recording processes had functioned correctly.

7.3.2. Hypothesis 1 Part I

Wimp Interfaces Will Have A Slower Throughput Of Work.

In this experiment the term 'Work' is defined as the number of commands the user issued, and the mean time between commands. Looking at the data generated by the resident recording processes an apparent non-significant superiority in 'work' is shown for the command line system. The command line system shows an average number of commands of 39.667 and an inter-command time of 64.524 seconds, while the wimp interface recorded 37.130 as its average number of commands, and an inter-command time of 68.135 seconds. However since the command system had a automatic entry to the help system, if the return button was depressed with no other action, a possible artifact here would be if there was a disproportionate number of help commands being issued in one interface and not in another (help is counted as a command type). To check against this possibility the types of command issued can be seen in the graph "Average number of command types". This histogram shows
the number of the various command types issued on both systems ('WIMP' denotes the direct manipulation interface, and 'Command' the command line interface totals). It can be seen that a difference does exist, there being more of the help command type in the command line interface.

**Total Number Of Help Commands Issued Over Both Interfaces.**

<table>
<thead>
<tr>
<th>Command Line</th>
<th>Wimp</th>
</tr>
</thead>
<tbody>
<tr>
<td>115</td>
<td>30</td>
</tr>
</tbody>
</table>

This shows that users of the command system used the on-line help facilities significantly (p < 0.05) more than users of the Wimp system. This unbalanced use of the help system would have led to misleading totals, and to evaluate hypothesis 1 in this study we must correct for this bias. So if we remove the number of helps issued over both interfaces we find:

**Average Number Of Commands**

<table>
<thead>
<tr>
<th>Command Line</th>
<th>Wimp</th>
</tr>
</thead>
<tbody>
<tr>
<td>36.928</td>
<td>36.478</td>
</tr>
</tbody>
</table>

From these totals we can adjust the mean time between commands to be:

<table>
<thead>
<tr>
<th>Command Line</th>
<th>Wimp</th>
</tr>
</thead>
<tbody>
<tr>
<td>66.712 seconds</td>
<td>65.650 seconds</td>
</tr>
</tbody>
</table>
After having corrected for the 'help bias' the difference between the two interfaces is seen to have disappeared for number of commands, and to have exchanged places for the rate of work. Where previously the Wimp system had shown a greater time gap between commands (68.155 seconds to commands 64.524 seconds), now after adjustment for helps the command interface can be seen to have been slightly slower (although not significantly so). This means that work throughput for the two interfaces was found to be almost identical, therefore part 1 of hypothesis number one has not been validated by this experiment.

**Time Between Commands.**

This figure shows the mean time between commands recorded by the system during the sessions for all the subjects over both interfaces (the graph terminates at the end of the session). The graph suggests that the command line system is faster in the early part of the study (as would be expected), but this graph also shows that the difference in the command line systems command rate decreased as the session progressed. There are at least two possible hypotheses that can be put forward to explain this, increased task complexity, and increased task boredom. Taking these in turn, the tasks that the tutorial sheets required the subjects to undertake became more and more complex as the session progressed. It is proposed that the users of the command line system displayed this increased task complexity by an increased inter-command time gap; this would be used as increased think time by the user. The Wimp interface user also has a rise in inter-command time gap towards the end of the session, but this is very slight.
Alternatively the increased task boredom interpretation would state that the user of the command system has far fewer stimuli, and is therefore subject to task boredom at an earlier time than the Wimp user who has much greater stimulation for the entire task duration. There is some support for the notion that the task boredom explanation is more likely and the comments given in the post session questionnaires and post experimental essay clearly indicate that most users found the Wimp system more stimulating.

### 7.3.3. Hypothesis 1 Part II

**Wimp Interfaces Have A Longer Session Time.**

Taking the recorded data we find the following

<table>
<thead>
<tr>
<th>Command Line</th>
<th>Wimp</th>
</tr>
</thead>
<tbody>
<tr>
<td>2463.595 seconds</td>
<td>2394.804 seconds</td>
</tr>
</tbody>
</table>

This shows a non-significant difference of 68 seconds in session length for the command line system. This means that part two of hypothesis one failed to be confirmed. During this experiment with computer naive keyboard literate users the command interface had a longer session length than the WIMP users.

### 7.3.4. Hypothesis 1 Part III

**Wimp Interfaces Will Be Preferred By Naive Users.**

The study has number of points at which the preference of the users can be detected. The last 4 questions of the post session questionnaires assess the users attitudes about the system used in that session. These questions came under the heading “What were your overall reactions to the system?”. By using a Wilcoxon matched pairs test a measure of user preference can be determined.

Subjects rated the WIMP system as better overall than the command line system ($N = 30 T = 138.500 Z = 1.933 P < 0.050$). The WIMP system was also rated as more stimulating ($N$
= 30 \text{T} = 111.000 \text{Z} = 2.499 \text{P} < 0.02), as being easier (N = 30 \text{T} = 137.000 \text{Z} = 1.964 \text{P} < 0.05), and as having more adequate power (N = 26 \text{T} = 108.000 \text{Z} = 1.714 \text{P} < 0.10).

Taking these findings as the "after use" affect we can clearly conclude that the subjects felt that the Wimp system was better in some very significant ways. It is not surprising that they found the 'feel' of the system more attractive and stimulating, but that they felt that the Wimp system had more 'power' is in direct contradiction to what most researchers in the HCI field would have anticipated (Shneiderman & Mayer, 1979). Power is generally equated with command line environments, and yet in this study we found a strong suggestion that touch typist users (to whom keyboard skills would be second nature), found the Wimp environment more powerful. It will be extremely interesting to see if this continues in the following studies with different subject populations.

Second Measure Of Interface Preference

A second measure is from the post experimental essay that the subjects composed on the subject of which of the two interfaces they preferred and which of the two interfaces they would recommend for business use. This data on the users preference is clouded by a possible artifact. Two things became obvious when the post experimental essays were read. The first was that the subjects all presented very coherent and comprehensive descriptions of the factors involved in the two systems, even the 17 students who did not take part and never attended any of the sessions. It must be assumed that some collaboration had taken place among the subject population about what to say in this essay. The other factor observed by S.G. was that the last interface used seemed to be preferred by the subjects. This was confirmed by an analysis, which split the data into presentation order.

Preferences From The Post Experimental Essays

Subjects who experienced the Wimp system last preferred the Wimp system in their essay, and subjects who experienced the command system last preferred the command system ($H(1,N= 31) = 9.475 \text{p} < 0.01$). It was a surprise to find that the command system was preferred by those people who experienced it last. It is of course possible that as the user gained experience with the system they changed to preferring the command system (as predicted by hypothesis 2). However our interpretation of this finding is that we are seeing the factor of the most recent experience influencing a compulsory essays content, which is
not unknown in an academic situation, and this must call doubt upon the preference data gathered from these essays.

7.3.5. Hypothesis 2 Part I

Once competence is reached in the command line interface, it will be preferred by experienced users and naive users who are 'keyboard literate', due to the interface’s greater information interchange rate, and the familiarity with the keyboard.

This subject group was keyboard literate (having been taught touch typing for 2 years previous to the experiment), and yet from the data we find strong evidence that these subjects preferred the Wimp system, thus disproving hypothesis 2 (part 1). It remains to be seen if an expert keyboard literate group also prefer the Wimp environment.

Hypothesis 3

In all users, over both interfaces, the error rates will be greater in the command line interface throughout the training.

If we look at the recorded data for errors we find the following :-

<table>
<thead>
<tr>
<th></th>
<th>Command</th>
<th>Wimp</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Number of Errors</td>
<td>29.9</td>
<td>12.9</td>
<td>6.82</td>
<td>&lt;&lt; 0.01</td>
</tr>
<tr>
<td>Mean Time between Errors</td>
<td>109.2 secs</td>
<td>237.8 secs</td>
<td>-5.45</td>
<td>&lt;&lt; 0.01</td>
</tr>
</tbody>
</table>

These results show that a highly significant difference exists between the number of errors (and the mean time between errors) over both interfaces, the Wimp interface having a highly significant advantage in both cases. This confirms experimental hypothesis number 3.
Time Between Errors\(^1\).

This figure shows that the average time between errors for the Wimp interface is longer than that for the command line interface. However this difference is reduced as the session length increases. The graph shows a dramatic peak in the beginning of the session, this would correspond with the time between reading and comprehending what was required of the subject from the tutorial sheet. After that initial period (the first to third error) the two graphs settle down to a fairly constant frequency pattern, except the two very large peaks on the Wimp system (errors 19, 26 & 27). Since the largest number of errors made on the Wimp interface was 33, the graphs show a zero error rate from that point on.

Types of Errors.

Syntax Errors.

The mean number of syntax error messages found on the WIMP system was 9.6, compared to 13.8 on the command line.

Semantic Errors.

The mean number of semantic errors found on the WIMP was 4.9, compared to 14.4 on the command line.

\(^1\) Note that these are the time between errors, and that the error number only indicates the order in which the errors occurred, not their identification number within the system.
**Command Line Error Types.**

This figure shows an almost even match between syntactic and semantic errors in the command line system. This suggests that the lack of a clearly provided 'mental map' (Nelson, 1980; Rutkowski, 1982; Hutchins, Hollan, & Norman, 1986) of how the system works in command line interfaces makes the cognitive load for both syntax and semantic detail about equal. The user is left to create their own mental model of what to use and how they are supposed to achieve the task.

**Wimp Errors.**

In contrast this figure shows a large difference between syntactic and semantic errors for users of the Wimp based system. We propose that this is caused by the clear and imposed mental model of how the system works that is presented by a graphical interface. Such a system drastically reduces the amount of cognitive load experienced by the user, and this reduces the number of semantic errors.
Overall.

This figure shows the detail from both interfaces combined. In this diagram, the difference in the number of semantic errors created by each interface is shown clearly. Although the Wimp system is better with regard to all error types, the difference is less marked for syntax errors.

7.3.6. Order Effects.

To test if any presentation effects had occurred we segregated the data into presentation order, and carried out a Kruskal Wallis Anova by ranks on the data.

We found slight (non-significant) increases in-session length and number of errors made in the second sessions, when compared to the first sessions. We also found an increase in the number of commands issued on both interfaces, consistent with a simple learning effect, and the increased relaxation that subjects would experience in their second session. It is proposed that the slight difference found in favor of the Wimp sessions is due to the increased "user friendliness" of that interface, which would enhance the learning/relaxation effect.
3 different screens were used in the study, the Amstrad 1512 mono, the Amstrad colour (CGA), and IBM PC green Monochrome screen respectively. Each of these screen types might have been a cause of some variation in performance and 'impact' on the subjects. The screens would also have had different speeds and influences upon the systems operation. The colour screen running the Wimp version would have been the slowest in it's refresh rate (machine numbers 7, 8, 9, and 10), and most 'pleasing' to the user due to the blue and white coloring used by the system. The next slowest would have been the Amstrad 1512 mono screens running Wimp (machine numbers 1, 2, 3, 4, 5, 6, and 12). The slowest of the command line systems would have been the IBM PC mono screen, run using the command line system (machine 11). However its amber screen did have much clearer character definition than any of the other screens when running the command line system.

During data analysis it occurred to the author that the colour screen/gem combination might have some affects upon the subjects performance, and satisfaction. He therefore carried out a post-hoc analysis looking at the strength of user satisfaction (from the post session gem questionnaire), and the recorded session data, with the screen type, which produced the following results.

7.3.7.1. Post Hoc Analysis One.

A Kruskal Wallis Anova by ranks was conducted on the Gem post session questionnaire 'satisfaction' (questions 19a thru to 19e) data looking for differences in subject response when they had a gem session using a colour screen as opposed to when they used a Mono screen. This revealed a suggestive (H(1, N = 44) = 3.324, p 0.065) link between the using a color screen and feeling that the GEM system was more stimulating.

7.3.7.2. Post-Hoc Analysis 2

Next A Kruskal Wallis Anova by ranks was conducted on the GEM interface behavioral data looking for differences when subjects used colour screens. This revealed two significant differences between the Gem session data for users of color and mono screens. The first was between the machine identification number and screen type (H(1, N = 45) = 3.720, p < 0.051) (which was expected), and the second was that users of the color screens issued more
commands than users of the mono screens \((H(1, N = 46) = 3.625, p < 0.055)\). This finding seems intuitively valid and could be of interest to the field. It suggests that colored graphical environments are significantly preferable to simple monochrome ones.

### 7.4. Parapsychological Findings

The overall psi measure from this experiment revealed no significant effect, with 1776 trials\(^2\) and 895 hits respectively \((Z = + 0.33)\).

![Psi trials](image)

**Figure 93. SSM Trials, Experiment 1.**

**Diagram of the SSM Trials.**\(^3\)

#### 7.4.1. Function/Malfunction Linked Operators.

Out of the 33 subjects who attended both sessions 10 people scored consistently in one direction over both sessions. 4 of these were 'hitters' (PA-FLP) and 6 were 'missers' (PA-MLP). Taking the data on Performance Alteration function linkage (using the codes

---

\(^2\) A trial is one attempt to guess the next direction in the b-tree, see chapter 5 for a detailed explanation.

\(^3\) Note graph plots cut off where N less than 10% of total.
0=PA-MLP, 1=PA-FLP), the author performed an exploratory nonparametric Spearman rank order correlation on the TAQ2 results and found the following correlations. MLPs were found not to be afraid of looking silly if they made a mistake while they used the computer (N= 10, R= -0.613, Z= 1.84, p 0.063), this was not in the expected direction. Flps reported never using a computer at home (N=10, R= -0.573, Z= 1.72, P 0.082), this was also not in the intended direction. Since these were the 2 only correlations which approached significance from the TAQ2 data it was concluded there was little evidence from the MLP or FLP correlations for any paranormal FAA (function alteration activity) in experiment 1’s data.

7.4.1.1. Function/Malfunction Linked Systems.

There was no evidence that either interface was found to be more or less Function linked. A Wilcoxon matched pairs test on the hit rate revealed no significance either overall (N=33, T= 249, Z= 0.56, P 0.55), or the selective FLP/MLP data (see table above) (N=10, T= 17, Z= 1.07, P 0.28).

7.4.1.2. What Does a MLP Look/Behave Like?

In the exploratory analysis which was conducted on the TAQ2 data from this experiment, 2 subjects fitted our expectations of a MLP, from their self reported machine interactions. Unfortunately there were none who fitted the category of an FLP. These MLP subjects had their data analysed in detail to see if any unusual traits could be detected which could be used as predictive measures in future analysis. The subjects concerned were subject number 29, and number 1 respectively, both of whom were females. There was one male who reported frequent
machine problems, but this turned out to be due to his having to renew his personal stereo batteries on a daily basis\(^4\).

Subject 29 although listing herself as MLP in the TAQ2, did not in fact show any evidence of Malfunction Linked Activity (MLA) during either of the two sessions, both her sessions were well within MCE. However Subject 1 also rated herself highly on the MLP scale in the TAQ2, and displayed significant MLA (This can be seen in the SSM performance graphs for Subject No. 1). This subjects 'psi' based score was highly significant (Command line system recorded 7 hits out of 24 trials ($Z = -2.0412$), the Wimp system recorded 2 hits out of 10 trials ($Z = -1.89$), giving an overall score of 9 hits out of 34 trials $Z = -2.70$), in the negative direction. After thoroughly checking the system no software based explanation for this highly significant deviation from MCE could be found.

The author concluded four possible explanations\(^5\):

1) Untraceable system malfunction. This is unlikely to have followed the same subject over two different systems and discs (disc gem1, and machine No. 7 on session no. 1; and disc com2 and machine number 7 on session no. 9). It is possible that machine number seven was to blame, but a detailed analysis over each machine revealed no significant bias for 'psi' to occur on machine 7.

---

\(^4\) The subject reported using the personal stereo 8 hours a day, so it was felt that he was getting normal performance from the power cells.

\(^5\) Noting that there are numerous other explanations.
2) Subject fraud of some nature. I find this an unpleasant thought, given the extreme security arrangements taken (and the covert nature of the study). This explanation would require either one of the Chair’s members or one of the two experimenters to have collaborated with the subject to commit the fraud. It would also have required a considerable amount of computing expertise, and resources. Although possible this option is extremely unlikely.

3) The subject is a true MLP demonstrating MLA.

4) The results are due to chance.

Because of the size of the effect it was felt that it would be worthwhile to follow this up by asking subject 1 to have another session. This would enable the author to interview her to find out which of the options listed above was most likely. This would of course have formed a separate investigation. Unfortunately it was found that she was in fact one of the German exchange students, and had returned to Germany. This made any follow up impracticable.

7.4.1.3. Decision Making Systems And Randomness.

Several decision making based recordings have been made in the study. This section attempts to give an analysis and interpretation of the decision making recordings from this study.

Histogram Decision Making Sources.

The SSM sources show the percentage accuracy of the various decision making sources within the experiment. None of the source’s success rates approach statistical
significance. In these diagrams '0' indicates MCE, positive values are indications of correct guesses, and negative values indicate incorrect guesses. None of them show any obvious pattern, except displacement. The unavoidable artifact associated with that source is discussed in depth in appendix 13, and a check of the in-session logs confirm that the subjects were indeed occasionally issuing the same query twice. Multiple attempts to retrieve the same record causes the displacement checker to 'learn' from the previous trial. However since the subjects are blind to the SSM, and the working of the displacement checker, the increase in the sources predictive accuracy is tolerable, and has been simulated in the control trials.

### 7.4.2. Questionnaire Findings.

Several analysis were performed on the data from the questionnaires These are detailed under separate headings.

**TAQ2 (N= 55).**

First the TAQ2 was correlated with the overall recorded session data, including obvious correlations, such as those detailed in the review of the pilot study. Within the TAQ2 data a non-parametric Spearman rank order correlation between the major factors of interest revealed the following :-

**TAQ2 and Recorded Session Data (N= 55).**

Session length.

Shorter session lengths were found to be associated with enjoying activities which involved doing things in a specific order (p < 0.061), not enjoying repairing machines (p < 0.058), not being allowed to use dangerous equipment when young (p < 0.05), being afraid of looking silly making a mistake while using the computer (p < 0.088), and having few problems with photographic equipment (p < 0.05).

Errors.

Making fewer errors was found to be associated with not being allowed to use dangerous equipment when they were young (p < 0.05), being afraid of looking silly making a mistake
while using the computer (p < 0.05), and being afraid of looking silly if seen not know how to use the computer (p < 0.05). Subjects who were bothered by the impersonal nature of computers made errors more slowly (p < 0.05).

Commands.

Issuing fewer commands was found to be correlated with having few problems with photographic equipment (p < 0.01), being allowed to use dangerous equipment when they were children (p < 0.05), being afraid of looking silly when making a mistake while using the computer (p < 0.05), and being afraid of looking silly if seen not to know how to use the computer (p < 0.05). Issuing commands more slowly was found to be correlated with being bothered by the impersonal nature of computers (p < 0.05), and preferring to have some one else use the computer for you (p < 0.05).

Function Alteration correlations.

Enhancing the SSM performance (psi hitting) was found to be correlated with making more hits (p < 0.01), wanting to be good at whatever one did (p < 0.05), never using a bank teller (p < 0.068), never using a micro at college (p < 0.07), being mildly injured by a machine in childhood (p < 0.68), and never helping one’s parents repair machines (p < 0.06).

TAQ2 Inter Questionnaire Correlations (N = 55).

Within the TAQ2 data a non-parametric Spearman rank order correlation between the major factors of interest revealed the following :-

Attitudes towards computers.

Subjects who enjoyed figuring out how things work, also liked to repair machines (p < 0.001), and did not think that computers were too hard for them to understand (p < 0.001). Understanding and dependability were closely correlated, such that subjects who reported that computers were not too complicated to understand felt that a computer could be more dependable than a human being (p < 0.015). Subjects who felt that computers were too complicated to understand were afraid of damaging the computer if they used it (p < 0.001), of looking silly if they made a mistake (p < 0.014), or were seen not to know how to use a
computer (p < 0.025). Thinking computers were too complicated also correlated with being bothered by the impersonal nature of machines (p < 0.01), and disbelieving that computers helped them to work better (p < 0.021). Subjects who read books or articles about computers above those required for their course said they also enjoyed figuring out how things work (p < 0.001), and performing tasks which require doing things in a specific order (p < 0.001).

Sex Differences.

Sex differences were found to be quite strong in the following areas: Males used digital watches more frequently (p < 0.01), played more video games (p < 0.031), helped their parents to repair faulty equipment (p < 0.053), and were more frequently allowed to use dangerous equipment (p < 0.01). Males were encouraged to explore how machines worked when they were young (p < 0.01).

Early life.

Subjects that helped their family repair equipment when they were children had the tendency to enjoy figuring out how things work (p < 0.05), and liked working with tools or machines (p < 0.001), and liked to repair machines (p < 0.00001). These subjects were also more likely to use a programmable calculator (p < 0.05). Subjects who were allowed to use dangerous equipment when they were young developed an enjoyment of repairing machines (p < 0.001), and were not afraid of looking silly if they made a mistake while using a computer in later life (p < 0.05), or were seen not to know how to use a computer (p < 0.05). The importance of early influences was also shown by the fact that subjects who were encouraged to explore how machines worked developed a liking for repairing machines in later life (p < 0.0001).

Function linkage.

The MLP/FLP syndrome, is one in which little work has been done since Prof. Morris outlined his conceptual framework (see previous chapters for details). This section presents the results from the TAQ2 questionnaire which tried to investigate this area. Throughout this section the reader should remember that the FLP/MLP syndrome can exist independently of any psychic component.
Subjects who said that at their house machines seemed to break down when they were young reported being afraid of looking silly if they made a mistake while using the computer ($p < 0.01$), or if they were seen not to know how to use the computer ($p < 0.01$). They also reported being bothered by the impersonal nature of computers ($p < 0.05$). Subjects who were frightened by some kind of equipment when they were young grow up being afraid of looking silly if they made a mistake using a computer ($p < 0.02$), and if they were seen not to know how to use the computer ($p < 0.02$). Subjects who helped their family to fix equipment when they were young had fewer problems with break downs with audio equipment ($p < 0.05$), and read more books about computers in their spare time ($p < 0.05$).

Subjects who were at least mildly injured by a machine when they were young had more frequent malfunctions with mechanical equipment ($p < 0.05$), personal electrical appliances ($p < 0.05$), and photographic equipment ($p < 0.05$). Unexpected correlations found with those who had been injured when they were young was that they read more about computers ($p < 0.05$), and were likely to have been allowed to help repair equipment with their parents ($p < 0.05$). Subjects who were encouraged to explore how machines work when young reported more problems with audio equipment breaking down ($p < 0.05$) this was not in the expected direction, but could be due to them dismantling the equipment. This tendency was certainly found in the final experiment with technocrats.

Subjects who reported that machines seemed to break at their home when they were young to have frequent break downs with mechanical equipment ($p < 0.05$), personal electrical appliances ($p < 0.05$). However these subjects tended to use multi-function digital watches ($p < 0.056$), and play video games ($p < 0.05$). These latter correlation are not in the expected direction.

Subjects who reported being frightened by some kind of equipment when they were young had more frequent malfunctions with mechanical equipment ($p < 0.5$), personal electrical appliances ($p < 0.001$), and photographic equipment ($p < 0.05$).

7.4.2.1. UEICS2 - Experiment 1

The next sections will cover the correlations found within the post session questionnaire data. These will cover both interfaces, first the Command line and then the GEM interface.
All these analyses were done using a non-parametric Spearman rank order correlation. These correlations are important from the view of designing of computer systems since they allow us to measure those effects which had the greatest impact upon the users.

**UEICS2 - Command Line Post Session Questionnaire (n = 42).**

Correlations within the command line in-session recordings.

There were fewer significant correlations in the command line data, than in the Gem system. This might be because most of the subjects did not like the command line system, and were therefore less forthcoming with their feelings about the system on the post session questionnaire.

**Sex Differences.**

Males issued suggestively fewer commands (p < 0.082).

**Errors**

Subjects who found the error messages helpful had slightly slower error rates (p < 0.9).

Subjects who found the system stimulating issued more commands (p < 0.05).

**Function Linkage.**

Enhancing the SSM performance was found to be correlated with the number of hits (p < 0.001), and rating the interface as preventing mistakes (p < 0.01).

**Correlations Within The Command System Post Session Questionnaire**

The command system post session questionnaire data was correlated against itself to get an idea of the reliability of the data, and to try and find which aspects of the system were most important to the users satisfaction.

**Sex Differences**

Females reported finding that the command line system had too many tasks per action (p < 0.05), and was too fast (p < 0.05). Both these are interesting since we have established
that the GEM system was in fact slightly faster, and the number of tasks per action was exactly the same in both systems. Males rated the command line system as providing too little feedback (p < 0.05).

The Use of the Keyboard.

The ease with which the subjects could use the keyboard had a significant effect on their reactions to the whole system. Subjects who found the keyboard easy also reported that the command names matched the task closely (p < 0.05), that the information presented was always sufficient (p < 0.058), that work was simplified by the display layout (p < 0.05), that learning the system was easy (p < 0.05), the system was good (p < 0.01), satisfying (p < 0.05), stimulating (p < 0.05), and they felt the system had enough power (p < 0.05).

The Importance of Appropriate Command Names.

In the command system subjects seemed to rate the importance of the command name matching the action it completed more highly than in the gem system. This was probably due to the reduced overhead that such command names have in being memorized. The closeness between command name and its action was found to correlate with command consistency (p < 0.001), clarity of (tutorial) instructions (p < 0.01), prevention of mistakes (p < 0.05), appropriateness of feedback (p < 0.05), simplification of the work to be completed (p < 0.01), helpfulness of error messages (p < 0.01), and clarity of error correction (p < 0.05). The user’s overall reaction to the system was also correlated to the closeness of command names matching their actions. Users who reported a close match between the command name and the action it performed made subjects rate the system as good (p < 0.01), easy to learn (p < 0.01), satisfying (p < 0.001), stimulating (p < 0.01), easy (p < 0.01), and powerful (p < 0.01).

Just as the choice of command name was found to be important in command line systems so was the inter-command consistency within the system. The analysis showed that when the subjects felt that the commands within the system were consistent with one another then the tutorial was rated as being easier to follow (p < 0.05), subjects had the impression of having fewer actions per task (p < 0.055), error messages were rated as being more helpful (p < 0.05), and the system was rated as easier to learn (p < 0.001). Inter-command
consistency was also found to be correlated with subjects rating the system as good ($p < 0.001$), satisfying ($p < 0.001$), and easy to use ($p < 0.001$).

Clarity of the Tutorial.

Subjects who found the tutorial easy to understand rated the system as good ($p < 0.001$) easy to learn ($p < 0.01$), satisfying ($p < 0.001$), easy to use ($p < 0.05$), and powerful ($p < 0.01$).

Display Layout and Feedback.

Subjects who rated the command line system as preventing mistakes also reported that the system had a helpful display layout ($p < 0.01$), was satisfying ($p < 0.5$), stimulating ($p < 0.06$), easy ($p < 0.05$), and more powerful ($p < 0.01$) to use. The system was rated as having adequate feedback if the subject felt the display simplified the work that had to be carried out ($p < 0.05$). Sufficient feedback made the subject have the impression that the system was faster ($p < 0.05$), that error correction was easy ($p < 0.05$), and that on-line help was clear ($p < 0.05$). Surprisingly subjects did not rate adequate feedback from the system as making the system easier to learn or to use, but they did rate it as making the system good ($p < 0.5$), satisfying ($p < 0.01$), and powerful ($p < 0.05$). Display layout was also found to correlate with the helpfulness of error messages ($p < 0.01$), clarity of the on-line help ($p < 0.05$), ease of learning ($p < 0.057$), and the reactions to the system. When the display layout simplified the work it made the subjects overall reaction to the system good ($p < 0.01$), satisfying ($p < 0.01$), stimulating ($p < 0.01$), easy to use ($p < 0.05$), and powerful ($p < 0.01$).

Error Correction and Prevention.

Error messages were judged to be helpful if they provided the corrective action ($p < 0.01$), were clear ($p < 0.01$), and if on-line help was clear ($p < 0.01$). The error messages were rated as helpful if the screen displayed all the information required to complete each task ($p < 0.05$). Good helpful error messages made the subjects rate the system as being good ($p < 0.01$), satisfying ($p < 0.001$), and adequately powerful ($p < 0.01$). Closely related to these factors was the clarity of error correction advice, this was found to be correlated with satisfaction ($p < 0.05$), and the powerfulness of the system ($p < 0.05$). If the command line
system was judged to have few tasks per operation then the subjects rated the system as being easy to learn (p < 0.05), and good to use (p < 0.05).

Questionnaire Reliability.

How good the system was to use was found to be correlated with satisfaction (p < 0.00001), stimulation (p < 0.0001), power (p < 0.001), and ease of use (p < 0.001).

7.4.2.2. UEICS - Gem Post Session Questionnaire (n = 46).

Correlations Within The Gem In-Session Recordings.

Sex Differences.

The subject population only contained 10 (18% of subject population) males who used the GEM system so any sex differences should be taken as indicative and not validated, unless subsequent experiments show a similar trend. Males had shorter sessions (p < 0.05), and had a shorter inter-error time gap (p < 0.05), this would seem to indicate that males were less careful in their work than females.

Previous use of a mouse.

Subjects who had used a mouse before were more proficient in their use of a mouse based system. Such individuals issued more commands (p < 0.05), and issued those commands more quickly (p < 0.053).

Subjects were asked directly how they rated the system, from items which correlated from these ratings we can get an idea of which factors are most important to the user. Subjects issued more commands if they felt the system was good (p < 0.05), satisfying to use (p < 0.05). The more stimulating subjects reported finding the system the longer their session length (p < 0.05), the more commands they issued (p < 0.01), and quicker they issued those commands (p < 0.01). Subjects who reported that the system had adequate power issued more commands (p < 0.05), and issued those commands significantly more quickly (p < 0.05). However there was a very weak correlation found with how satisfying the users found
the system and how much they degraded SSM performance (p < 0.1), such that the more satisfying subjects found the system the more they degraded the system performance. This is not in the expected direction.

Session length was also influenced by the command complexity, subjects who felt that a command required very few actions in order to complete it had a longer session length (p < 0.05). Session length was also found to correlate with how good subjects rated the interface’s feedback (p < 0.05), display layout (p < 0.01). Such that subjects who rated the feedback and display layouts as better had longer sessions. Subjects who reported that the display layouts simplified their work also issued more commands (p < 0.061).

Error Correction.

Subjects who rated error correction as clear had longer session lengths (p < 0.05), issued more commands (p < 0.05).

Function Linked Correlations.

Enhancing SSM performance was found to be correlated with the number of hits (p < 0.001), clear error messages (p < 0.055), and rating the interface as being stimulating (p < 0.1).

Correlations within the GEM post session questionnaire.

The questions within the gem post questionnaire were also correlated against one another, first to try to get an idea of the reliability of the data, and secondly to try and find which aspects of the system were most important to the users satisfaction and dissatisfaction within the system. One of the factors of greatest interest to the study was why some people find one system preferable to another.

Acceptance of the mouse.

The mouse and its acceptance were found to be most important to the acceptance and ease of use of the wimp based environment. If the subject found the mouse easy to control then they also found the pull down menus easy (p < 0.05), and the simultaneous use of the keyboard and the mouse (p < 0.01). Ease of using the mouse made subjects report that the
tutorial sheet was easier to understand (p < 0.05), and error messages were rated as being more helpful (p < 0.01). If the mouse is found to be easy to use then the whole system was rated as being easier to learn (p < 0.05), good (p < 0.01), satisfying (p < 0.01), and easy to use (p < 0.05).

Other Aspects of the WIMP interface.

Subjects who found the pull down menus easy reported finding the work was simplified by the system (p < 0.05). Subjects who found that using the mouse and keyboard was easy also found the system satisfying (p < 0.05), and easy (p < 0.01).

Command Name Consistency.

It was reassuring that subjects who found a closeness of the word to the command name found that the commands were consistent (p < 0.001), had a good reaction to the system (p < 0.05), found it satisfying (p < 0.054), and stimulating (p < 0.05). Command consistency was also found to be important in other aspects. Subjects who found the gem system to have consistency in its commands also found levels of feedback to be sufficient (p < 0.052), display layout simplified tasks (p < 0.05), error messages helpful (p < 0.055). Command consistency was also found to correlate with the subject reporting that the system was good (p < 0.05), satisfying (p < 0.05), stimulating (p < 0.057), and easy to use (p < 0.01). The data was also suggestive that command consistency was made the system easier to learn (p < 0.075).

Feedback, Error Correction and Prevention.

Prevention of mistakes (as rated in the questionnaire by the subject) was clearly shown to be positively affected by the amount of feedback from each command (p < 0.01), helpful error messages (p < 0.01), and by the information required to complete each task being visible on the screen (p < 0.01).

Information about which factors are relevant to creating an impression of good feedback were helpfulness of error messages (p < 0.01), error messages indicating the corrective action (p < 0.001), and if the process of error correction itself was found to be clear (p < 0.01). The
importance of feedback is shown by the fact that if the feedback was considered good then the system was rated as being good \( (p < 0.01) \), and satisfying \( (p < 0.05) \) to use.

The helpfulness of the error messages was found to correlate with how accurately it indicated the corrective action to be taken \( (p < 0.001) \). Simply stating the thing that was wrong is not enough, the system should tell the user what to do to correct the error. The clarity of the error messages has an affect upon the ease of learning the system. The better users rate the error messages the easier they rate learning the system \( (p < 0.05) \), and the satisfaction derived in using the system \( (p < 0.05) \) is increased. The helpfulness of the error messages were also found to be correlated with the ease of learning the system \( (p < 0.05) \), how satisfying the system was to use \( (p < 0.05) \), and how easy the system was to use \( (p < 0.05) \).

The ease of error correction was found to be connected with how satisfying \( (p < 0.05) \), stimulating \( (p < 0.05) \), and easy it is to use the system \( (p < 0.063) \). Also of importance was the how clear subjects found the on-line help. The clarity of on-line help was found to correlate with how easy subjects found pull down menus \( (p < 0.05) \), how much subjects reported that display layouts simplified their work \( (p < 0.01) \), and how good the subjects felt the system was \( (p < 0.055) \).

Display layout.

The questionnaire data showed that a good display layout (one which helped simplify work) also made the system seem fast \( (p < 0.05) \), and easier to learn \( (p < 0.05) \). Good display layout also made the system appear good \( (p < 0.01) \), and satisfying \( (p < 0.05) \) to use. The users impression of the speed of the system was found to influence their satisfaction (more speed, more satisfaction) \( (p < 0.05) \).

Questionnaire Reliability.

Evidence for inter-question reliability was found in those questions which had asked the subjects to rate the system. So that good systems correlated with how satisfying \( (p < 0.0001) \), stimulating \( (p < 0.01) \), and easy \( (p < 0.01) \) systems were to use. Satisfaction and easiness were also found to be correlated \( (p < 0.001) \). Confirmation of this inter-question reliability
was also taken to be shown by clarity of task instruction (tutorial) being correlated with the ease of learning the system (p < 0.001). Ease of learning the system was found to be correlated with how good (p < 0.01), satisfying (p < 0.001), and easy (p < 0.001) the system was to use.

7.4.2.3. Post-Hoc Comparison Of The Data From The Two Questionnaires.

Looking at the information from the two post session questionnaires together we can try to glean something about the differences people felt about the two systems.

It is reassuring that subjects ratings for command consistency over the two post session questionnaires matched very closely (the commands were identical on both interfaces) (p < 0.05). Also reassuring from a consistency point of view is the finding that subjects who found on-line help confusing did so on both systems (p < 0.01).

Ease of Use.

Females found the mouse was easier to control (p < 0.05), but regardless of gender if the subject had used a mouse based system before then they found using the keyboard in the command line system was hard (p < 0.05). Those subjects who found the command system dull also found that the mouse was easy to control (p < 0.05). Subjects who had used a mouse driven system before found that the command system was too fast (p < 0.05).

7.5. Discussion.

7.5.1. HCI Interface Discussion.

We have described the results from the in-session recording software from a specifically designed HCI interface comparison tool. The small sample size (55) reflects the increasing difficulty in obtaining truly computer naive user/subjects. Our findings with keyboard literate computer naive users strongly suggest that WIMP interfaces are faster, easier, less error prone, and more stimulating for naive touch typists than a functionally identical command line interface. Further research is needed to see if the WIMP interface’s superiority was maintained if the subjects used the system over a long period. We have also
found a marked difference in the types of errors (semantic and syntactic) produced in the two systems, even after they have been adjusted to exclude biases.

7.5.2. **Parapsychology**

The overall SSM FAA measure gave no indication of any consistent function alteration ($Z = +0.33$), either in the hitting or missing direction. This was a disappointment to the author, who would have liked to have seen the SSM system he spent 8 months creating 'in action'. The two correlations which were found from the Hitter/Misser correlations were both weak ($p < 0.06$) and in unexpected directions. It was concluded that they were most probably artifacts from the analysis, but that a definitive conclusion would have to await verification from the analysis of subsequent experiments. It was unfortunate that the subject who was apparently an MLP (subject 1) had returned to Germany. Any investigation of MLPs would have to await further experiments. The SSM sources (other than displacement) show no obvious pattern which could indicate that anyone of them was better than any other. However given the overall SSM score this would be expected.

7.5.3. **Questionnaires.**

7.5.3.1. **TAQ**

It was interesting that the subjects early influences made such an impact upon their later technology based work styles. All of the major indexes of performance, session length, commands, and errors correlated with being allowed to use dangerous equipment when young. It would seem to be an important responsibility for parents to ensure that young children have adequate encouragement to explore technology. The FAA/TAQ correlations were weak, apart from wanting to be good at whatever one did, which was significant ($p < 0.05$), and in the expected direction. However the author suspects that these correlations (including the significant one), are probably noise from the analysis. As with the parapsychological FLP/MLP correlations these kinds of judgments have to await further empirical evidence from subsequent experiments.

The TAQ correlations provide evidence for the role of early experiences in later technology acceptance. This acceptance is most likely to be responsible for the amount of perceived success subjects felt that they had with machines. Subjects who had a negative attitude
towards computers were more likely to report more frequent machine breakdowns. However it is possible that when these subjects have to cope with machine breakdowns they will tend to cope less well. In this state the scale of the breakdown becomes more dramatic than it would appear for someone with a positive attitude towards technology. The author's informal interviews with people with very positive attitudes towards new technology have shown that they have as many (if not more\textsuperscript{6}) breakdowns with equipment than the MLPs, but that the breakdowns do not assume any significance. The subjects concerned can usually cope with the problems without resorting to the advice of another individual. The fact that the MLP has to resort to getting another individual, rather than solving the problem themselves, does two things. It reinforces their self image as being 'no good with machines', and creates a 'significant event' which is more likely to remain in their, and others memory. To break this vicious circle the MLP should get the other individual to talk them through what they are doing when they solve the problem. The author has noted that within informal groups of computer 'Hackers' there is taboo against peers within the group asking one another to do something 'for them'. It is however acceptable for newcomers to the group to ask such favors. The author (along with his supervisors) believe there is a great deal of useful psychological and sociological data which could be gained from investigating and comparing the norms of both 'high-tech' and 'low-tech' social groups.

7.5.3.2. Command Line Interface

The large sex difference bias in the group (nearly 5 female to 1 male), makes any sex related findings suspect. However the findings which did emerge were interesting because the females rated the command line system as too fast. This could be taken as very weak evidence of an inappropriateness of the command line system to a 'female cognitive style'. Although the author is not convinced that such a difference actually exists in any form which would show as strongly as this finding\textsuperscript{7}. The importance of finding the keyboard easy on a command line system was highlighted as was the importance of the choice of command names, as suggested by the work of Barnard et al (1982). The display layout, feedback, and

\textsuperscript{6} A factor which emerged was that FLPs enjoyed 'tinkering' with equipment, and unless they are in a hurry may even enjoy repairing faults.

\textsuperscript{7} The sex differences from the combined data are more revealing since they reflect a more equal gender ratio sample.
error correction findings were not surprising, and merely served to confirm the guidelines of good interface design suggested by HCI researchers.

7.5.3.3. GEM Interface

The ease of using and coordinating the mouse were found to be an important factor with regard to the ease with which people operated the Gem system. It may be that 'clumsy' people (who have poor hand/eye coordination), are less likely to prefer Wimp environments, and might be more likely to be MLPs. It was this suggestion from the data which led to the introduction of some kinesthetic measures in the final version of the TAQ (TAQ3), which is discussed in later chapters.
8. Experiment 2

Quote:

A man who has exercised his mind to the utmost, knows his nature. Knowing his nature, he knows heaven.
Mencius.

8.1. Introduction.

This experiment ran from the 20th of September 1988 to the 11th of October 1988.

8.2. Method, Design And Planned Analysis.

These were as outlined in the initial section.

8.2.1. Subjects

These were 26 first year postgraduates in a systems analysis and design course. The tutorial sessions formed part of their coursework, and attendance was compulsory. The subjects had both typing and previous computer experience\(^1\). Of the 26 who took part, 21 attended both sessions, 25 used the command line system, and 21 used the Wimp system. 14 of the group were male and 12 were female. 24 of the group were right handed, 1 was left handed, and 1 claimed to be ambidextrous. The youngest member of the group was 19 and the oldest was 36.

This was an 'intermediate' group of subjects, a third of whom had done some touch typing and 4/5ths of whom had some previous experience of using computers at various levels of competence.

\(^1\) All of these subjects had used a computer before.

As outlined in the initial section (chapter 6).

Machine positions are given in the diagram.

8.4. Procedure.

There were 5 one hour sessions, proceeded by a one hour lecture on human computer interaction given by S.G. In this lecture the subjects were informed on the background to HCI, and its problems. This lead to a description of the nature of the study (that it was a HCI interface comparison), and that they were taking part to experience the two versions of the system, and give their comments. This lecture also covered how to fill in the questionnaires, and use the mouse and pull down menus, for those subjects who had not used a Wimp system before. Since this subject group were starting a systems analysis and design course they were also asked to view the interfaces from a systems design point of view. This systems analysis task formed part of their course work, but since this was the students induction week it was not a vital part of their grading.
8.5. Results

8.5.1. In-Session Details.

8.5.1.1. Experiment 2 Command Interface Session Details, Parametric Correlation ($n = 25$).

Longer sessions were correlated with more errors ($p < 0.072$), and issuing more commands ($p < 0.05$). Making more errors was correlated with issuing more commands ($p < 0.05$), making errors more quickly ($p < 0.01$), and issuing commands more quickly ($p < 0.05$). Issuing more commands was found to be correlated with attempting more trials ($p < 0.01$), and issuing commands more quickly ($p < 0.01$). A faster error rate was found to be correlated with a fast command rate ($p < 0.05$).

8.5.1.2. Experiment 2 Gem Interface Session Details, Parametric Correlation ($n = 22$).

A longer session length was found to be correlated with a higher number of errors ($p < 0.073$), issuing more commands ($p < 0.01$), and attempting more trials ($p < 0.01$). Making more errors was found to be correlated with making errors more quickly ($p < 0.01$). Issuing more commands was found to be correlated with attempting more trials ($p < 0.01$), and issuing commands more quickly ($p < 0.01$).

There were some unusual correlations within the GEM session data, such that the subject number (order the Ss turned up to the experiment, see chapter 6 for details) correlated with the session length ($p < 0.05$). What is possibly more interesting is that the session length correlated with the psi component within the skeptical software monitor ($p < 0.01$). These correlations were such that earlier subjects had longer sessions, and degraded the performance of the SSM ($p < 0.01$). From this it would seem that it is a possibility that later subjects did better on the covert psi task, however enhancing the SSM was not found to be correlated with the subject number (as would be expected) ($p < 0.5$), so the finding would appear to be the result of a statistical artifact.
8.5.1.3. Experiment 2 Overall Session Details, Parametric Correlations (N = 26).

Longer sessions were correlated with making more errors (p < 0.89), issuing more commands (p < 0.01), and attempting more trials (p < 0.05) Making more errors was found to be correlated a faster error rate (p < 0.01). Issuing more commands was found to be correlated with attempting more trials (p < 0.01), and issuing commands more quickly (p < 0.01).

8.5.2. HCI Hypotheses.

1) WIMP interfaces would be preferred by naive users, but would have a slower throughput of work and a longer overall session time.

In these studies 'work' has been defined as the number of commands issued and mean time between those commands (rate of throughput).

<table>
<thead>
<tr>
<th></th>
<th>Command Line</th>
<th>Wimp Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Number of Commands</td>
<td>29.88</td>
<td>27.68</td>
</tr>
<tr>
<td>Mean Time Between Commands</td>
<td>98.938 secs</td>
<td>90.90 secs</td>
</tr>
<tr>
<td>Mean Session Length</td>
<td>2537.64 secs</td>
<td>2105.31 secs</td>
</tr>
</tbody>
</table>

The difference between session lengths is significant at (p < 0.05) level (using parametric single-t test).

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>p</th>
<th>n</th>
<th>E(X-Y)</th>
<th>D(X-Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session Lengths</td>
<td>2.742</td>
<td>&lt;&lt; 0.05</td>
<td>21</td>
<td>409.76</td>
<td>668.27</td>
</tr>
</tbody>
</table>

However (as in the previous experiments which used the KMDB system), we have to consider the difference in entry to the lowest levels of the help system before we can test the difference in work throughput.
See figure average number of command types experiment 2. This shows that there was a large difference between the number of help commands issued on the two interfaces. However we can adjust for this by excluding the number of help commands.

Table Showing Number of Commands

<table>
<thead>
<tr>
<th>Command Type</th>
<th>Command Line</th>
<th>Wimp Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select</td>
<td>433</td>
<td>447</td>
</tr>
<tr>
<td>Help</td>
<td>119</td>
<td>28</td>
</tr>
<tr>
<td>Modify</td>
<td>30</td>
<td>22</td>
</tr>
<tr>
<td>Add</td>
<td>68</td>
<td>66</td>
</tr>
<tr>
<td>Delete</td>
<td>16</td>
<td>26</td>
</tr>
<tr>
<td>Exit</td>
<td>25</td>
<td>16</td>
</tr>
<tr>
<td>Quit</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Save</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td><strong>Total (- number of helps)</strong></td>
<td><strong>(695 - 119) = 576</strong></td>
<td><strong>(611 - 28) = 583</strong></td>
</tr>
</tbody>
</table>

To calculate the new adjusted average number of commands, and mean time between commands we can do the following:
Command line interface.

Average number of commands = 23.04 commands.
Mean time between commands = 110.14 secs.

Wimp interface

Average number of commands = 26.5 commands.
Mean time between commands = 79.446 secs.

From this we can see that the Wimp interface has a greater throughput of work (110.1 seconds with the command as compared to 79.4 seconds with the Gem) with this subject group (p < 0.05), and the command line system has a longer session length (2537.64 seconds with the command as compared to 2105.31 seconds with the Gem), (p < 0.05).

This diagram shows the mean time between each command on the two interfaces. There seems to be a trend for the wimp inter-command times to decrease as the session progresses, while the command line system data shows a leveling and then a dramatic increase at the later part of the session. The initial peaks are where the subjects are reading and thinking about the tutorial sheets. The patterns of both interfaces then settle down to a more or less steady level, except that in the GEM interface’s case the trend is always to a decreasing inter-command time. The command line inter command times shows a similar trend to that shown in experiment 1.

Figure 99. Time Between Commands, Experiment 2.
Hypothesis 2

Once competence is reached in the command line interface, it will be preferred by experienced users and naive users who are ‘keyboard literate’, due to the interface’s greater information interchange rate, and the familiarity with the keyboard.

The subjects in this group produced a very varied response to the computer system. Some of the group had a definite ‘hacker’ mentality and spent any unobserved moment trying to work out how to break the system. For example subject 4, session 3 told the experimenter that the system should have been written in BASIC so he could get at the workings. In contrast subject 15, session 5, was so afraid of the system that she sat and just looked at it until coaxed by the experimenter to try some of the tutorial exercises.

Subjects had to complete an essay in which they were supposed to describe their feelings toward the two interfaces. We did not bother to analyze these, due to our experience with the essays in experiment one. Instead we relied upon the comparison of the last 5 questions of the post session questionnaires (Com21a - 21e & Gem 23a - 23e), and used the non-parametric alternative to the T test called the Wilcoxon matched pairs test. This revealed that subjects did not rate either interface as being better (N= 21, T= 69.00, Z= 1.62, P 0.11), or more powerful (N= 11, T= 20.50, Z= 1.11, P 0.3) and provided only marginal evidence of the GEM interface being more satisfying (N= 21, T= 68.00, Z= 1.65, P 0.095). However subjects did rate the GEM interface as being significantly more stimulating (N= 21, T= 53.00, Z= 2.17, P < 0.05), and easier to use (N= 21, T= 42.00, Z= 2.55, P < 0.05).

Hypothesis 2 part ii

These findings are most interesting when they are compared with those from experiment 1. These later findings are much restrained and probably reflect the greater experience of this subject group with computer systems. The subject found the Wimp system to be slightly more satisfying, significantly more stimulating, and significantly easier to use. It is interesting that they did not think the Wimp system was better or more powerful as the naive subjects in experiment 1 reported.
Hypothesis 3

In all users, over both interfaces, the error rates will be greater in the Command line system.

If we look at the recorded data for errors we find the following:

<table>
<thead>
<tr>
<th>Number of Errors</th>
<th>Mean</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>St.Dev</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Command line</td>
<td>36.44</td>
<td>25</td>
<td>9.0</td>
<td>90.00</td>
<td>18.62</td>
<td>911.00</td>
</tr>
<tr>
<td>Wimp Interface</td>
<td>15.04</td>
<td>22</td>
<td>3.0</td>
<td>54.0</td>
<td>11.76</td>
<td>331.00</td>
</tr>
</tbody>
</table>

Mean Time Between Errors.

<table>
<thead>
<tr>
<th>Mean</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>St.Dev</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Command line</td>
<td>86.29</td>
<td>25</td>
<td>28.067</td>
<td>43.122</td>
<td>2157.22</td>
</tr>
<tr>
<td>Wimp Interface</td>
<td>230.432</td>
<td>22</td>
<td>17.574</td>
<td>168.06</td>
<td>5069.50</td>
</tr>
</tbody>
</table>

The Single-t tests for these are as follows:

<table>
<thead>
<tr>
<th>t</th>
<th>p</th>
<th>N</th>
<th>E(X-Y)</th>
<th>D(X-Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Errors 4.809</td>
<td>&lt;&lt; 0.001</td>
<td>21</td>
<td>20.905</td>
<td>19.442</td>
</tr>
<tr>
<td>Mean Time Between Errors -3.88</td>
<td>&lt;&lt; 0.01</td>
<td>21</td>
<td>-134.263</td>
<td>154.745</td>
</tr>
</tbody>
</table>
This clearly shows that a highly significant difference exists between the number of errors (and the mean time between errors). The Wimp interface having a highly significant advantage in both cases. This confirms experimental hypothesis number 3.

Diagram time between errors.

This shows a very similar pattern to that found in experiment 1, with the Gem system having a larger time between each error once the subjects had settled into the experiment.

Totals of error types from experiment 2.

The following table shows the error categories after they have been corrected by an analysis which only counts those errors that are possible in both systems. The details of these processes are given in the previous chapters, and in much greater depth in the system technical reference data.

The adjusted totals are:

<table>
<thead>
<tr>
<th>(Adjusted)</th>
<th>Command Line</th>
<th>Wimp Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntax Errors</td>
<td>21.32</td>
<td>11.72</td>
</tr>
<tr>
<td>Semantic Errors</td>
<td>18.6</td>
<td>7.63</td>
</tr>
<tr>
<td>System Errors</td>
<td>2.08</td>
<td>2.0</td>
</tr>
</tbody>
</table>

The adjustment process has reduced the difference between the two interfaces, but the GEM interface is still only has 54.9% of the syntax errors, and 41% of the semantic errors present in the command line system. These differences are most clearly seen in the pie charts of error types. The differences strongly resemble those found on the first experiment. This similarity between the errors could be taken as reinforcing the proposed effect of the mental maps provided by the GEM system (as originally proposed in experiment 1).
Figure 101. Wimp Error Types, Experiment 2.

Figure 102. Command Line Errors, Experiment 2.

Figure 103. Overall Error Types, Experiment 2.

Types of Errors (adjusted) Exp2

Wimp Interface

Syntax 11.72

Command Interface

Syntax 21.3

System 2.0

Semantic 18.6

Types of Errors (adjusted) Exp2

Combined

Com Syntax 21.32

Com Semantic 18.6

Gem System 2.0

Gem Semantic 2.0

Gem Syntax 7.63

11.72
8.5.3. Review of presentation effects.

Since a counter balanced within subjects design was being used the data can be split into those who experienced the command system, then the GEM, and vice versa. This is shown in the following table.

<table>
<thead>
<tr>
<th></th>
<th>First Session</th>
<th>Second Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>Command Line Session</td>
<td>2427.0</td>
<td>2657.5</td>
</tr>
<tr>
<td>Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Command Line Number</td>
<td>31.84</td>
<td>41.41</td>
</tr>
<tr>
<td>of Errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Command Line Number</td>
<td>27.46</td>
<td>32.5</td>
</tr>
<tr>
<td>of Commands</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wimp Interface Session</td>
<td>1893.46</td>
<td>2411.33</td>
</tr>
<tr>
<td>Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wimp Interface Number</td>
<td>15.23</td>
<td>14.77</td>
</tr>
<tr>
<td>of Errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wimp Interface Number</td>
<td>19.53</td>
<td>39.44</td>
</tr>
<tr>
<td>of Commands</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Using a Kruskal Wallis anova by ranks on this data significant presentation differences were found between the command line (p < 0.05), and GEM (p < 0.01) session lengths. The number of commands in the GEM sessions (p < 0.01), and the mean time between commands in the GEM sessions (p < 0.01). These findings were such that subjects had longer sessions in their second use of the system, on both interfaces. While the other correlations are restricted to the Gem systems, so that subjects who experienced the Gem system after the command line system issued more commands, and issued them faster. These can be explained by simple learning effects.

8.5.4. Attempted replication of the findings on colour screens from experiment 1.

One of the interesting findings from the first experiment had been the post hoc discovery that first time WIMP system users found colour screens suggestively more stimulating (p < 0.06), and issued more commands in WIMP systems when the used colour screens (p < 0.05). However none of the findings from the first experiment about colour screen were
replicated. This could either be due to the subject group being more computer experienced, or simply due to a statistic artifact.

8.6. Parapsychological Findings.

The main SSM measure revealed no significant psi effects, with 355 hits in 696 trials \((Z=+0.53)\).

![Psi Trials Graph](image)

**8.6.1. Psi Trials.**

The graph of psi trials shows little pattern, except perhaps that the command line data shows a much less jagged profile.

**8.6.2. FLP/MLPs (Empirically Determined Via Their SSM Measure).**

Looking at the subjects we use the same method as used in experiment 1 to determine 'hitters' and 'missers'. If a subject scored consistently in the same direction (hitting/missing) over both interfaces then they are labelled as being a hitter or misser. This allows us to do a multi-way non parametric anova (Kruskal Wallis) on the in-session recorded data. The only two variables that show any significance are the psi-hitting missing variables (Command line hit rate \((p<0.05)\), and the GEM hit rate \((p<0.05)\)), used to make the original categories.

**8.6.3. SSM Decision Sources.**

The graphical representation of the decision making systems shows that the sources were predominately hitting. In experiment 1 they had been predominately negative (missing),
Experiment 2

SSM Sources

Exact Missing
Displacement
Reg. Wave
Quick Coin
Quick Stab
Unigram
Coin

Success

Command
Gem
Control

Figure 107. SSM Sources, Experiment 2.

except displacement. In the current experiment none of the sources reached significance. It is highly probable that this change from missing to hitting is merely a random fluctuation. However further experimentation is necessary to clarify this issue. The displacement source is again found to be the most successful of the sources, but investigation showed many repeated commands in both interfaces, and this would produce the so called displacement artifact (see appendix 13).

8.7. Questionnaires.

Based upon the results which were found in experiment one's data the questionnaire data was analysed looking for similar effects; in the TAQ, Sex differences, Working style, Computer Attitudes, Use of machines, Early life, and Function linkage (both FLP & MLP); in both versions of the UEICS2, Sex differences, Previous experience with WIMP systems, Error correction, Feedback, Error Correction &
Prevention, Display layout, Use of the keyboard, and Command Name Consistency. All analyses were done using a non-parametric Spearman rank order correlation.

8.7.1. In Session Recording Correlations With The TAQ2 (n = 26).

Early Parental Influence.

Subjects whose parents allowed them to help repair broken equipment had shorter sessions (p < 0.05). Shorter sessions were also associated with those subjects whose parents had allowed them to use dangerous equipment (p < 0.05), these subjects also issued fewer commands (p < 0.05).

Gender Differences.

There were some weak gender differences, all in favor of the female subjects. Females issued more commands (p < 0.086), more trials (p < 0.05), and issued their commands more quickly (p < 0.05).

Working Style.

Those subjects who enjoyed following instructions issued more commands (p < 0.05), and those who enjoyed figuring out how things work tended to have longer sessions (p < 0.072).

Attitudes Towards Technology.

Shorter session lengths were associated with subjects who felt that computers were not too complicated to understand (p < 0.05), and who liked to repair machines (p < 0.05). Subjects who like to work with machines issued their commands more slowly (p < 0.053), while those subjects who were unafraid of looking silly if they made a mistake issued their commands more quickly (p < 0.05). Subjects who preferred to use computers themselves made fewer errors (p < 0.059), and made errors more slowly (p < 0.05). While in contrast subjects who were not afraid of damaging the computer if they used it seemed to make errors more quickly (p < 0.064).
Use of Technology.

Subjects who frequently used digital watches had shorter sessions (p < 0.05), seemed to issue fewer commands (p < 0.081), and seemed to be associated with faster error rates (p < 0.082). Subjects who frequently played video games had shorter session lengths (p < 0.05), issued fewer commands (p < 0.06), and made errors more quickly (p < 0.05). Subjects who used a micro at college had shorter session lengths (p < 0.05), and issued fewer commands (p < 0.06). Subjects who used the mainframe at college had shorter session lengths (p < 0.01), and issued fewer commands (p < 0.01).

Function Linkage.

Enhanced SSM performance was found to be correlated with enjoying figuring out how things work (p < 0.05), disagreeing that computers were too complicated to understand (p < 0.05), frequent use of a programmable calculator (p < 0.06), frequent reading about computers (p < 0.077), and being mildly injured by a computer (p < 0.05).

8.7.2. Inter-Question Correlations From Within The TAQ2 (n = 26).

Age.

Perhaps surprisingly older subjects (oldest in the group was 36) viewed the computers as being more fun to use than younger subjects (p < 0.051). Older subjects used digital watches less (p < 0.05), played fewer video games (p < 0.05), and used the college microcomputers less frequently (p < 0.01). Older subjects seem to have been less likely to have helped their parents repair equipment (p < 0.09), to have received less encouragement to explore how machines worked (p < 0.05), and to have been significantly less likely to have had parents who repaired equipment themselves (p < 0.01).

Gender Differences.

The only gender difference from within the TAQ2 was that females tended to feel that computers were more dependable than human beings (p < 0.05).
Working Style.

Subjects who enjoyed figuring out how things worked also enjoyed working with (p < 0.05) and repairing (p < 0.01) machines. The enjoyment of figuring out how things work also correlated with not feeling that computers were too hard to understand (p < 0.05), and preferring to use the computer themselves, rather than letting others use it for them (p < 0.01). Subjects who enjoyed figuring out how things worked felt that computers were fun to use (p < 0.05) and that they helped them to work better (p < 0.01). Enjoyment in figuring out how things work also correlated with frequent playing of video games (p < 0.05), more frequent use of microcomputers at college (p < 0.05), and reading of books about computers (p < 0.05). Subjects who liked following instructions also liked working with machines (p < 0.01), but did not frequently use the college mainframe (p < 0.5). Subjects who liked working with machines also liked to repair them (p < 0.01), were not bothered by machines impersonal nature (p < 0.052), felt that computers were fun to use (p < 0.01), and helped them work better (p < 0.05).

Computer Attitudes.

Subjects who felt that they could understand computers were unafraid of damaging the computer if they used it (p < 0.05), or of looking silly if they made a mistake while they used a computer (p < 0.5). Feeling that one could understand computers also correlated with feeling that computers were fun to use (p < 0.01). Subjects who liked to repair machines used a mainframe computer more frequently (p < 0.05), and read more about computers (p < 0.01). Frequent playing of video games was found to be weakly correlated with being unafraid of damaging the computer (p < 0.066).

8.7.2.1. Function Linkage.

Remembering that function linkage does not assume a psychic component, 3 main factors emerged as having importance in function linkage, current use of equipment, current attitudes towards equipment, and parental influence.
Current Use Of Equipment.

Subjects who used multi-functional digital watches had fewer problems with their photographic equipment (p < 0.05), and subjects who used auto-bank tellers had fewer problems with HiFi equipment (p < 0.05). However, subjects who frequently played video games tended to have more frequent malfunctions with their personal electrical appliances (p < 0.06), and television (p < 0.05). This could be due to the somewhat aggressive attitude video games could induce in the subject's handling of the equipment.

Current Attitudes Towards Technology.

Subjects who enjoyed hobbies that involve doing things in a specific order had fewer breakdowns with their TV & Video equipment (p < 0.05), and seemed to have fewer problems with their HiFi (p < 0.073). Subjects who preferred to use a computer themselves reported fewer breakdowns with mechanical equipment (p < 0.05). Subjects who reported being unafraid of damaging the computer if they used it had fewer breakdowns with their personal electrical appliances (p < 0.01). In a similar way subjects who reported feeling that computers helped them work better had fewer mechanical breakdowns (p < 0.05).

Parental Influences.

The parental influence upon attitudes in later life were found to be very strong. Subjects whose parents had encouraged them to explore how machines worked developed a whole series of technology positive traits in later life. These subjects liked to work with machines (p < 0.05), frequently used microcomputers (p < 0.001), and mainframes (p < 0.01). Finally, they also reported being unafraid of damaging the computer if they used it (p < 0.05). Subjects whose parents had allowed them to help repair equipment developed a similar series of characteristics. These subjects played more video games (p < 0.01), used microcomputers and mainframes more frequently (p < 0.01, and p < 0.05 respectively), and developed an immunity to fear of damaging the computer if they used it (p < 0.05). This trend, of more frequent computer (micro and mainframe) use, was shared by subjects who had been allowed to use dangerous equipment when they were children (p < 0.05) for micro, and (p < 0.01) for mainframe use respectively. Subjects who had been allowed to use dangerous equipment read more about computers (p < 0.05), and were unafraid of
damaging the computer when they used it (p < 0.05). Subjects whose parents had ensured that broken equipment was promptly repaired reported feeling that computer could be more dependable than human beings (p < 0.05). While subjects who had seldom being frightened by equipment when they where children were less bothered by the impersonal nature of machines (p < 0.05).

Subjects who had been allowed to use dangerous equipment when they were children felt that they could understand computers (p < 0.05), and seemed to have fewer breakdown with their TVs (p < 0.075), HiFi (p < 0.062), or personal electronic appliances (p < 0.05). There was also some evidence that function linkage seems to persist over various items of equipment. Subjects whose personal electrical appliances rarely broke down also rarely had problems with photographic equipment (p < 0.05), Tv & Video (p < 0.01), or HiFi (p < 0.05). In contrast to our expectations it was found that subjects who reported frequent break downs in their childhood home had few breakdowns with mechanical devices in later life (p < 0.01), but this could be a statistical artifact due to the number of analysis.

**Use of New Technology.**

Subject who used a multi-function digital watch were more likely to use a microcomputer at college (p < 0.05). However this use of microcomputers both at home and at the college correlated with the more frequent playing of video games (p < 0.01). This is confirmed by the finding that subjects who frequently played video games also read computer based books beyond those required for their course (p < 0.059). From these correlations it would seem that playing of video games was one of the major factors in the use of computers in this subject group.

**Summary Discussion On A Possible Subject Number Effect.**

We found some weak (non-significant) effects which correlated with the subject number, so later subjects disliked working with machines more (p < 0.06), and were afraid of looking silly if they were seen by others not to know how to use the machine. These are similar to the weak 'psi' effects which were found in the recorded behavioral data from the systems, and may be a sign of some weak FLP/MLP effects. Weak evidence for this is that the subject number correlated with the use subjects made of terminals and the college mainframe (p < 0.055).
8.7.3. UEICS2 Post Session Command Line Questionnaires (n= 25).

8.7.3.1. Correlations Within The Command Line In-Session Recordings.

Sessions and Commands.

Subjects who had used a command interface before had longer sessions (p < 0.01), and issued more commands (p < 0.05). Subjects who found that on-line help was clear issued also more commands (p < 0.01), and issued them more quickly (p < 0.01).

Errors.

Subjects who found the keyboard easy felt that error messages were always helpful (p < 0.05), however they also made errors more rapidly (p < 0.09). Subjects who found that error messages were helpful made fewer errors (p < 0.09), and had longer time between commands (p < 0.05). While subjects who found that error correction was clear had fewer errors (p < 0.05), and had a longer mean inter error time (p < 0.05). Subjects whose overall reactions to the system were good had a longer inter-error time (p < 0.082). Subjects who found that the system was stimulating made fewer errors (p < 0.05), and had a longer mean inter error time (p < 0.05).

Function Linkage.

Enhancing SSM performance was found to be correlated with greater age (p < 0.075), and rating that the information needed to complete each task had to be memorized rather than being present on the screen (p < 0.065).

8.7.3.2. Inter-Question Correlations From Within The Command Line Questionnaire.

Age

Age correlated with the feeling that error messages indicated the corrective action to be taken (p < 0.054), such that older subjects thought the error messages were better.
Sex Differences,

Males seemed to find the task instructions easier to understand \( (p < 0.066) \), that the system prevented mistakes \( (p < 0.052) \), was more stimulating \( (p < 0.05) \), and easier to use \( (p < 0.085) \).

Previous Experience With Command Line Systems,

Subjects who had used a command interface before found that the number of actions per task too large \( (p < 0.088) \), and that the system was too slow \( (p < 0.05) \). These same subjects rated the error messages as being more helpful \( (p < 0.052) \).

Feedback,

Subjects who reported that feedback was sufficient felt the way the system displayed its information simplified work \( (p < 0.01) \), and that on-line help was clear \( (p < 0.054) \). These subjects rated the system as being easy to learn \( (p < 0.068) \), satisfying to use \( (p < 0.08) \), and as having adequate power \( (p < 0.061) \).

Error Correction & Prevention,

Subjects who felt that the system prevented mistakes felt feedback was sufficient \( (p < 0.054) \), error messages were helpful \( (p < 0.055) \), that error messages indicated the corrective action to be taken \( (p < 0.05) \), and that on-line help seemed to be helpful \( (p < 0.08) \). Error prevention was found to correlate with rating the system as being good \( (p < 0.05) \), satisfying \( (p < 0.05) \), easy to learn \( (p < 0.01) \), easy to use \( (p < 0.01) \), and as having adequate power \( (p < 0.05) \). Subjects who felt that error messages indicated the corrective action to be taken also felt that error messages were helpful \( (p < 0.01) \), error correction was clear \( (p < 0.01) \), and that it was easy learning the system \( (p < 0.01) \). Rating the error messages as indicating the corrective action was found to correlate with feeling the system was good \( (p < 0.001) \), satisfying \( (p < 0.01) \), stimulating \( (p < 0.05) \), and easy to use \( (p < 0.01) \).

Subjects who felt that error correction was clear also felt that on-line help was easy \( (p < 0.05) \), and that it was easy to learn the system \( (p < 0.01) \). Clarity of error correction was found to be correlated with feeling that the system was good \( (p < 0.01) \), satisfying \( (p < 0.05) \), and easy to use \( (p < 0.01) \).
Display Layout.

Subjects who felt that the display layout simplified work seemed to feel that error messages were helpful (p < 0.092), that the error messages indicated the corrective action (p < 0.057), that error correction was clear (p < 0.05), and that on-line help was clear (p < 0.05). Good display layouts were found to correlate with rating the system as good (p < 0.069), satisfying (p < 0.017), easy to use (p < 0.05) and easy to learn (p < 0.01). Clarity of the on-line help was found to correlate with feeling the system was good (p < 0.05), satisfying (p < 0.01), easy to use (p < 0.01), and easy to learn (p < 0.01). Subjects who felt that the information needed to complete each task was present felt the system was good (p < 0.056), satisfying (p < 0.05), stimulating (p < 0.05), and easy to use (p < 0.062).

Use Of The Keyboard.

Subjects who found the keyboard was easy to use found that the words used as commands matched the actions performed (p < 0.05), that the system was easy to learn (p < 0.05), use (p < 0.053), and had adequate power (p < 0.05).

Command Name Consistency.

Subjects who felt that the words used as commands matched the actions felt feedback was sufficient (p < 0.055), work was simplified by the display layouts (p < 0.01), on-line help was clear (p < 0.05), learning the system was easy (p < 0.05), and that the system had adequate power (p < 0.05). Subjects who felt there was command consistency rated the tutorial instructions as clear (p < 0.05), the system as preventing mistakes (p < 0.05), feedback as sufficient (p < 0.05), the display layouts as simplifying work (p < 0.05), error messages as helpful (p < 0.05), and on-line help as clear (p < 0.05). Command consistency was found to be correlated with feelings that the system was satisfying (p < 0.01), easy to learn (p < 0.05), and easy to use (p < 0.05).

Clarity Of The Tutorial.

Subjects who found the instructions describing the tasks they had to perform easy to follow felt that the system prevented mistakes (p < 0.05), that feedback was sufficient (p < 0.05), that display layouts simplified work (p < 0.01), that error messages were helpful (p < 0.01),
and indicated the corrective action (p < 0.05). Those subjects who found the tutorial easy also felt that error correction (p < 0.05), and on-line help were clear (p < 0.01). Good ratings for the tutorial were found to be correlated with rating the system as good (p < 0.05), satisfying (p < 0.01), easy to use (p < 0.01), and easy to learn (p < 0.05).

**Inter Questionnaire Reliability.**

Subjects who felt that it was easy to use the system felt the system was good (p < 0.001), satisfying (p < 0.001), and had adequate power (p < 0.05).

**8.7.3.3. EUICS2 - Gem Post Session Questionnaire (N = 22).**

Based upon the results which were found in experiment one’s data the questionnaire data was analysed looking for similar effects. Sex differences, Previous experience with WIMP systems, Error correction, Feedback, Error Correction & Prevention, Display layout, Use of the keyboard, and Command Name Consistency.

**8.7.3.4. Correlations Within The Gem In-Session Recordings.**

Females issued more commands (p < 0.054), and issued them more quickly (p < 0.075). Similarly, subjects who found the mouse and keyboard easy to coordinate issued more commands (p < 0.05), and issued those commands more quickly (p < 0.05). Faster commands were also associated with subjects who rated the system as being easy to learn (p < 0.05).

**8.7.3.5. Inter-Question Correlations From Within The Gem Questionnaire.**

**Age.**

Older subjects felt there were too many actions per operation (p < 0.05), and that error correction was confusing (p < 0.058).
Previous Experience With WIMP Systems.

Subjects who had used a mouse before seemed to have found pull down menus easy ($p < 0.087$), and the use of the keyboard and mouse together easy ($p < 0.078$). However these subjects found the tutorial confusing ($p < 0.05$), and felt there were too many actions per task ($p < 0.096$). Subjects who found the mouse easy to control found it easy to master pull-down-menus ($p < 0.05$), and the simultaneous use of the keyboard and mouse ($p < 0.05$). These subjects also seemed to have found that the words used to describe tasks matched the operations closely ($p < 0.082$), that the system always prevented mistakes ($p < 0.071$), and that feedback was always enough ($p < 0.08$). These subjects rated the system as being good ($p < 0.05$), and stimulating ($p < 0.08$) to use. Subjects who found pull down menus very easy found that words matched command action ($p < 0.01$), there were too many actions per task ($p < 0.05$), that the work was always simplified by the display layout ($p < 0.05$), and that error messages were always helpful ($p < 0.054$).

Feedback.

Subjects who felt that feedback was sufficient also felt that display layouts simplified work ($p < 0.01$), that error messages were always helpful ($p < 0.05$), and that error messages indicated the corrective actions to be taken ($p < 0.05$). Subjects felt that systems with good feedback were easy to learn ($p < 0.05$), good ($p < 0.05$), satisfying ($p < 0.065$), and stimulating to use ($p < 0.07$).

Error Correction & Prevention.

Those subjects that felt the system prevented mistakes felt that feedback was sufficient ($p < 0.05$), display layouts simplified tasks ($p < 0.05$), error messages were helpful ($p < 0.05$), and that error messages indicated the corrective actions to be taken ($p < 0.05$). Systems which prevented mistakes were rated as having clear error correction ($p < 0.053$), and were easy to learn how to use ($p < 0.056$). Subjects who felt the system prevented mistakes felt the system was good ($p < 0.05$), and satisfying ($p < 0.078$). Subjects who found that error messages helpful reported that they indicated the corrective action to be taken ($p < 0.01$), that error correction was clear ($p < 0.05$), and that this made the system easy to learn ($p < 0.05$). Subjects who found error correction was clear found the system was easy to learn ($p
< 0.01), and good to use (p < 0.05). Subjects who found the system prevented mistakes thought the system had adequate power (p < 0.066), as did those subjects who thought that error messages were helpful (p < 0.05).

**Display Layout.**

Subjects who felt that display layout simplified work found the on_line help was clear (p < 0.01), the system easy to learn (p < 0.01), and that the information needed to complete each task accessible on the screen (p < 0.01). Good display layout correlated with finding the system was good (p < 0.01), easy (p < 0.05), satisfying (p < 0.01), and stimulating to use (p < 0.05).

Subjects who found that the information to complete each task was accessible on the screen thought the system was good (p < 0.059), satisfying (p < 0.01), easy (p < 0.088), and stimulating (p < 0.05). These subjects also felt that on-line help was clear (p < 0.05).

**Command Name Consistency.**

Those subjects who rated the words used for commands as closely matching their tasks found the instructions describing the tasks they had to perform clear (p < 0.069), the system always prevented mistakes (p < 0.05), and that feedback was always sufficient (p < 0.01). Command/action match was found to correlate with the display layout simplifying work (p < 0.05), the system being easy to learn (p < 0.05), the information required to complete each task being present on the screen (p < 0.05) and the system being good to use (p < 0.094).

Subjects who rated the system as having command consistency seemed to have found feedback was sufficient (p < 0.065), that the display layout simplified work (p < 0.05), that the information required to complete each task was accessible on the screen (p < 0.05). Command consistency seemed to make the subjects describe the system as good (p < 0.067), satisfying (p < 0.05), and stimulating to use (p < 0.092).

**Possible Subject Number Effects.**

The subject number seemed to be correlated with how closely subjects rated the command/action match (p < 0.05) (so later subjects felt they were better), and negatively with the amount of feedback (so higher subject numbers felt feedback was worse) (p < 0.05).
Clarity Of The Tutorial.

Subjects who found that the instructions clear found feedback was always sufficient \((p < 0.05)\), display layout simplified work \((p < 0.05)\), error correction was clear \((p < 0.05)\). Clear instructions lead to subjects rating the system as easy \((p < 0.01)\), good \((p < 0.01)\), and satisfying \((p < 0.05)\).

Ease of Learning.

Subject who found the system was easy to learn thought the system was good \((p < 0.01)\), and satisfying to use \((p < 0.05)\).

Questionnaire Reliability.

Subjects who found the system good also found it satisfying \((p < 0.001)\), and stimulating \((p < 0.01)\). Those that found the system was easy to learn rated the system was easy \((p < 0.01)\). Ease of use and adequate power also correlated \((p < 0.05)\).

8.8. Discussion.

8.8.1. HCI

The in session records appear to confirm that the recorded process were operating correctly. The correlation which was found within the Gem data with regard to session numbers and psi influence would seem to have an artifact due to the number of analysis performed on the data set (since the expected cross correlation between subject number and psi did not appear). This subject group performed more work in the Gem interface, and had longer sessions in the command line version. A similar trend to that found in the first experiment was found in the inter-command time data, such that the Gem system shows the times decreasing to plateau. In contrast the command line data shows an increase towards the end of the session. These more computer experienced subjects rated the Gem interface differently to those in the previous experiment. Reporting no significant difference in how good or powerful they rated the interface. However they did find the Gem interface more stimulating, and easier to use. As in the first experiment the largest differences between the two interfaces were found in the number and rate of errors made in the two systems. The Gem interface was
found to be significantly better in all error measures. The ratio of semantic to syntactic error types was similar to that found in the previous subject group. Presentation effects showed some significant variations due to simple learning effects, these were present in both interfaces. There was also evidence of more exploratory behaviour in the second session. The attempt to replicate the screen colour differences was unsuccessful, this was a shame since the author felt the post-hoc finding from the experiment one had made logical sense. In conclusion this second experiment which used experienced computer users with keyboard skills strongly suggests that the WIMP interface is faster, easier, less error prone, and more stimulating to use, even for the sophisticated computer user. Whether this trend would continue given prolonged use of the interface, remains to be determined.

8.8.2. Parapsychological

The overall SSM measure showed no evidence of any paranormal SSM FAA. Although both systems were marginally above chance this was most probably due to the effects of the displacement artifact (see SSM Sources below). The author was disappointed that this group did not show significant enhancement of the SSM FAA, since they were the 'expert' user population and therefore theoretically more likely to have FLP tendencies. There was no evidence from the analyses for the existence of either FLPs or MLPs within this subject group, in either the TAQ and session data. The SSM sources show no pattern except for the displacement artifact. In conclusion the second experiment revealed little evidence for computer function enhancement. The very weak data which has emerged over the past two experiments showed the opposite trend to that expected. Subjects who we would have expected to be linked with function enhancement seem to be degrading the performance of the system. It must be noted that this effect is so very weak it could well be due to chance, or some unidentified artifact in system. Although the system was extensively tested, and the recorded data has proved to have a very high degree of internal consistency, it is still possible that some artifacts remain within the 71,000 lines of code.

8.8.3. Questionnaires.

8.8.3.1. TAQ

There was an apparent reversal of the strong effects associated with the early encouragement to explore machines which had been found in the first experiment within the session data.
Subjects in this group who reported having such encouragement issued fewer commands and had shorter sessions, this was not predicted. However the original trend of FLP tendencies being formed by early experiences were confirmed by the later TAQ correlations. The correlations with enhancement of the SSM were all in the predicted direction, for example enjoying solving problems, and having a positive attitude towards machines. The correlation about being hurt by a machine was also found in the previous experiment, and might prove to be due to early use of dangerous equipment. The FLP correlations were much stronger than in the previous experiment, as would be predicted by the groups greater computer expertise. However it is possible that these FLP subjects were trying to understand how the system worked and duplicated their trials more frequently than the other subjects\(^2\). This would cause them to have increased the displacement source’s accuracy. If this was the case it these subjects would be expected to have issued more trials. However no evidence for this emerged from an inspection the session log data. We are left with the possibility that this could be a genuine indication of FLP tendencies enhancing SSM FAA.

It is interesting that females tended to view computers as being more dependable than human beings. This subject group was more balanced in it’s gender ratio (almost 50:50), and we can therefore take more notice of any resulting gender effects. However these females could be un-representative, since they are experienced computer users.

Finally we should note that similar trends appeared for subjects who reported fewer breakdowns as those which we reported in experiment 1. The same factors which we discussed obviously still apply in this group.

**8.8.3.2. Command Line Questionnaire.**

Previous use of a command line interface, acceptance of the keyboard, and appropriate command naming seemed to be the most important factors within the command line system. They were associated with increased productivity and satisfaction. The males seemed to find the command line system easier to use than the females, although these differences did not reach significance. As in the previous experiment the factors which HCI researchers

---

2 Since the knowledge of which commands activated a psi trial was covert this hypothesis would be problematic.
recommend as good interface design were found to enhance both the performance, and satisfaction of the users. The clarity of the tutorial was also found to be a very important factor in the ease of using the interface.

8.8.3.3. GEM Interface.

Females seemed to have performed better on the Gem interface than males. The importance of being able to use and control the mouse was again found to be a major factor in the acceptance of the Wimp interface. There also appeared to be some evidence that later subjects felt the Gem interface was better than earlier ones. It is possible that these later subjects were less experienced computer users.
9. Experiment 3.

Quote:

Most accidents in well designed systems involve two or more events of low probability occurring in the worst possible combination.

Law of Titanic Coincidence, Robert Machol.

9.1. Introduction.

This experiment ran from the 17th of November 1988 to the 3rd of February 1989. The experiment had to be split into 2 sub-series\(^1\), with a break between the 5th of December 1988 and the 31st of January 1989\(^2\).


These were as outlined in the initial section.

9.3. Subjects

The subjects were split between 26 first year HND office studies students, and 75 first year business studies degree students. 25 subjects (split between both groups), completed the experiment. 8 subjects from the HND completed both interfaces, and the remaining 3 were taken from the second group. Of the 25 subjects 21 were female, and 4 were male. They were aged between 17 and 40 approximately\(^3\), the average age being about 18. All but one of the subjects were right handed.

---

1 There were problems getting enough subjects to complete this study.
2 During this forced break the author conducted experiment 4 (see chapter 10).
3 Some of the female subjects declined to provide their personal details.
This experiment proved to be among the most trying aspects of the whole study for the author due to problems with the subjects. Unlike the other problems that had been faced in the study this one could not be solved by hard work, or logic. The subject's attendance of the sessions was purely optional, and most of the subjects opted not to attend. With hindsight this was not surprising, since the subjects, most of whom disliked computers, had been given large amounts of course work, and were in the midst of exam preparation. However the author had to take whatever subject populations he was given, and could not afford to miss an opportunity to run computer and keyboard naive subjects, since they are extremely difficult to obtain. When the subjects did not attend the sessions the author tried to encourage them by (on the group's suggestion) offering to buy the class a drink (which they accepted), or even to throw a party in exchange for their cooperation. Finally (again upon the group's suggestion), the author worked full time at Napier Polytechnic for 2 weeks offering potential subjects an hours individual tuition in any subject\(^4\), in return for their completing the one hour experimental session. Unfortunately even this failed because subjects took the 1 hours tuition, and then failed to turn up for their experimental session. These subjects then 'hid' in the coffee rooms to avoid, as they put it, 'being captured' by the author. The author did not bother 'hunting' for these missing subjects. After an incident where an 'un-captured' subject literally screamed and ran away from the author, rather than pass him on the street\(^5\), it was decided to abandon working with these subjects.

To try to make up the numbers to the minimum required to conduct meaningful non-parametric statistics, a second series of experiments were conducted in the February of the following year (1989). This time the subjects were students in their first year of a degree in business studies (these students were also computer naive, non-touch typists). A lecture was arraigned where the entire intake of 275 students could potentially become subjects. Of these

---

Subjects covered included accounting, computing, business administration, (T-line) shorthand, industrial psychology, wordprocessing, and mathematics.

The author would not have recognised the subject had she not made such a dramatic gesture! Indeed the author was in the process of walking straight past her. Some colleagues have speculated that it was the very fact that the author was ignoring the student concerned which caused this behaviour. The author would not have forced any 'un-captured' subject into attendance.
275 students, 75 attended the lecture, and of these 10 agreed to participate. Only 4\(^6\) of these 10 actually took part in the experiment.


As outlined in the initial section.

9.5. Procedure

13 one hour sessions, split into 2 parts, one before Christmas, the other at the start February. 2 one hour lectures proceeded these two parts. These lectures were provided by SG, and informed subjects on the nature of the study, how to fill in the questionnaires, and use the mouse and pull down menus.

---

One of these 4 students was an exchange student from Hong Kong, and it is uncertain if he understood that attendance was voluntary. The author tried to explain to the subject that he did not have to attend, but it is uncertain if the subject fully understood.
9.6. Results

9.6.1. In Session Recording Correlations.

9.6.1.1. Experiment 3 Command Line Session Details, Parametric Correlation (n = 18).

Longer session lengths were found to be correlated with issuing more commands (p < 0.01), attempting more trials (p < 0.01), and issuing commands more quickly (p < 0.05). More errors were found to be correlated with making errors more quickly (p < 0.01). More commands were found to be correlated with attempting more trials (p < 0.01), and issuing commands more quickly (p < 0.01).

9.6.1.2. Experiment 3 Gem Interface Session Details, Parametric Correlation (n = 18).

Longer session lengths were found to be correlated with issuing more commands (p < 0.01), attempting more trials (p < 0.05), issuing commands more quickly (p < 0.05). More errors were found to be correlated with a faster rate of issuing errors (p < 0.01). More commands were found to be correlated with attempting more trials (p < 0.01), and issuing commands faster (p < 0.01).

9.6.1.3. Experiment 3 Overall Session Details, Parametric Correlation (n = 25).

Longer sessions were found to be correlated with making more errors (p < 0.01), issuing more commands (p < 0.01), attempting more trials (p < 0.001), and issuing commands more quickly (p < 0.05). Making more mistakes was found to be correlated with issuing more commands (p < 0.01), attempting more trials (p < 0.06), and having faster error rates (p < 0.01). More commands were found to be correlated with attempting more trials (p < 0.01), and issuing commands more quickly (p < 0.01).

These correlations can be seen as confirmations that the resident recording systems were functioning normally.
9.6.2. Experimental Hypothesis.

Hypothesis 1 - Wimp Interfaces Will Have A Slower Throughput Of Work.

The command line system showed an average number of commands of 31.333, and an average inter command time of 81.365 seconds. This is in contrast to the Wimp system which shows an average number of commands of 26.056, and an inter command rate of 85.645 seconds. However we have the famous command line 'help' bias, which can be more clearly seen by looking at the 'Average number of commands' graphs.

<table>
<thead>
<tr>
<th>Command Type</th>
<th>Command Line</th>
<th>Wimp Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select</td>
<td>384</td>
<td>361</td>
</tr>
<tr>
<td>Help</td>
<td>85</td>
<td>24</td>
</tr>
<tr>
<td>Modify</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>Add</td>
<td>45</td>
<td>44</td>
</tr>
<tr>
<td>Delete</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Exit</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>Quit</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Save</td>
<td>0.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Totals</td>
<td>565</td>
<td>469</td>
</tr>
</tbody>
</table>

The difference between the use of helps is in fact significant at the (p < 0.01) level.
Single t-test

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>p</th>
<th>n</th>
<th>E(X-Y)</th>
<th>D(X-Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Helps</td>
<td>-3.233</td>
<td>&lt;&lt; 0.01</td>
<td>18</td>
<td>-3.38</td>
<td>4.322</td>
</tr>
</tbody>
</table>

Adjusted for help bias this becomes:

\[
\text{Gem} = 445 \text{ with an N of 18 produces a mean number of commands of } 24.722
\]

\[
\text{Com} = 480 \text{ with an N of 18 produces a mean number of commands of } 26.667
\]

From these we can adjust the mean time between commands to be

\[
\text{Mean Gem Session length} = 82.931
\]

\[
\text{Mean Com Session length} = 84.159
\]

This shows that the users of the command line system issued more commands, although the difference does not reach significance. In contrast the users of the Gem system issued fewer commands, but they issued them more quickly (although again this does not reach any level of significance).

The graph time between commands shows this more clearly. The command line users are initially faster, but the mental map provided by the Gem system enables the Gem users to reduce their average inter command times as the session goes on. This trend runs through the past 2 experiments.

Figure 110. Time Between Commands, Experiment 3.
9.6.2.1. Hypothesis 2

Once competence is reached in the command line interface, it would be preferred by experienced users and naive users who are 'keyboard literate', due to the interface's greater information interchange rate, and the familiarity with the keyboard.

This subject group was the rarest of those investigated, computer and keyboard naive. Since the post experimental essays had been found to provide little useful data (other than how closely the class worked together), no post session essay was required from these subjects. Because of the problems we had in getting these subjects to attend the sessions it would have been futile to have requested an essay anyway. Instead we relied upon the comparison of the last 5 questions of the post session questionnaires (Com21a - 21e & Gem 23a - 23e), and used the non-parametric alternative to the T test called the Wilcoxon matched pairs test. This revealed that subjects did not seem to rate either interface as better\(^7\) (N= 11, T= 20.50, Z= 1.11, P 0.25), more satisfying (N= 11, T= 21.5, Z= 1.02, P 0.3), more stimulating (N= 11, T= 15.00, Z= 1.60, P 0.1), easier (N= 11, T= 26.5, Z= 0.57, P 0.5), or more powerful (N= 8, T= 11.0, Z= 0.98, P 0.3). It is very interesting to notice that this group had less preference than the previous groups towards either system. The author speculated that this group may have been actively avoiding computers all their lives and therefore dislike them intensely. It was hoped that the correlations from the TAQ2 would give a much clearer insight into this possibility.

9.6.2.2. Hypothesis 3

In all users, over both interfaces, the error rates will be greater in the Command line system.

If we look at the recorded data for errors we find the following:-

\(^7\) Although the trends were in favor of the Gem system.
Number and rate of Errors

<table>
<thead>
<tr>
<th></th>
<th>Command Line</th>
<th>Wimp Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Number of Errors</td>
<td>22.50</td>
<td>13.05</td>
</tr>
<tr>
<td>Mean Time Between Errors</td>
<td>171.78</td>
<td>271.18</td>
</tr>
</tbody>
</table>

A single t-test reveals this to be a highly significant difference (N=11).

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>p</th>
<th>n</th>
<th>E(X-Y)</th>
<th>D(X-Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Number of Errors</td>
<td>-4.86</td>
<td>&lt;&lt; 0.001</td>
<td>11</td>
<td>10.09</td>
<td>6.55</td>
</tr>
<tr>
<td>Mean Time Between Errors</td>
<td>-2.91</td>
<td>&lt;&lt; 0.05</td>
<td>11</td>
<td>-132.54</td>
<td>143.74</td>
</tr>
</tbody>
</table>

It is interesting that even though the users were making significantly fewer errors in the Gem interface this had no statistically significant influence on their stated preference with regard to interfaces (see Hypothesis 2 above).

**Diagram Time Between Errors.**

This shows a similar pattern to that produced by subjects in both experiment 1 and 2, with the time between Gem based errors becoming longer and longer, and the time between command line errors becoming more constant as the session continues.

Using the error system developed for this study we find the following:
Figure 112. Command Line Errors, Experiment 3

Figure 113. Wimp Error Types, Experiment 3

Figure 114. Error Types from both Interfaces, Experiment 3

Adjusted Figures.

<table>
<thead>
<tr>
<th>(Adjusted)</th>
<th>Command Line (N=18)</th>
<th>Wimp Interface (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntax</td>
<td>11.38</td>
<td>9.61</td>
</tr>
<tr>
<td>Semantic</td>
<td>11.33</td>
<td>8.72</td>
</tr>
<tr>
<td>System</td>
<td>2.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

These are more clearly shown in the pie charts which show a much closer match between error categories than those produced for experiments 1 & 2.

9.6.3. Presentation Effects

To investigate any order effects within the data a Kruskal Wallis anova by ranks was carried out on the data, splitting the subjects data into order of presentation. The code 0 was used to represent Command line first and 1 was used to represent Gem system first.
This analysis revealed that subjects who experienced the command line system second (after using the Gem system) had significantly longer sessions (p < 0.01), issued more commands (p < 0.05), and completed more trials (p < 0.05) in their second session. This was initially interpreted as indicating that subjects were able to carry over the mental model from the GEM system over to the command line system. Within the Gem systems those who completed the command system first and then experienced the Gem system had significantly longer Gem session lengths (p < 0.01), issued more commands (p < 0.01), and attempted more trials (p < 0.05). Those who experienced the Gem system second issued their Gem commands more quickly (p < 0.01), but had a suggestively worse SSM performance (p < 0.0993). However this is probably statistical noise.

9.7. Parapsychological Data.

9.7.1. Overall SSM Effect.

The in session recording details for this experiment showed that subjects made 552 trials with 287 hits (Z = 0.93), which is within MCE.

Diagram showing SSM trials.

This shows a similar pattern between the two interfaces, but, as the SSM performance figure above shows there is no evidence for significant FAA.

Figure 115. SSM Trials, Experiment 3.
9.7.2. Hitters And Missers For Experiment 3.

As per the last experiments the TAQ2 and session data were split into those subjects who consistently hit and those who consistently missed\(^8\). A Kruskal Wallis anova by ranks was then carried out.

The only significant results from this analysis was those of the hit rates themselves, and since these were the variables used to determine if the subject was a hitter or misser this is of no interest.

**Experiment 3**

<table>
<thead>
<tr>
<th>SSM Sources</th>
<th>Exact Missing</th>
<th>Displacement</th>
<th>Proc. Speed</th>
<th>Reg. Wave</th>
<th>Quick Coin</th>
<th>Quick Stab</th>
<th>Unigram</th>
<th>Coin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Diagram showing SSM Sources, Experiment 3](image)

Figure 116. SSM Sources, Experiment 3.

9.7.3. SSM Decision Sources.

This shows an overall trend over all conditions (control and both interfaces) of hitting, but without any obvious pattern.


Based upon the results which were found in the two previous experiments the questionnaire data was analysed looking for similar effects. In the TAQ, Sex differences, Working style, Computer Attitudes, Use of machines, Early life, and Function linkage (both FLP & MLP). In both versions of the UEICS2, Sex differences, Previous experience with WIMP systems, Error correction, Feedback, Error Correction & Preven-

---

\(^8\) A curious effect was observed with this data. The author could only get 11 subjects who were prepared to complete sessions on both interfaces. These 11 subjects were very committed to helping the author, usually in return for the help he had given them with their coursework (see notes under subjects). Of these 11 subjects 6 consistently hit or missed over both interfaces (54%). This is an unusually high proportion, for example the first experiment had only 10 such people out of 33 (30%). It is possible that this could be due to some form of paranormal experimenter effect, or merely chance.
tion, Display layout, Use of the keyboard, and Command Name Consistency. All analyses were done using a non-parametric Spearman rank order correlation.

9.8.1. In Session Recording Correlations With The TAQ2 (n = 25).

Session Length

Longer sessions were found to be correlated with ownership of a computer at home (p < 0.05), and not reading about computers (p < 0.08). Subjects with longer sessions tended to report more problems with mechanical breakdowns (p < 0.08).

Commands

Subjects who used a computer at home issued more command (p < 0.07). More commands were associated with never using an autobank teller (p < 0.087).

Errors

Making more mistakes was found to correlate with having a microcomputer at home (p < 0.05), read books about computers (p < 0.05), or have frequent problems with camera breakdowns (p < 0.05). Slower error rates were associated with subjects who wore a digital watch (p < 0.063), had problems with photographic equipment (p < 0.05), and read about computers (p < 0.093).

Function Linkage.

Enhancing SSM performance was found to correlate with not wanting to be good at everything subjects did (p < 0.01), preferring to have someone else use the computer rather than using it themselves (p < 0.05), and being afraid of damaging the computer if they used it (p < 0.059).
9.8.2. Inter-Question Correlations From Within The TAQ2.

Age

Younger people appear to enjoy doing things in a specific order ($p < 0.068$), and following instructions ($p < 0.052$), they also had a tendency to think that computers helped them to work better. Older people were more likely to have a computer at home ($p < 0.053$), and were less likely to be bothered by the impersonal nature of computers ($p < 0.05$). Older subjects were found to use an autobank less frequently ($p < 0.01$). Sex differences

Higher subject number were associated with male subjects ($p < 0.01$), while females were less likely to be taking the course to get a good job ($p < 0.057$). The former of these correlation is because the final batch of subjects used to make up the numbers were mature male students.

Working Style

An enjoyment of figuring out how things work tended to correlate with liking to do things in a specific order ($p < 0.01$), to work with tools or equipment ($p < 0.05$), machines ($p < 0.079$), and thinking that computers are fun to use ($p < 0.084$). Those who liked working with tools or equipment liked working with machines ($p < 0.01$), repairing machines ($p < 0.08$), and agreed that computers could help them to work better ($p < 0.06$). Those that liked working with machines were not afraid of damaging computers if they used them ($p < 0.05$), and felt that computers were fun to use ($p < 0.09$). Subjects who reported that computers were too complicated for them to understand felt that computers were more dependable than human beings ($p < 0.05$), were bothered by the impersonal nature of computers ($p < 0.06$), and felt that computers were not fun to use ($p < 0.05$). Those people who wanted to be good at everything they did preferred to use the computer themselves rather than have someone else use it for them ($p < 0.05$). Those people who preferred to use the computer themselves were not afraid of damaging the computer if they used it ($p < 0.01$), or of looking silly if they made a mistake while using the computer ($p < 0.05$).

Computer Attitudes

Those who were afraid of damaging the computer if they used it were afraid of looking silly if they made a mistake while using the computer ($p < 0.05$), and were bothered by the
impersonal nature of computers \( (p < 0.05) \). The two questions which addressed fear of looking silly by making a mistake and by being seen not to know how to use the computer correlated quite strongly \( (p < 0.01) \). Subjects who had a fear of looking silly if seen to make a mistake, or if they were seen not to know how to use a computer tended not to use a programmable calculator \( (p < 0.05) \) and \( (p < 0.01) \) respectively. Subjects who were bothered by the impersonal nature of computers felt that computers did not help them to work better \( (p < 0.08) \). In contrast those subjects who felt that computers were fun to use felt that computers helped them work better \( (p < 0.05) \).

**Use Of Machines**

People who played video games did not think that computers were too complicated for them to understand \( (p < 0.08) \), and were not bothered by the impersonal nature of computers \( (p < 0.05) \). Subjects who reported using a computer at college were afraid of damaging the computer if they used it \( (p < 0.05) \), and were more afraid of looking silly if they made a mistake while they used the computer \( (p < 0.062) \). Those who used the mainframe computer frequently did not enjoy doing things in a specific order \( (p < 0.05) \), and were unafraid of damaging the computer if they used it \( (p < 0.05) \).

**Early Life**

Those that helped their family to repair equipment liked working with equipment \( (p < 0.079) \), machines \( (p < 0.05) \), and did not think that computers were too complicated for them to understand \( (p < 0.077) \). Subjects who often helped their family repair equipment were more likely to use a micro computer at home \( (p < 0.07) \) and to play video games \( (p < 0.01) \).

Those subjects who were allowed to play with dangerous equipment were more likely to use a micro computer at home \( (p < 0.05) \), and play video games \( (p < 0.05) \). However subjects who had helped their families to repair broken equipment had more problems with their audio equipment breaking down \( (p < 0.05) \), and were less likely to read about computers \( (p < 0.05) \). Subjects whose parents ensured that broken equipment was promptly repaired also tended not to read books about computers \( (p < 0.05) \). These last two correlations are not in the expected direction, and may be a by product of the general levels of 'computer phobia' in the subject group.
Subjects who used a microcomputer at college on a regular basis were encouraged to explore how machines worked \((p < 0.05)\), helped their families repair equipment \((p < 0.05)\), and were allowed to use dangerous equipment \((p < 0.05)\).

Subjects whose parents repaired equipment themselves frequently used a microcomputer at college \((p < 0.068)\), and reported\(^9\) fewer mechanical breakdowns \((p < 0.05)\). Frequent users of computers at college also tended to have parents who ensured that broken equipment was repaired promptly \((p < 0.07)\).

**Function Linkage (both FLP & MLP)**

Subjects who reported frequent problems with mechanical equipment felt that computers were too complicated for them to understand \((p < 0.05)\).

Left handedness\(^{10}\) was found to be correlated with having frequent problems with electrical appliances \((p < 0.05)\), TV \((p < 0.05)\), and feeling that computers did not help them work any better \((p < 0.097)\). Subjects who played video games had more breakdowns with their personal electrical appliances \((p < 0.07)\), photographic equipment \((p < 0.05)\), or their HiFi \((p < 0.08)\).

Subjects who reported having frequent breakdowns with personal electrical appliances also experienced problems with their video and TV equipment \((p < 0.05)\). Subjects who experienced problems with their TV and video also experienced problems with their HiFi and audio equipment \((p < 0.05)\). Those subjects who had been allowed to use dangerous equipment when they were children tended to have more problems with their photographic equipment \((p < 0.05)\).

---

\(^9\) One of the points which we have already covered in the previous chapter is that people who are competent with machines may underestimate the number of minor adjustments they make to the equipment they use. In contrast people who feel helpless, or intimidated by equipment will tend to overestimate the number of breakdowns. This could be the major factor in reported function and malfunction linked scenarios.

\(^{10}\) The sample size is too small to allow this to be any more than anecdotal.
9.8.3. UEICS2 Post Session Command Line Questionnaires (n = 18).

9.8.3.1. Correlations Within The Command Line In-Session Recordings.

Session Length, Commands and Errors.

Longer sessions were associated with rating the error correction as clear (p < 0.051), while rapid command rates were associated with rating the system as having few actions per operation (p < 0.076), and as having good feedback (p < 0.05)

Function Linkage.

Enhancing the SSM performance was found to correlate with the number of hits (p < 0.05), the system being rated as never preventing mistakes (p < 0.06), inadequate feedback (p < 0.01), slow performance (p < 0.069), and finding on-line help clear (p < 0.05).

9.8.3.2. Inter-Question Correlations From Within The Command Line Questionnaire.

Age

Subject number correlated with age so older subjects got higher subject numbers (p < 0.5).

Sex Differences

Males tended to get higher subject numbers (p < 0.05), and tended to rate the interface as having less power (p < 0.07). Females tended to be right handed (p < 0.05), rated the interface display layout as simplifying work (p < 0.052), and rated the command line system as easy to learn (p < 0.05).

Previous Experience With Command Line Systems.

Subjects who had never used a command line system before rated the interface as simplifying work by the way it, displayed its information (p < 0.05), that error messages indicated the corrective action (p < 0.05). Naive Command line users rated the system as stimulating (p
<0.083), and as having adequate power (p < 0.01). Subjects who rated the system as having too many actions per operation found the system slow (p < 0.05), and on-line help confusing (p < 0.09).

Feedback

Rating the system as having provided sufficient feedback was found to be correlated with feeling the work was simplified by display layout (p < 0.05), and system was too fast (p < 0.05).

Error Correction & Prevention

Rating that the system prevented mistakes correlated with feelings that feedback was sufficient (p < 0.05), work was simplified by the display layouts (p < 0.05), information required to complete each task was present on the screen (p < 0.072), and that the system had adequate power (p < 0.05). Finding that the system was too fast correlated with finding error messages helpful (p < 0.05), and the system easy to learn (p < 0.064). Feeling that error messages were helpful was found to correlate with error messages indicating the corrective action (p < 0.05), clear error correction (p < 0.05), clear on-line help (p < 0.05), and surprisingly the system being found difficult to use (p < 0.05). The corrective action being provided in error messages tended to be associated with the system being rated as being more stimulating to use (p < 0.01).

Display Layout

Feeling that work was simplified by good display layout was found to correlate with rating error correction as clear (p < 0.05), with the system as good (p < 0.069), and satisfying to use (p < 0.05). If the information required to complete each action was provided on the screen then subjects rated the system as having adequate power (p < 0.05).

Use Of The Keyboard

Subjects who found the keyboard easy to use rated the match between command name and action as being close (p < 0.05), command consistency to be high (p < 0.01), error messages as indicating the corrective actions to be taken (p < 0.05), and the system as easy to learn (p < 0.05), and easy to use (p < 0.05).
Command Name Consistency

Closeness of command name and action was found to correlate with command consistency (p < 0.01), prevention of mistakes (p < 0.05), and that the information needed to complete each command was available on the screen (p < 0.05). Closeness of command to action was found to make the subjects rate the system as easy to use (p < 0.05), and as having adequate power (p < 0.051).

Command consistency was found to correlate with clarity of tutorial (p < 0.078), helpful error messages (p < 0.05), and surprisingly the error messages not indicating the corrective action (p < 0.9). Command consistency correlated with subjects finding the system easy to use (p < 0.01), and as having adequate power (p < 0.077).

Clarity Of The Tutorial.

Subjects who rated the tutorial as clear felt the displays were good (p < 0.01), error correction was clear (p < 0.05), that the system was good (p < 0.01), satisfying (p < 0.01), stimulating (p < 0.05), easy to use (p < 0.05), learn (p < 0.056) and had adequate power (p < 0.084).

Internal Validity Of The Questionnaire.

Subjects who rated the system as easy to use felt the system was good (p < 0.01), satisfying (p < 0.05), stimulating (p < 0.05), and easy to use (p < 0.05).

9.8.4. EUICS2 - Gem Post Session Questionnaire (N = 18).

9.8.4.1. Correlations Within The Gem In-Session Recordings.

Session Length

Older subjects tended to complete longer sessions (p < 0.08) The more stimulating subjects found the system the longer their session (p < 0.01).
Commands

The more difficult subjects rated using the keyboard and mouse at the same time the more commands they made (p < 0.07). More commands were also associated with how stimulating subject found the interface (p < 0.057). While in contrast issuing fewer commands was found to be correlated with feeling that the system prevented mistakes (p < 0.01), and that error messages indicated the corrective actions (p < 0.05). It could be that when an interface is found to be too safe and helpful it no longer stimulates the user.

Males issued their commands more quickly (p < 0.09). Slower command rates were found to be associated with finding the keyboard and mouse easy to use (p < 0.064), that error prevention was good (p < 0.05), and that error messages indicated the corrective actions (p < 0.094).

Errors

Making many errors were associated with finding pull down menus hard (p < 0.05), with rating the match between command name and action as poor (p < 0.05), rating command consistency as poor (p < 0.01), on-line help as poor (p < 0.08), error messages as unhelpful (p < 0.06), and display layouts as poor (p < 0.05). A system was rated as good (p < 0.05), and satisfying to use (p < 0.01) if the subject made few errors on it.

Slow error rates were associated with feeling the system was slow (p < 0.64), command names matching their action (p < 0.05), command consistency (p < 0.01), and good screen layouts (p < 0.01). Slow error rates correlated with subjects rating the system as good (p < 0.01), satisfying (p < 0.01), and easy to learn (p < 0.05). However a difficulty in understanding the concept of pull down menus made the inter error rate worse (p < 0.05).

Function Linkage.

Enhancing the SSM performance was found to be correlated with previous use of a mouse system (p < 0.05), clear tutorial instructions (p < 0.05), helpful error messages (p < 0.05), error messages indicating corrective action (p < 0.05), clear error messages (p < 0.05), clear on-line help (p < 0.05), and the information required to complete each action being present on the screen (p < 0.05).
9.8.4.2. Inter-Question Correlations From Within The Gem Questionnaire.

Age.

Older subjects reported finding the mouse harder to control (p < 0.06), and the joint control of the mouse and keyboards harder as well (p < 0.05). Older subjects reported that they felt there were too many actions in each operation (p < 0.092), that error messages were unhelpful (p < 0.09), but that the gem system was more stimulating to use (p < 0.065).

Sex Differences

There were more left handed females than males (p < 0.06), but then the samples gender distributions were highly biased anyway. Females rated the system as preventing mistakes more than their male counterparts (p < 0.05). Females also felt that error messages were more helpful (p < 0.09), and that these error messages were more likely to give the corrective action (p < 0.05).

Previous Experience With Wimp Systems

Those subjects who had used a mouse before felt the system required too many actions per operation (p < 0.065), and that the systems displays never simplified the work (p < 0.087).

Those subjects who found the mouse hard to control also found the pull down menus hard to understand (p < 0.063), and found it hard to coordinate the mouse and keyboard simultaneously (p < 0.05). These subject also felt that the system did not prevent mistakes (p < 0.05). Finding the mouse hard to control correlated with feeling that the system did not simplify the work by the way it displayed its information (p < 0.085). Not surprisingly subjects who found the mouse hard to control also found it hard to learn the system (p < 0.052), and felt that the system demanded that they had to memorize the information needed to complete each task (p < 0.06). Subjects who found the concept of pull down menus hard to understand found using the mouse and keyboard simultaneously (p < 0.01). These people also reported there was little command consistency (p < 0.01), that it was hard to learn the system (p < 0.05), that the system was bad (p < 0.05), and that it was frustrating to use (p < 0.05).
Use Of The Keyboard

Those subject who felt that using the mouse and keyboard simultaneously was easy rated command consistency as high (p < 0.05), the system as preventing mistakes (p < 0.06), on-line error messages as helpful (p < 0.05), and that they indicated the corrective actions to be taken (p < 0.07). These people felt the system was good (p < 0.05).

Feedback

Those subject who felt that feedback about commands was sufficient felt that on-line help was clear (p < 0.05), that the system was easy to learn (p < 0.05), and satisfying to use (p < 0.05).

Error Correction & Prevention

Those subjects who felt that the error messages were helpful also felt the error messages indicated the corrective actions to be taken (p < 0.01). These users found the system was satisfying to use (p < 0.05). Those subjects who found error correction clear also reported that on-line help was clear (p < 0.05), that the information needed to perform each task was available on the screen (p < 0.084), and that error messages indicated the corrective actions to be taken (p < 0.056). Subject who felt that the system prevented mistakes felt that error messages were not helpful (p < 0.05). This is not in the expected direction.

Display Layout

Those subjects who found on-line help clear reported that the system presented the information they needed to complete each task on the screen (p < 0.05), that the system good (p < 0.01), and satisfying to use (p < 0.05).

Command Name Consistency

Subjects who felt that the word used as a command closely matched the action it carried out reported command consistency (p < 0.05), that the system was good (p < 0.08), and that the way the displays were laid out made the work easier (p < 0.05). However these people rated the system speed as too slow (p < 0.05). Command consistency correlated with ease
of learning the system (p < 0.05), good overall reactions to the system (p < 0.01), and finding the system satisfying to use (p < 0.05).

Clarity Of The Tutorial.

Those subjects who found the instructions easy to follow found the system feedback sufficient (p < 0.05), error messages helpful (p < 0.05), on-line help clear (p < 0.05), the information needed to complete commands accessible on the screen (p < 0.01), and the system satisfying to use (p < 0.05). Clear instructions were correlated with the ease of error correction (p < 0.08).

Excessive Complexity.

Those subject who felt the system demanded too many actions per operation found error messages were never helpful (p < 0.087).

Difficulty In Using The System.

Difficulties when using the system correlated with being left handed (p < 0.087), the tutorial being confusing (p < 0.05), there being a lack of error protection (p < 0.054), insufficient feedback (p < 0.05), error messages which did not indicate the corrective action (p < 0.065), and confusing on-line help facilities (p < 0.01). Subjects who rated the system as having these faults experienced difficulty in learning the system (p < 0.093), frustration (p < 0.05), had bad overall reactions to the system (p < 0.078), and felt using the system was dull (p < 0.064).

Correlates With System Power.

The system was rated as having adequate power if they found pull down menus easy (p < 0.06), the instructions easy (p < 0.05), error messages helpful (p < 0.05), error messages which indicated the corrective actions (p < 0.055), error correction clear (p < 0.055). Adequate power was also associated with the system being good (p < 0.05), satisfying (p < 0.01), easy to learn (p < 0.05), and easy to use (p < 0.05).
9.9. Discussion Of Findings From Experiment 3.

9.9.1. HCI

The in-session recordings seem to indicate that the recording processes were functioning correctly. The Gem interface was found to be faster, but the command line system had a greater number of commands issued to it, although these differences are not significant. The command time graph shows that the command line system was initially faster, but then the Gem interface showed a decline to a steady plateau, which was lower than that of the command line system. This was a similar trend to that found in the previous experimental data. It was not possible to administer the post session questionnaires to these subjects, so the only index of preference was from the last 5 questions of the UEICS2. This subject group showed no preference for either of the interfaces in terms of satisfaction, stimulation, ease of use, or rated power. This could be due to a number of factors, the simplest of which is that the subjects were totally disinterested in the experiment and took no active involvement in its activities. Alternatively the lack of preference could be due to the subject group having no previous experience of computers, or them simply disliking all computer systems regardless of the interface type. The last of these possibilities would be an interesting finding since it would indicate that computer phobics were in a specialised sub-group. However the few subjects who did turn up for the experiments cannot have been truly computer phobic, or totally disinterested in the experiment, otherwise they would have simply stayed away. The possibility that the subject group had no preference because they had no previous computer experience is also felt to be unlikely, since the subjects from the first experiment were also computer naive, and yet had the strongest preferences of any of the experimental groups. We are left with the possibility that the subject group were uncertain to the meanings of these words in the computing context, and therefore rated both interfaces similarly. To check this possibility the variance figures for the two interfaces were compared, and showed there was a very small amount of variance in the Gem interface, but this would not explain the lack of difference between the two interfaces. There was found to be a significant difference in the number and rate of errors in the two interfaces, in favor of the Gem interface. The semantic and syntactic error ratios showed a similar pattern to those found in the previous two experiments. A significant learning effect was found in the analysis for presentation effects.
9.9.2. **Parapsychology**

The overall SSM measure was again well within chance, and showed no evidence for the existence of paranormal FAA. However there was an unusually high ratio of consistent hitters and missers within the subject group. Unfortunately this extra proportion of hitters and missers did not reveal any evidence of FLP/MLP tendencies in the TAQ2 data. There was nothing of interest in the SSM source diagram, except that the Gem interface displacement source indicated that the subjects repeated very few of their commands, this resulted in reducing the displacement artifact.

9.9.3. **Questionnaires.**

9.9.3.1. **TAQ2 data.**

The TAQ2 session data showed no surprising results, except some weak correlations which were interpreted as being statistical noise. Those SSM function enhancement correlations which did appear were all consistently in an unexpected direction. If this was a paranormal effect, it is likely to have been due to those subjects wishing to 'help' the experimenter, or to some form of experimenter effect. The author believes that these are more likely to be due to some form of artifact, or statistical noise than to some form of paranormal action. As in all the previous subject groups the role of parental encouragement to explore machines was found to be the strongest indicator of function linkage in later life.

9.9.3.2. **Command Line System**

Similar findings as those from the previous experiments emerged with these subject groups. The strongest factor to emerge from this groups appraisal was that their satisfaction with the command line interface was found to be related to the error prevention and correction facilities. The other factors such as command consistency and ease of using the keyboard were also found to influence the subject feelings about the interface, but not to the extent of the previous groups.

9.9.3.3. **Gem Interface System**

The ease of using the mouse and keyboard were found to be of great importance to this subject group. not just in the questionnaire data, but also in the more objective measures of...
system use. In fact within the Gem interface the in session details provided more useful information, from a software designer’s point of view than the questionnaire data. This could be an important factor in dealing with totally computer naive subjects. Females seemed to prefer the Gem interface, since they rated its error prevention and correction facilities more highly. The other factor to emerge was the importance of the on-line help facilities to this subject group.

9.10. Findings From The Comparison Of The Overall Results From Experiments 1 to 3.

9.10.1. Combining the in-session recording differences and conducting a parametric correlation we found the following:


Longer session lengths were found to be correlated with an increased number of errors (p < 0.01), more commands (p < 0.01), and more trials (p < 0.01). An increased number of errors were found to be correlated with issuing more commands (p < 0.01), and increased error rates (p < 0.01). An increased number of commands was found to correlated with attempting more trials (p < 0.01), and a faster rate of issuing commands (p < 0.01).

9.10.1.2. Gem Overall Totals Parametric Correlations (N = 86).

Increased session lengths were found to be correlated with issuing more commands (p < 0.01), attempting more trials (p < 0.01), and issuing commands more quickly (p < 0.01). More errors were found to be associated with faster error rates (p < 0.01). More commands were found to be correlated with an increase in the number of attempted trials (p < 0.01), and a faster command rate (p < 0.01).

9.10.1.3. Overall Session Details Parametric Correlations (n = 106).

Longer sessions were found to be correlated with more errors (p < 0.01), more commands (p < 0.01), more attempted trials (p < 0.01), and a faster error rate (p < 0.01). More errors were found to be associated with having issued more commands (p < 0.01), attempted more
trials (p < 0.01), and having a faster inter-error rate (p < 0.01). More commands were found to be correlated with having attempted more trials (p < 0.01), and having faster inter-command (p < 0.01) and error rates (p < 0.01).

These correlations can be interpreted as confirmations that the resident recording systems were functioning normally throughout all three of the experiments.

9.10.2. Investigation Of Experimental Hypothesis From The Overall Data.

9.10.2.1. Hypothesis 1 -

Wimp Interfaces Will Have A Slower Throughput Of Work.

Work differences between interfaces.

<table>
<thead>
<tr>
<th></th>
<th>Command Line</th>
<th>Wimp Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session Length</td>
<td>2438.92</td>
<td>2248.62</td>
</tr>
<tr>
<td>Mean Number of Commands</td>
<td>35.02</td>
<td>32.39</td>
</tr>
<tr>
<td>Mean rate of Commands</td>
<td>78.21</td>
<td>77.62</td>
</tr>
</tbody>
</table>

A Parametric single t-test (n = 65) on this data shows no significant differences between the session lengths (p < 0.1), the number of commands (p < 0.5), or the rate of issuing commands (p < 0.8) over the two interfaces. In these unadjusted figures the command line system shows a slight superiority, but we have not taken the 'help bias' into account.

Total Number Of Helps.

<table>
<thead>
<tr>
<th></th>
<th>Command Line</th>
<th>Wimp Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td>115</td>
<td>30</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>119</td>
<td>28</td>
</tr>
<tr>
<td>Experiment 3</td>
<td>85</td>
<td>24</td>
</tr>
<tr>
<td>Total Number of Helps</td>
<td>319</td>
<td>82</td>
</tr>
<tr>
<td>Total Commands</td>
<td>2977 (n=85)</td>
<td>2786 (n=86)</td>
</tr>
<tr>
<td>Less Help Commands</td>
<td>2658</td>
<td>2704</td>
</tr>
</tbody>
</table>
Gives command averages of Wimp = 31.44 & Command = 31.27 respectively.

The mean time between command becomes 71.52 seconds for the Wimp, and 77.99 seconds for the Command line systems respectively. This shows that the two interfaces did a similar number of commands, but the difference in session length made the Wimp system faster in the rate with which it processed commands (p < 0.05).

9.10.2.2. Hypothesis 2

Once competence is reached in the command line interface, it would be preferred by experienced users and naive users who are 'keyboard literate', due to the interface's greater information interchange rate, and the familiarity with the keyboard.

Questionnaire Preferences Between The Two Interfaces.

The planned analysis using the post session questionnaires could not be conducted since the essays had been dropped from the protocol due to its lack of validity, and the problems involved in its administration. This would have been impossible with the last group of subjects anyway.

Other Measure Of Preferences Between The Two Interfaces.

The post session command and gem system questionnaires were combined and a non-parametric wilcoxon matched pairs test (N=70) was carried out on the last 5 questions on the two questionnaires which asked the subjects to rate the systems with regard to goodness, satisfaction, stimulation, ease of use, and 'power'.

This analysis revealed that subjects rated the Wimp interface as better (p < 0.01), more satisfying (p < 0.05), more stimulating (p < 0.001), and easier to use (p < 0.01). However subjects rated no overall difference in regard to power (p < 0.2). This confirms Macleod's argument (Personal communication Macleod, 1988) that the notion of 'power' was too abstract for these subject populations. Indeed it is probably too vague a concept to be of any use regardless of subject population. It should be dropped from any future questionnaires.
9.10.2.3. Hypothesis 3

In all users, over both interfaces, the error rates will be greater in the Command line system.

Looking at the overall number of errors made in the interface we find:

<table>
<thead>
<tr>
<th></th>
<th>Command Line</th>
<th>Wimp Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Number of Errors</td>
<td>30.23</td>
<td>13.46</td>
</tr>
<tr>
<td>Mean Rate of Errors</td>
<td>115.712</td>
<td>242.92</td>
</tr>
</tbody>
</table>

A parametric single t-test (n = 65) on this data shows a highly significant difference for both the number of errors (p < 0.001), and the mean time between them (p < 0.001). This confirms hypothesis two.

9.10.3. Presentation Effects Over All The Experiments 1 To 3.

To look at the presentation effects over all the experiments numbered 1 to 3, we carried out Kruskal Wallis anova by ranks on the in-session recorded data with the codes 0 = Com first, and 1 = Gem first. This revealed the following effects.

Differences In Subjects Behaviour When They Experienced The Command Line System Last.

The command line environment session lengths and errors were found to have varied significantly, so subjects who experienced the command line system in the second session had longer sessions (p < 0.001), and made more mistakes (p < 0.05) than subjects who experienced it as their first interface.

Differences In Subjects Behaviour When They Experienced The Gem Interface System Last.

The session lengths and command rates for gem sessions was also found to vary with presentation order, so subjects who experienced the gem interface second had longer session times (p < 0.001), issued more commands (p < 0.0001), and issued their command
significantly faster ($p < 0.001$). However unlike the command line interface they made only suggestively more errors ($p < 0.075$).

## 9.10.4. Parapsychology.

### 9.10.4.1. SSM Measures

The total number of trials was 3024 and out of that there were 1537 correct guesses ($Z = +0.909$), which is within MCE, and at 50.8% is smaller than most of parapsychological effect sizes reported in the review in chapter two.

**Function Enhancement Over Interface Types.**

- Command hit rate 2.078
- Gem hit rate 2.097

The combined data for the SSM performance shows virtually no difference for the two systems. Both are so close to mean chance expectation (2.0) that it is not necessary to conduct any statistical tests upon them.\(^\text{11}\)

**Looking For Evidence Of Function Linkage In The Combined TAQ2 Data.**

We combined the TAQ2 data and the overall (combined gem & com) session recorded data in to see what differences existed between those subjects who were defined as hitters, missers, and those that scored at MCE. This was done using a Kruskal Wallis 3 way anova by ranks. We split the data up so $0 =$ 'misser', $1 =$ MCE, $2 =$ 'hitter'. This only revealed significant correlations for the hit rates used to form the categories (in both cases: $p < 0.001$, $N = 65$).

### 9.10.4.2. SSM Decision Making Sources For All Three Experiments.

- Exp 1-3 SSM Sources - Command Line Systems
- Exp 1-3 SSM Sources - Gem Systems

---

\^\text{11} The difference could well be due to rounding errors within the statistical package used.
Decision making systems and the SSM.

Several decision making recordings have been made in the study. These diagrams show an analysis of the success or otherwise of each of the decision making elements within the experimental SSM system\(^2\). Unfortunately none of the sources of decision making showed

---

\(^2\) The control process took account of the displacement artifact, and shows a similar trend.

---
any significant deviation from MCE, the only point of interest being the dramatic success (in comparison), of the displacement detector (under both control and test conditions). The reasons for this is the unavoidable displacement artifact (discussed in appendix 13). It had been hoped that the experiment would be able to show which sources of decision making and randomness performed best, in order to give an insight into how any 'psi' interaction took place.

9.10.5. Parapsychological Discussion (Smart Noise).

Unfortunately these experiments provided no evidence for the existence of smart noise, this might be due to several factors, which we will briefly discuss. Some parapsychologists (Von Lucadou, 1987) have postulated that any totally defined system cannot display any 'psi'. Since the SSM tried to exclude as many uncontrolled variables as possible it might fall into this category. Some fellow parapsychologists in the Koestler chair laboratory attributed the author's attitude as being responsible for making the SSM results come out in the 'no psi' direction. The experimenter has no fixed belief system which demands that 'psi' exists, and has adopted the attitude that it's existence would have to be empirically demonstrated before he was willing to accept that it existed. Until that time 'psi' remains only one of many possible explanations which the author considers could be used to explain supposed anomalies. The other possibility raised by some fellow researchers was that the author had a deep unconscious 'fear of psi' (Tart & Lahore, 1986). Unfortunately, as most parapsychologists admit, both of these explanations are as unhelpful as they are unfalsifiable. Instead it is better to consider more testable possibilities. For example it is possible that the application did not give enough feedback with regard to SSM success, or that the feedback was too ambiguous. It is also possible that subjects have in some way to know they are in a 'psi usable' situation before they make use of these supposed abilities, although there is experimental evidence (Stanford, 1974) against this notion. Experiment 4, the so called 'function linked' experiment tried to rectify some of these possible shortcomings. This experiment is discussed in chapter ten.
9.10.5.1. Skeptical Software Monitor.

The SSM was an attempt to use logic to track down the working of 'psi' and to use it to enhance system performance. It is a development of the concept of smart noise which keeps track of the effectiveness of the SNS. The SSM increases or decreases the amount of notice that the system takes of the various noise sources, dependent upon the sources success in the most recent past. Under ideal conditions it would take the most notice of the successful decision making systems, and least notice of the unsuccessful sources. The SSM also checked for displacement and exact missing, which should in theory have found any mis-directed 'psi' at the moment it was occurring and adjust the system to take most advantage of it. What actually happened was that the FAA in most systems was so random that the SSM spent most of its time adjusting decision weightings. In conclusion the SSM and it’s associated experimental design worked beyond expectations, all that was missing was the 'psi'.

9.10.6. Investigation Of The Interface Hit Rate Correlations Within The Questionnaire Data.

A series of analysis were conducted to look at possible effects within the questionnaire data, and the respective hit rates from the session data. To do this the questionnaire data was combined and a non-parametric spearman rank order correlation carried out using the respective hit rates from each session, or in the case of the TAQ2 the overall hit rates.

9.10.6.1. Hit Rate Correlations From Within The TAQ2 Combinedc Data (N= 106).

Enhancing the SSM performance was found to correlate with more hits (p < 0.01), never using a bank auto teller (p < 0.075), never using a micro at college (p < 0.094), often being injured by a machine when young (p < 0.01), and machines often breaking down in the childhood home (p < 0.05).
9.10.6.2. Function Linkage Correlations From Within The UEICS2 Combined Data.

**Command Line Systems** (n = 85).

Enhancing the SSM performance on the command line system was found to be correlated with an increased number of hits \( (p < 0.001) \), and having clear on-line help \( (p < 0.093) \).

**Gem Interface Systems** (n = 86).

Enhanced SSM performance was found to be correlated with an increased number of hits \( (p < 0.001) \), error messages being found to be helpful \( (p < 0.05) \), clear error correction \( (p < 0.01) \), clear on-line help \( (p < 0.05) \), and adequate power \( (p < 0.05) \).

9.10.7. Inter-Group Comparison Of The Questionnaire Data.

The three TAQ2 data files were combined into one large one, and a code to define which data came from which experiment was added (one = Experiment 1, two = Experiment 2, three = Experiment 3). After this a non-parametric Kruskal Wallis anova by ranks \( (N = 106) \) was conducted. The results were split into the same categories which were used in the three main experiments. In these results group one will refer to the subjects from experiment one, who were touch typists with little computer experience. Group two will refer to the postgraduate systems analysts, who were experienced computer users. Finally group three will refer to the non typist, computer naive subjects in experiment 3.

**Age.**

Group two were significantly older than the other two groups \( (p < 0.01) \).

**Sex Differences**

Group two had significantly more males than the other two groups \( (p < 0.01) \).

**Working Style**

Group two were significantly more inclined to enjoy working out how things work \( (p < 0.05) \) than the other two groups. Group three were least likely to enjoy figuring out how
things work. This spread is more likely to be correlated with the groups educational attainment than computing background. For example postgraduate students would be more likely to enjoy problem solving. This response pattern was reversed for the groups rating of the question of if they were taking the course to get a good job. The postgraduates were least likely to answer in the affirmative, while the final group were the most likely to be completing the course to get a good job (p < 0.01). Subjects in group one were less likely to want to be good at everything they did than those from the other two groups (p < 0.077).

Computer Attitudes

Level of education was also found to play a strong part in computer attitudes. The subjects of group three were most likely to feel that computers were too complicated for them to understand. While subjects from group two were likely to feel that they could easily understand computers (p < 0.05). Subjects from group two were also much more likely to dislike having someone else use the computer for them (p < 0.073).

A surprising result emerged for the ratings from how much fun the groups felt computers were. Group one were strongest in their disagreement with this statement, and group three most strongly in agreement with the proposition (p < 0.05). It was expected that the systems analysts would rate computers as being most fun to use.

Use Of Machines

Group one disagreed most strongly with the proposition that computers helped them to work better, while subjects from group two strongly agreed with the statement (p < 0.01). Subjects from group three were non-committal in their replies. Ratings of the use of a microcomputer at college significantly divided the subject populations (p < 0.01). This division was such that the first group used a micro most, then the third group (they obviously included the experiments), and finally the second group. Due to the timing of the experimental sessions with the second group their experiments took place on the week of their induction to the college, and they therefore had no previous use of the computing facilities. In the normal course of events computer scientists would be among the heaviest users of the computing faculties at Napier. This timing artifact was also responsible for the finding that subjects in group two did significantly less wordprocessing (p < 0.01), and usage of the mainframe (p < 0.01) than the other groups. These corrections highlight how honestly the subjects
answered their questionnaires. As expected the subjects from group two did significantly more reading about computers than those in the other two groups (p < 0.001).

**Early Life**

There were two intriguing correlations from the early life questions. Subjects from group three (the nearest we came to computer phobics), had received less encouragement from their parents to use dangerous equipment (p < 0.058) than those in groups one and two. The subjects from the other two groups had approximately the same amount of use of dangerous equipment.

The answers to the question that asked how promptly the subjects families repaired broken equipment also showed a slight trend towards group three’s parents being most prompt in repair. Group one’s parents being less prompt, and group two’s parents slowest in ensuring that broken items were promptly repaired (p < 0.073).

**Function Linkage (both FLP & MLP)**

The question which asked about the perceived number of breakdowns which occurred with their personal electronic belongings revealed that group two had the most breakdowns, and group one the least (p < 0.05). This is most probably due to a 'tinkering effect', where people with an interest in high technology tend to investigate how the equipment operates. This tendency was most noticeable in the findings from the fourth experiment.

9.10.7.1. In Session Recordings.

**Session Length.**

There was a significant difference in the session lengths between the three groups (p < 0.05). Subjects in group two had the longest sessions, with group three having the shortest session length.

**Errors.**

The number of errors showed a significant split over the three groups such that group two showed the most errors, group one the next largest, and group three showing the least (p <
The increased number of errors made by the group two is due to their investigating the systems capabilities. The discomfort felt by subjects in group three towards computers resulted in them being over cautious, and hence getting the least number of errors. This trend was confirmed by the finding that group two also had the fastest error rates (group three having the slowest (p < 0.05)).

**Commands**

Subjects from group one issued the most commands, and those from group three the least (p < 0.05). This was probably due to group one subjects following the tutorial exercise more rigorously.

The mean rate of issuing command showed an unexpected difference between command rates (p < 0.0001). Subjects from group one issued their commands significantly faster than those of the other two groups, with subjects from group two being the slowest. This was not expected, and could be due to group two’s subjects thinking about what the commands were doing, rather than mechanically following the tutorial instructions.

Function linkage and SSM system.

The number of 'psi' trials (which were covert), had a significant difference in their distribution (p < 0.05). Group one issued the most, group two the next, and group three issued considerably less than the other two groups. This probably reflects the degree to which subjects followed the tutorial sheet, rather than any clairvoyant preference.

**9.10.7.2. Summary Of In Session Differences.**

To contrast these figures let us look at extracts from the data gathered in the previous three studies, all of which used an identical system.
Experiment 1 (n = 46) Trained touch typists (not used a computer before).

- Mean time between errors = 237.837 seconds
- Mean time between commands = 68.135 seconds
- Psi Influence Index (hit rate) = 2.230

Experiment 2 (n = 22) Postgrads in systems analysis (used computer & keyboard).

- Mean time between errors = 230.432
- Mean time between commands = 90.908
- Psi Influence Index (hit rate) = 2.283

Experiment 3 (n = 18) Never used a computer before & not touch typists.

- Mean time between errors = 271.184
- Mean time between commands = 85.645
- Psi Influence Index (hit rate) = 1.699

9.10.7.3. Post Session UEICS2 Questionnaires Experiment 1 to 3.

The three experiments used the same post session questionnaires for both the command line and gem environments. The next analysis looks at the differences that emerge when a three way Kruskal Wallis anova by ranks is carried out on the combined questionnaire data.


Age

The age of subjects who completed a command line session varied significantly, such that subjects from experiment 2 were considerably older than those from the other two experiments (p < 0.01). This confirmed the age correlations found in the TAQ2 data (above).
Sex Differences

The sex difference for the three groups was also significant, such that the populations of groups one & three were more predominately female, and those of group two were more male ($p < 0.01$). This confirmed the TAQ2 finding.

Previous Experience With Command Line Systems

The number of subjects in each group who answered that they had never used a command line system before was highest in group three, closely followed by group two. Most previous command line users came from subjects within group one ($p < 0.05$). The command line system’s speed was also a case for some significant divisions among the subject groups ($p < 0.05$). Group one rated it as being fastest, then group three, and finally group two who rated it as being slow. This would merely be a reflection of the amount of previous computing experience held by the three groups.

Clarity Of Tutorial.

There was a clear division on the question of which subject groups found the instructions which described the tasks they had to do as being clear. Subject group one found them clearest, group two the next clearest, and group three found them confusing ($p < 0.05$). The same instructions were used for all groups.

Feedback

The adequacy of the feedback provided by the system after each command also varied significantly over the three groups ($p < 0.05$). This such that group three rated the feedback as being good, group two less good and group one as being worst.

Error Correction & Prevention

There was a clear division between the groups on how much they felt the system prevented mistakes ($p < 0.01$) Group one rated the system worst (most error prone), group two next
worst, and group three rated it as preventing mistakes. Again this probably reflects the amount of previous experience the subject had with computer systems (the experienced subjects had seen better and worse systems than the KMDB system). The rating as to how much error messages indicated the corrective action varied significantly (p < 0.05). Group one felt most strongly of the three groups that the error messages did not indicate the corrective action. In contrast groups two & three felt that error messages did indicate the corrective action.

9.10.7.4.1. In Session Recorded Data Command Line Systems.

Commands

The number of commands issued by the three different subject groups varied significantly (p < 0.01). Group one issued the most, group three the next most, and group two issued the least number of commands. The rate at which subject groups issued commands also varied significantly over the command line system (p < 0.001). Such that group one subject issued them fastest, group three the next fastest, and group two the slowest.

Errors

The number of errors made on the command line system varied significantly over the three subject groups (p < 0.05). Group two made the most errors, then group one, and finally group three who made very few errors in comparison to the other two groups. The rate at which the different subject groups made errors varied suggestively over the command line systems (p < 0.055). The subjects in group three were the slowest to make errors, those in group one the next slowest, and the subjects in group two were the fastest at making errors on the command line system.

Function Linkage

The number of 'psi' trials on the command line system varied significantly (p < 0.01) over the three groups. Remember these were covert trials, but probably only reflect how strictly the subjects followed the tutorial instructions. Group one issued the most trials, group three the next most, and group two the least.
9.10.7.4.2. Summary.

The in session recorded details from the command line system shows a similar trend to those found overall. This can be taken to indicate the data is a measure of the general behaviour of the groups towards computer systems, and not a specific response to a particular environment.

9.10.7.5. Post Session Gem UEICS2 Questionnaire Data (n = 86).

We next conducted a similar Kruskal Wallis anova by ranks comparison on the data from the Gem experimental sessions over all three experiments. There were similar effects on Age, and Sex, as was found on both the command line data and the overall totals.

Learnability

The ease with which groups felt they could learn the system also varied suggestively ($p < 0.089$), such that group one thought the Gem system was easy to learn, group two less easy, and group three felt it was hard.

Error Correction & Prevention

Subjects from group three rated the Gem system’s error correction as being slightly more confusing than the other groups ($p < 0.1$). This was probably due to their dislike of computer systems.

Finding The Gem Interface Stimulating.

The amount the subjects rated the system as either dull or stimulating varied quite suggestively ($p < 0.07$). The subjects of group two feeling it was most stimulating, followed by group three, and then group one.
9.10.7.5.1. In Session Recorded Data Gem Interface Systems.

Session Length

The Gem based session lengths varied significantly between subject groups (p < 0.05). The sessions generated by group one were longest, then came the sessions of group two, and finally the sessions of group three, which were the shortest.

Commands

The number of commands issued during gem sessions also varied significantly over the three groups (p < 0.01). This was such that group one issued the most commands, group two the next, and then group three with the least number of commands. The rate at which subject groups issued commands varied quite significantly over the three sessions (p < 0.01). This was such that group one was the fastest, group two the next fastest, and group three the slowest.


TAQ & Overall Session Data Correlations With Sex Differences.

To look at the sex differences within the subject populations we combined the TAQ data with the overall session recordings, allocated a sex code of 0 for females, and one for males respectively, and then performed a Kruskal Wallis anova by ranks with the following results:

In Session Recordings.

Females were found have made fewer errors (p < 0.05), had a slower inter-error time (p < 0.05), issued their commands faster (p < 0.085), and have been younger (than the male subjects) (p < 0.01),

Working Styles.

Females were found to be less likely to enjoy hobbies which involved doing things in a specific order (p < 0.064), less likely to enjoy repairing machines (p < 0.06), were more
afraid of looking silly if they were seen by other not to know how to use the computer (p < 0.05). Females used digital watches less (p < 0.01), played fewer video games (p < 0.05), and read less about computers (p < 0.05).

**Early Childhood.**

In their early childhood females were less likely to have been allowed to use dangerous equipment (p < 0.01), were less likely to have been encouraged to explore how machines worked (p < 0.05), more likely to report having been frightened by some kinds of equipment (p < 0.05).

**9.10.8.1. Post Command Session Questionnaire And Command Session Data Correlations With Gender.**

The post Command line system questionnaires were combined with the command line in-session recordings. The resulting data file was then analysed on gender using a Kruskal Wallis anova by ranks, with the following results:

**In Session Recordings.**

Females were found to have issued more commands (p < 0.01), attempted more trials (p < 0.01), issued their commands faster (p < 0.01), and tended to be younger than the males (p < 0.01).

**Comments About The System.**

Females felt there were too few operations required to complete a task (p < 0.05), and that the system was too fast (p < 0.05).

**9.10.8.2. Post Gem Session Questionnaire Data And Session Data Correlations With Sex.**

The post Gem system questionnaires were combined with the Gem in-session recordings. The resulting data file was then analysed on gender using a Kruskal Wallis anova by ranks, with the following results:
In Session Details.

Females tended to have longer sessions ($p < 0.05$), made fewer errors ($p < 0.05$), issued more commands ($p < 0.064$), made errors more slowly ($p < 0.05$), and tended to be younger ($p < 0.05$).

Using The Wimp System

Females found the mouse harder to control ($p < 0.05$).

The conclusions derived from these analyses are given in the following chapter.
10. Overall Discussion on Experiments 1 to 3.

Quote:

To see the world in a grain of sand And a heaven in a wildflower: Hold infinity in the palm of your hand, And eternity in an hour.

William Blake.

10.1. Conclusions From The Comparison Of The Overall Results Of Experiments 1 To 3; Experiment 4, The TAQ Factor Analysis And The Overall Conclusions And Recommendations For Future Research.

10.1.1. Overall Discussion on Experiments 1 to 3.

10.1.1.1. HCI

The in-session recordings demonstrated that the interface recording systems functioned correctly. Hypothesis one was falsified because although both interfaces had a similar number of commands, the rate at which commands were issued was faster in the Gem interface. Overall the subjects rated the Gem interface as significantly better, more satisfying, more stimulating, and easier to use than the command line system. The Gem interface was also found to be significantly superior in all recorded error measures, thus confirming hypothesis 3. Since the experimenters we not able to conduct any long term trials with the system many aspects of hypothesis 2 have not been empirically tested. However we can state that over an introductory period the Gem interface was superior in all the key areas of usability. In the author’s opinion there are few grounds to suppose that all the Gem interface’s superiority in error rates would diminish with increased exposure. It is however likely that the Command line system’s command rate would be reduced. Looking at the data from the UEICS2 questionnaires we consistently found error prevention factors among the most important factors related to satisfaction. It is felt there is evidence that the Gem interface’s error rate and subjective superiority would remain even after extensive exposure. There was
evidence of a significant learning effect within the presentation order analysis, however these factors appear to be consistent over both presentation conditions, and can be considered to have cancelled each other out.

10.1.2. Parapsychology

10.1.2.1. Smart Noise And The SSM System.

The overall SSM measure was well within MCE ($Z = +0.909$) and must regarded as giving little empirical support to the smart noise hypothesis. The SSM hit rate difference between the two interfaces was so slight (Gem was 2.097, Command was 2.078, and MCE = 2.00.) that statistical analysis was felt to be unnecessary, since the difference could merely be due to noise from computer based numerical representation errors (rounding errors) in the statistical package. Overall the SSM decision making systems were found to show no evidence of any consistent pattern, except for the displacement and quick coin sources. The consistency in regard to the direction of the latter is so close to MCE that the author regards it as a chance effect (see figures in chapter 9). In conclusion the 3 experiments provided little evidence of any parapsychological effects, and little support for smart noise effects within the experimental situations which were investigated. There are many possible factors which could be held as being responsible for the lack of results, these will be discussed in the final conclusions.

10.1.2.2. The FLP/MLP Hypothesis.

The data from the hitters/missers TAQ2 analysis can be regarded as providing little empirical evidence for many of the paranormal components of the FLP/MLP hypothesis. The overall hit rate correlations with the combined questionnaire data revealed significant trends within the TAQ2 data in unpredicted directions, with 'frequent machine breakdowns', and 'being injured by a machine when young' being linked with enhanced SSM performance. Unfortunately any conclusions drawn from these two correlations would be purely speculative, and probably unjustified because of the large number of analysis performed upon the data. The two UEICS questionnaires showed improved error messages, and on-line help to be significantly correlated with enhanced SSM performance. These are both in the expected direction, but a more consistent overall pattern of effects would be required (as we have found in the HCI data), before these could be taken seriously. However, the non-paranormal
aspects of the FLP/MLP hypothesis proposed by Morris (1985) have received very strong confirmation in the HCI TAQ2 data.

10.1.2.3. The KMDB System a Possible MLP 'Net'.

At this point we should also mention the possibility that MLPs and FLPs could be very rare individuals, and that the one possible case of an MLP which we found in experiment 1 might have been a genuine one. If this is the case then the KMDB system, as it is presently set up might be more suitable as a 'net' for detecting such individuals, rather than as a smart noise development. Certainly the combination of the TAQ questionnaire, and the KMDB system provide what might be regarded as an ideal tool for any future investigations which specifically set out to look for MLPs.

10.1.3. Inter group comparison of the questionnaire data.

10.1.3.1. TAQ2 data.

There were few surprises from the inter group findings. The postgraduate group were more exploratory (made more errors), took longer over their commands (not expected), had more favorable attitudes towards computers, did more computer based reading and were older. There was only slight confirmation of the influence of early exploratory activity with machines effecting gross aspects of later life, such as academic achievement. The 'computer phobics' had received less parental encouragement to explore how machines worked. However there was found to be little difference between the levels of parental encouragement in machine exploration given to subjects from group one and two respectively. This difference might have been more pronounced had we sampled a computer science degree course, as the subjects from group two were in a 'conversion course', and were not necessarily from a computing background. However this result does remain unexpected, and suggests there may be other factors involved in developing positive attitudes towards computers, above and beyond parental encouragement. The postgraduate students had the highest rate of reported machine breakdowns of any of the three groups, but this is probably due to the so called 'tinkering effect'.
10.1.3.2. Post Command Line Session Questionnaire Data

There was little of interest to emerge from these analysis, except the confirmation of the
greater computing experience of the postgraduate subjects. The overall interface behaviour
reported in the TAQ correlations were replicated in the command line system, with subjects
in group one issuing the most commands, most quickly, and the postgraduates issuing
commands the slowest, and making the most errors. This last correlation is taken to be an
indication of that groups increased experimentation, and 'think time'.

10.1.3.3. Post Gem Interface Session Questionnaire Data.

There were few correlations of interest from the Gem post-session differences either. The
subjects of group one found the Gem error correction significantly harder to understand than
subjects from other groups. The command issuing behaviour (rank ordering on number of
commands issued by the three groups) was different on the two interfaces. This was so that
the subjects of group three issued the fewest commands in the Gem interface (switching
their command line ranking with subjects of group 2), while those of group one maintained
their relative position as issuing the most commands. This could be taken as weak evidence
that subjects from group three were more uncomfortable in the Gem interface, than the
command line system. However this would require much more empirical evidence before
it could be accepted.

10.1.3.4. Post-Hoc Sex Differences.

In this analysis the females were found to have made fewer mistakes, perhaps by being more
accurate and careful, but also less exploratory. the traditional sex stereotypes discussed in
chapter 4 were also found. Females were less favorable in their attitudes towards computers,
and more concerned about looking silly if they were seen to make a mistake. These sex
stereotyping were most strongly shown by the fact that females were significantly less likely
to have received encouragement to use dangerous equipment, and explore how machines
worked. On the negative side, and probably connected to their lack of exposure to machines
females were significantly more likely to have been frightened by a machine in their childhood. There was also some (weak) evidence that females found the command line system easier to use.

10.2. Conclusions, Biology Questionnaire, and Experiment 4.

In this section we will discuss the HCI and parapsychological findings from the study, the directions forward, and the final developments from the 3 year investigation. This section will also outline and discuss 2 minor studies which were conducted in the closing stages of the dissertation. One of these was a final attempt to provide evidence for the existence of paranormal FAA using the hypothesis developed from the previous experiments, the so-called 'Function Linked Experiment'. The other is the use by a final year psychology student Ms Lucy Smith, of version 3 of the TAQ in her honours dissertation. This was a HCI study designed by Hamish Macleod, involving correlating the TAQ3 and computer usage details of the first year biology undergraduates at Edinburgh University. The author's involvement in this study was limited to creating the questionnaire, discussing the design of the study with Macleod, and performing various manipulations on the subject's computer usage data. Macleod and Smith were involved with investigating any HCI correlations within this subject population. In return for performing these tasks the author was given access to Smith's data to perform a factor analysis to independently determine the internal validity of the TAQ. This was to see how appropriate a measure it was for looking at computer attitudes and function linkage. This study became known as the Biology Questionnaire Study. This factor analysis provided significant evidence for the empirical reliability of the TAQ in detecting aspects function linkage. A break down of the various factors and their respective reliability coefficients are provided in appendix 3.

---

Morris (personal communication, 1989) has noted that the females apparent problems with the mouse could be due to their spatial ability.

This involved removing any methods of identifying the subjects, and any spurious data from the set before it was handed to Smith and Macleod for combination with the TAQ3 data they had already gathered.

For details see Smith's honours dissertation (1989).
10.2.1. Experiment 4 - The Function Linked Experiment.

10.2.1.1. Introduction.

In the previous chapter we discussed the possible reasons why no evidence could be found of parapsychological phenomena in the SSM system. These can be briefly summarized under the headings of the experimenter’s influence, the subject’s influence, or something inappropriate within the task. There are other possible reasons, such as psi not existing, or it being an acausal force, but the assumption of either of these would have made conducting any of these experiments questionable.

To give the SSM one last try under different conditions the author designed a mini series of informal experiments with a modified version of the Gem KMDB system. This experiment was designed to try and remove or reduce as many of the possible reasons which had been proposed for the lack of paranormal FAA in the previous experiments. These objections and there proposed solutions within this experiment were:

10.2.1.1.1. The Experimenter.

One of the reasons proposed by parapsychologists as to why the previous experiments failed to demonstrate psi have concentrated on the role of some form of experimenter influence. Although the author finds the 'fear of the paranormal' an unhelpful suggestion, there is a possibility that he had an emotional investment in seeing the system perform without any events which could indicate a possible 'bug', or a flaw in the security systems. To remove this possible weakness it was felt to be important to allow some aspect of the study to produce significant FAA, without having compromised either the security, or the integrity of the software system.

Considerable thought was given to this and it was decided to make the study totally informal\(^4\), and to have the implicit assumption that if significant FAA was demonstrated under informal conditions, a follow up study would be conducted which implemented total controls. As we have stated this study was an attempt to see if the dissertation could generate any evidence for significant FAA. The subjects were therefore allowed feedback, and total

\(^4\) Informality also removed the 'classroom' atmosphere, which may have been 'psi' inhibitory.
freedom to issue as many commands as they wished. This allowed any significant FAA to be due to a conscious or unconscious optional stopping artifact. This took any conscious or unconscious pressure off the experimenter and subject to restrict FAA to within baseline chance, since there would be a non-paranormal explanation for any effects which were found. It must be emphasized that the follow up study would have given the subjects a predetermined series of commands to issue, thereby removing the possibility of this artifact. The introduction of this tutorial sheet would be the only difference in the protocols between the informal and formal series, thereby minimizing the possible restricting psychological effect of introducing extra controls.

10.2.1.1.2. The Subjects.

Two main factors were considered as possible 'problems' with the subjects in the previous experiments. They did not consciously know there was a parapsychological element in the study, and they were not likely to possess outstanding FAA abilities. This final study enlisted 12 individuals who had impressive reputations within their working environments for being both 'computer experts', and intensive computer users in their own right. These individuals were Super-FLPS, in the non-paranormal sense, and were therefore felt to be highly likely to have developed paranormal FAA abilities (if they exist). The subjects were drawn from the local computing industry, the 'hacking fraternity' and Edinburgh University. Each subject was told that they were participating in a computer enhancement parapsychology experiment. They also had the workings of the SSM task and operation specified to them in as much detail as they requested. These detailed explanations could only be appropriate with subjects who had an extensive\(^5\) background in computing. The only potential problems with this subject group was that they were from a 'hard science' background, and might therefore have a negative view about the likelihood of the existence of paranormal phenomena. They did however agree to hold an 'open mind' during the session, and were given various suggestions as to possible function linked strategies. Subjects were free to adopt one of these strategies or to use one which they had already developed themselves.

The Possible 'Function Linked' Strategies Suggested to the Subjects.

---

Many of the subjects had international publications in computing and computer science.
Try to take what can be best described as a 'whole' view of your use of the computer, and its functions. Include yourself as being part of a system or 'team' that includes the computer, your actions and its functions. The 'goal' of this system is to perform the allotted task. The task in this experiment is to retrieve the information recorded in the records.

Don't try too hard! (This will be natural if the first suggestion is adopted.) Too much concentrated (focused) mental attention on the physical actions to be carried out (task based activities) seems to reduce the amount of apparent function linkage.

Have confidence that you can produce function linkage with machines. Recall highly successful, and satisfying moments of your past computer use.

Try to hold an open mind as to the possibility of the existence of 'psi'. Whatever your thoughts and feelings towards parapsychology, try to keep an open and positive attitude during the experiment.

Try to forget this is 'an experiment'. Concentrate instead on 'playing' with a new and different piece of software!

10.2.1.1.3. The Task.

The other possibility which the study tried to address was that there was insufficient ambiguity in the previous experiments, since everything possible had been defined, and controlled. As we have discussed, some parapsychologists (Batcheldor, 1979, Von Lucadou 1987) have proposed that some unknowns are needed to allow a psi process to occur. The author refused to remove the experimental controls from the SSM part of the system, since to do so would allow FAA to exist, but outside the monitored domain (see chapter 5). However, the availability of an optional stopping artifact was felt to provide enough ambiguity in the experiment\(^6\) to allow significant FAA to occur (see above).

---

\(^6\) Note that this possible artifact was not included in the briefing given to the subjects.
10.2.1.2. Development Of A Possible Theoretical Background To Function Enhancement Of Computer Systems.

During the process of this dissertation informal models and observations were developed which formed the rationale behind much of the design of the final experiment. They are included in appendix 17.

10.2.1.3. Planned Analysis.

These were as for the parapsychological parts of experiments 1 to 3, and have been outlined in chapter 6. The only additions were the pre and post session questionnaires, which would be analysed by using a non-parametric spearman r correlation (n = 12). The comparisons between this subject group and those of the previous 3 experiments would be done via a non-parametric Kruskal Wallis Anova by ranks. However since these were not under identical experimental conditions, any resulting differences would have to be regarded as purely anecdotal, and have to await empirical validation.

10.2.1.4. Method

10.2.1.4.1. The System, and Equipment.

The equipment used throughout the experiment was an IBM XT286 with a 20MB hard disc, 1 x 1.2Mb floppy, an EGA screen and adapter. The mouse was an optical 3 button 'Summa' mouse. The machine was against a north facing wall, and had an identical machine residing on its left hand side. To the right hand side of the experimental machine there was an Epson LQ800 dot matrix printer. At no time during the sessions was the printer operated.

10.2.1.4.1.1. Note Of Environmental Control.
During each session of these series of experiments the experimenter checked the experimental environment for ultrasonic radiation and other electromagnetic effects which might have affected the performance of the computer. This was achieved by monitoring a series of 3 battery powered silicon (CMOS) based quartz crystal timepieces which were synchronized with a Rolex™ Oyster 'GMT-MASTER'\textsuperscript{7}, before and after each session (the mechanical chronometer had received a full manufacturer’s service 3 months previously). The target computer’s battery based real time clock was included separately in these checks. The quartz timepieces were distributed around the experimental area near the target computer in S35, worn by the experimenter (as a second watch), and resident inside the experimental computer system Parapsi004’s unit case. At no time during any session did the quartz timepieces deviate abnormally from the mechanical chronometer’s synchronized time. The mechanical chronometer was checked against the National Physical Laboratory’s atomic clock at 08:00Hrs (GMT) and 18:00Hrs (GMT) respectively\textsuperscript{8}, every day during the experimental period. During this experiment the mechanical chronometer did not require resetting, and it was not found to deviate by more than (+ or -) 0.5 of a second with the NPL atomic clock over the 9 day experimental period. The quartz timepieces needed resetting twice over the same period, but their time gain (1 second per week) was within their expected performance specification. The experimenter wore the mechanical chronometer continuously during the 9 day period (even during showers, or his early morning run\textsuperscript{9}).

\textsuperscript{7} These are individually tested, and have each been awarded the coveted title of 'officially certified superlative chronometer'. This is the highest award given by the independent Swiss standards institute. These mechanical chronometers are constructed to be virtually impervious to environmental effects (Rolex describe them as being 'virtually indestructible'), and are standard navigational instruments. They have successfully been used as the sole means of navigation on such treks as the British joint-services east-west crossing of the Sahara desert, the trans-globe expedition and the recent recreation of Scott’s polar expedition.

\textsuperscript{8} If an experimental session clashed with the 18:00HRS time check then the next available hourly time check was used.

\textsuperscript{9} These early morning runs were probably responsible for the 0.5 second time deviations which were occasionally noted. Hence the author’s use of the phrase 'virtually impervious to environmental effects' in the footnote above.
10.2.1.4.1.2. Target Machine Circuit Board Control.

The internal boards and main chip numbers were recorded and checked on a daily basis during the experimental period\(^{10}\). There was no trace of any components being changed during the 9 day experimental period. As we have already discussed in chapter 3, if someone had wished to produce an anomaly, replacing machine parts would be an ideal way to do it. To try to reduce the possibility of fraud subjects were asked to give a declaration that they would not try to cheat, and they were never allowed to be with the equipment unattended by the experimenter (during each session). Therefore they had little chance of making alterations to the system during the experiment. The computer’s internal part numbers were checked when no other person was present (usually late at night), so that no other individual knew the details of the security checks. This is important since it reduced the potential possibility of anyone collaborating with a fraudulent subject to produce an apparently anomalous effect.

10.2.1.5. The Questionnaires.

10.2.1.5.1. The Pre-session Questionnaire\(^{11}\)

This covered the areas of the subjects beliefs regarding paranormal influence of equipment functioning. This covered questions such as: whether the subject had ever witnessed an 'inexplicable' form of function linkage, how many hours the subject worked on a computer each day, if they worked consistently with one type of computer system (and if they did what type of system it was), if the subject had ever had a correct hunch or instinct about the source of system problem they were working with, did the subject feel that computing was a lonely or antisocial business, and finally did they ever use a computer system purely for the pleasure of using it.

---

The details of every major component of the system were recorded to be as follows: Power supply number - 40071085700/C, A43458, DATE = 47/86; Hard Disc (type 2 - 20 MB) - 1789701, 600001789701; 1.2 MB floppy - serial number = 1346-0745, Date = 1986 - 42; Motherboard - # = 55w48045585 (62X11568.A43499 V0011), INTEL 80286 = L5270101 INTEL '84; Egacard = 6280539 ECA 33217 V0011; Serial and Parallel Port - W08063591; 6135932 340301, MB 8641 ET PR 34 4B; Hard & Floppy disc controller - B1L41062224, PN = 62X1132 X10 VE330.

A copy of this is reproduced in appendix 10.
10.2.1.5.2. The Technology Attitudes Questionnaire.

The TAQ3 (reproduced in appendix 9) tried to extract the subjects' personal details, their self-rated adjustment to change, self-esteem, kinesthetic sense, Interpersonal skills, stress (new technology based and work based), early attitudes, self-rated frequency of damaging mechanical and electrical devices, working style in new situations, family and immediate social groups' history and attitudes with machines, and the individuals' own attitudes towards machines.

10.2.1.5.3. The Post-Session Questionnaire\textsuperscript{12}

This asked how much the subject had enjoyed their session, if it had changed their views on the likelihood of human influence on computer systems, how often they used a direct manipulation interface, how they viewed such interfaces, whether they used a strategy to enhance their attempted influence, and if they did what the strategy was, and how successful they felt it that strategy was for them. They were also asked to specify any feelings or suggestions they had which might improve upon the strategy, or strategies they used. Finally, they were asked to rate the influence they felt the experimenter had upon their performance.

10.2.1.6. Description of Experimental Method.

12 Subjects agreed to take part, all of whom were selected by the experimenter as being 'Intensive Users' of computer systems. They were told that the study was attempting to investigate the concept of function linkage, in both its normal and parapsychological aspects. All the sessions took place on machine Parapsi004, in room S35 of the Psychology Department, except for one subject who due to a disc reading fault had to fill in his questionnaire in one location and then move to the psychology department for the experimental session. When subjects first arrived, the experimenter tried to make them feel as much 'at home' as possible. All subjects were then briefed before starting the session on a number of topics. First they were told of the nature of the study, and the FLP hypothesis. This explanation included a detailed background of the possible nature of the system.

\textsuperscript{12} A copy of this is reproduced in appendix 11.
enhancements, and what this experiment was investigating with specific reference to the SSM. Next the experimenter detailed the workings of the SSM to the subject in as much detail as they requested. The level of detailed requested by each subject varied with the amount of interest they had in software engineering, or parapsychology. The subjects were then told the way in which the system would display its feedback from the SSM, and exactly what this feedback showed in terms of success rates. Finally they were given the details (see above) of some possible function linked strategies (which they were to free to adopt or reject as they saw fit), and they were told to freely adopt any strategy which they normally found enhanced system performance in their day to day computer use. The subjects then completed the two pre-session questionnaires; a new pre-session questionnaire, and the Technology Attitudes Questionnaire (TAQ3)). These recorded the subject’s views towards the possibility of paranormal influence with machines, their attitudes/background to computing, in addition to the TAQ3. The TAQ3 had been amended to include additional sections on self esteem, stress, reaction to change, and kinesthetic sense (these questionnaires are all provided in the appendix). Subjects were then allowed exactly 15 minutes to investigate and literally 'play' on the system, while trying to attempted as many psi based trials as they wished. The version of the KMDB system used for this study had been amended to give feedback to the subjects at the end of each record request. This was in addition to the temporal feedback (how long the system took to perform the task), which had been used in the previous studies.

After the subjects had completed their strictly timed 15 minute session (during which they received trial by trial feedback of their 'psi' performance, and their final overall 'psi' result), subjects completed a post-session questionnaire. This tried to capture their feelings towards the experiment, their 'function linked' strategy, and the experimenter. Finally, the experimenter answered any questions, or dealt with any reactions which the subject had noticed. The experimenter made contact with all subjects at a period
after the experiment to ensure that they had no remaining problems produced by participating in the experiment\textsuperscript{13}.

\textbf{10.2.1.6.1. Subject Population.}

These were 10 males, and 2 females, whose lowest age was 23, and highest age was 39 (mean = 31.00). The subjects occupations included Systems Analysts, Computer Consultants, Computer Sales Support Personnel, Full-time Researchers, Computer Communications Consultants, and University Lecturers.

\textbf{10.2.1.7. Results.}

Since the subjects were all friends of the experimenter the data was combined and analysed, in such a way that the experimenter was blind to which subject had contributed which questionnaire and session data.

\textbf{10.2.1.7.1. Correlations On In-Session Recorded Data.}

\textbf{10.2.1.7.1.1. Experiment 4 Overall Session Details, Parametric Correlation (n = 12).}

Longer sessions were found to be correlated with issuing more commands (p < 0.053). More errors were found to be correlated with attempting more trials (p < 0.65), and making errors more quickly (p < 0.01). More commands were found to be correlated with attempting more trials (p < 0.01), and issuing commands more quickly (p < 0.05).

\textbf{10.2.1.7.1.2. Using A Non-Parametric Spearman Rank Order Correlation The Following SNS FAA Relationships Were Found.}

A non significant but suggestive negative correlation was found between the number of errors a subject made and the number of psi trials they attempted (p < 0.06). This suggests that the more mistakes a person made the fewer psi trials they attempted. The number of commands issued by each subject was also found to correlate weakly with the number of psi trials they attempted (p < 0.087). There was also a weak correlation found between the number of trials each subject attempted and their mean time between errors (p < 0.094),

\textsuperscript{13} No subject reported any problems associated with having participated in the study.
and their mean time between commands \((p < 0.05)\). This suggests that subject who worked more quickly issued more psi based trials.

**The in-session recording details**\(^{14}\).

**Numbers Of Command Types Issued.**

<table>
<thead>
<tr>
<th>Mean Number of Commands</th>
<th>Wimp Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select</td>
<td>22.25</td>
</tr>
<tr>
<td>Help</td>
<td>0.75</td>
</tr>
<tr>
<td>Modify</td>
<td>0.83</td>
</tr>
<tr>
<td>Add</td>
<td>2.75</td>
</tr>
<tr>
<td>Delete</td>
<td>0.33</td>
</tr>
<tr>
<td>Exit</td>
<td>1.0</td>
</tr>
<tr>
<td>Quit</td>
<td>0.00</td>
</tr>
<tr>
<td>Save</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>27.91</td>
</tr>
</tbody>
</table>

This is more clearly shown in the diagram average number of commands.

![Average Number of Commands](image)

**Figure 122.** Average Number of Commands, Experiment 4.

**Notes:** All time based totals are in seconds. In all cases \(N = 12\). Psi Influence Index has mean chance expectancy (MCE) of 2.0, any higher values than 2.0 indicate a degrading of system performance (worse than MCE), and any values that are lower than 2.0 indicate an enhancement of system performance (better than MCE)
The in-session recorded data (Averages)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Session Length</td>
<td>964.16</td>
</tr>
<tr>
<td>Number of Errors</td>
<td>7.83</td>
</tr>
<tr>
<td>Number of Commands</td>
<td>27.91</td>
</tr>
<tr>
<td>Number of Psi Trials</td>
<td>37.91</td>
</tr>
<tr>
<td>Number of Psi Hits</td>
<td>18.33</td>
</tr>
<tr>
<td>Mean Time Between Errors</td>
<td>265.19 secs</td>
</tr>
<tr>
<td>Mean Time Between Commands</td>
<td>34.54 secs</td>
</tr>
<tr>
<td>Psi Influence Index</td>
<td>2.15</td>
</tr>
</tbody>
</table>

To allow us to contrast these figures let us look at extracts from the data gathered in the previous 3 studies, all of which used an almost identical system.

**Smart noise study (n = 12) Expert computer users.**

mean time between errors = 265.19 seconds mean time between commands = 34.54 seconds
Psi Influence Index = 2.15

**Experiment 1 (n = 46) Trained touch typists (not used a computer before).**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Time Between Errors</td>
<td>237.83</td>
</tr>
<tr>
<td>Mean Time Between Commands</td>
<td>68.13</td>
</tr>
<tr>
<td>Psi Influence Index</td>
<td>2.23</td>
</tr>
</tbody>
</table>

**Experiment 2 (n = 22) Postgrads in systems analysis (used computer & keyboard).**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Time Between Errors</td>
<td>230.43</td>
</tr>
<tr>
<td>Mean Time Between Commands</td>
<td>90.90</td>
</tr>
<tr>
<td>Psi Influence Index</td>
<td>2.28</td>
</tr>
</tbody>
</table>

**Experiment 3 (n = 18) Never used a computer before & not touch typists.**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Time Between Errors</td>
<td>271.18</td>
</tr>
<tr>
<td>Mean Time Between Commands</td>
<td>85.64</td>
</tr>
<tr>
<td>Psi Influence Index</td>
<td>1.69</td>
</tr>
</tbody>
</table>
When a Kruskal Wallis Anova by ranks is applied to this data it is found that the Experiment 4’s inter-command rate is significantly faster (p < 0.01) than that of the other groups.

The graph showing Experiment 4’s inter command times shows a steady decline over the experimental period. This evens out around 25 seconds per command, which would be close to the human performance limit on this system.

The inter-error rates for this group are shown on the graph Time between Errors. This graph is quite peculiar in the way in which the time between errors peaks early in the session and then settles down. This is due to subject number 4 who only made 2 mistakes! One 1145 seconds into the session, the second 211 seconds after that!

If that subjects data is removed then we find that the data is much more uniform and hangs around the 100 second mark.

### 10.2.1.7.2. Error Types.

<table>
<thead>
<tr>
<th>Adjusted figures (divided by N)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntax</td>
<td>5.83</td>
</tr>
<tr>
<td>Semantic</td>
<td>3.0</td>
</tr>
<tr>
<td>System</td>
<td>1.83</td>
</tr>
</tbody>
</table>
Types of Errors

<table>
<thead>
<tr>
<th>Exp 4</th>
<th>Syntax</th>
<th>System</th>
<th>Semantic</th>
</tr>
</thead>
</table>

Figure 125. Types of Errors, Experiment 4.

The most noticeable difference is that the average number of semantic errors is almost half that of the syntax, and that the syntax and semantic totals are much smaller than those of previous studies.

10.2.1.7.3. Questionnaire Data.

Pre-session questionnaire findings (n = 12). Non-parametric Spearman R correlation.

A significant relationship appeared for the subject number and the length of session (p < 0.05). This reflects the length of time the subject took to fill in the 2 pre-session questionnaires. Since the system was started at the moment the subject started the questionnaires, and the actual use of the system was restricted to 15 minutes for each subject. This finding indicated that the later subjects took longer to complete their questionnaires. The subject number also correlated with the number of commands issued (p < 0.058), such that later subjects issued more commands\(^\text{15}\).

The length of time the subject used a computer each day correlated with how they viewed computing. Those who worked longer on the machine tended to disagree more strongly with the statement that computing can be a lonely and antisocial business (p < 0.082). Possibly they used electronic communication during their working day.

\(^{15}\) These effects could have been due to the experimenter becoming more proficient at running the experiment.
Those subjects who worked consistently with only one system made more errors \((p < 0.056)\), than those subjects who used a variety of systems, and they seemed to make those errors more rapidly \((p < 0.085)\).

Subjects who often used computers 'just for the enjoyment or 'feel' of it', made more errors \((p < 0.05)\) than those who did not, and they made those errors more quickly \((p < 0.05)\). These same subjects who often used computers for the enjoyment of the system, also seemed to have a greater tendency to 'miss' (slow down the system) in the psi element of the recordings \((p < 0.08)\); but this was not significant.

**Post Session Questionnaire Findings. Non-Parametric Spearman R Correlation \((n = 12)\).**

The subject number (which is the order with which the subjects attended the experiment) correlated with the influence subjects reported the experimenter had made on their performance \((p < 0.08)\). This was such that later subject felt that the experimenter had made a more favorable influence than had earlier subjects. This is probably due to the practice the experimenter gained at running this experiment. The degree to which the subjects enjoyed the experimental session correlated with how much they felt their attitude on the likelihood of 'function linkage' had been changed by the experiment \((p < 0.05)\). This effect was such that the less the subject felt their attitude had changed the more they enjoyed the session.

Those subjects who enjoyed the session most had a weak tendency to be those who had adopted a strategy to enhance their 'psi' performance during the session \((p < 0.06)\). Enjoyment of the session also correlated with how successful the subject felt their strategy had been \((p < 0.05)\). This was so the less successful they felt the strategy had been in the experiment the more they enjoyed the session. This was also confirmed by a weak correlation that occurred between the subject's enjoyment of the session and the degree to which the subjects 'psi' performance went in the missing direction \((p < 0.08)\). These findings would seem to indicate that these subjects were not trying to produce FLA.

The more frequently a subject used a graphical or mouse based system the more useful they regarded such systems \((p < 0.05)\). While those subjects who seldom used graphical or mouse based systems took longer to complete their pre-session questionnaires \((p < 0.05)\). However
those people who viewed graphical interfaces as being useful issued fewer commands than those subject who did not view such interfaces as being useful (p < 0.056).

Those subject who viewed their strategy to enhance the machines performance as being successful seemed to attempt fewer 'psi' based trials (p < 0.08). There was no correlation between the subjects feeling of success in their strategy and their actual 'psi' based performance (p = 1.0 - totally random).

10.2.1.7.4. Technology Attitudes Questionnaire (Version 3).

There was a large amount of data generated by the TAQ. Due to space restrictions it is not possible to present all of it, so the reader must understand that the following is a summary of the major findings.

Non-parametric Spearman R Correlation (n = 12).

10.2.1.7.4.1. Machine Breakdowns.

The data from the questionnaire that referred to breakdowns revelled the following patterns:

**Situational or Physical.**

Those subjects who reported frequent problems with computers tended to own a computer at home (p < 0.06). Subjects who reported problems with computers also reported more frequent breakdowns with their mechanical equipment (p < 0.05), and with their cameras and film (p < 0.06).

**Biological**

Those subjects who were right handed had fewer problems with machines breaking down (p < 0.05). It also appears that younger subjects had more machine breakdowns (p < 0.05). Left handed subjects reported having more breakdowns with their personal computers (p < 0.052).
Psychological

Subjects who experienced problems with machines breaking down tended to get depressed when they were criticized (p < 0.05). As did people who avoided social situations where there was the chance for criticism or rejection (p < 0.05).

Subjects who avoided social situations where there was the chance for criticism or rejection liked working with equipment (p < 0.05), but admitted to being bothered by the impersonal nature of computers (p < 0.05). This group tended to disagree that computers were too complicated for them to understand (p < 0.05), wanted to be good at whatever they did (p < 0.05), and had a tendency to feel that humans were more reliable than computers (p < 0.08). Subjects who reported having problems with machines also reported being more afraid of looking silly if they were seen not to know how to use the computers (p < 0.05). Those subjects who reported having breakdowns with personal computers felt that computers were too complicated for them to understand (p < 0.05), even though they acknowledged that computers could help them to work better (p < 0.05). Subjects who reported having problems with their personal electrical appliances often felt they could not do as well as others (p < 0.074). However problems with machines did not seem to affect the enjoyment of life, since subjects with problems with their computers tended to be happy with their life (p < 0.09).

Background

Those who were encouraged to explore how machines worked when they were children had few breakdowns with equipment (p < 0.05). Those whose immediate family’s preferred occupations involved new technology reported having fewer problems with their personal computer (p < 0.05), or their motor car (p < 0.074).

Other Findings.

Stress due to fear.

Those subjects who reported that they found it hard to relax around new technology also reported having a fear of damaging equipment if they used it (p < 0.01).
Subjects who reported being so pressured in their work that they could not cope, tended to have a slow rate of adjustment to change (p < 0.05), and found it hard to relax around new technology (p < 0.05). These same subjects reported being uncertain how they could cope with new or uncertain situations (p < 0.01).

**Early Experiences, A Major Factor In Later Life.**

Subjects who reported that they helped their family repair equipment when they were children developed an enjoyment of figuring out how things work (p < 0.05), and tended to feel that computers were fun to use (p < 0.05). While those subjects whose family repaired broken equipment promptly when they were children tended to feel that computers were not too complicated for them to understand to understand (p < 0.07). Those subjects who as children, had been encouraged to explore how machines worked felt that computers were fun to use (p < 0.05).

Subjects who had parents who repaired equipment themselves tended to be sure they could cope with new or uncertain situations (p < 0.01), and they seldom got so pressured with their work that they could not cope (p < 0.01).

Subjects whose parents let them help repair broken equipment tended to be unperturbed by social situations where there was a chance of rejection (p < 0.05).

**10.2.1.7.4.2. Correlations With In-Session Recorded Data Non-Parametric Spearman R Correlations (n = 12).**

**Errors**

Subjects who reported being from homes where machines broke down frequently made fewer errors (p < 0.05). Subjects who felt that computers were too complicated to understand made more errors on the system (p < 0.05). Those subjects who reported being afraid of looking silly when they made a mistake while using the computer made more errors (p < 0.062).
Subjects who read books about computers also tended to make more errors (p < 0.05), and those subjects who reported being more satisfied with their life made more errors (p < 0.054).

Those subjects who felt that computers were not too complicated for them to understand had a slower inter-error rate (p < 0.056). This tendency to make errors more slowly was shared by those subjects who reported feeling that humans were more dependable than machines (p < 0.05), and by those subjects who were not afraid of looking silly if they made a mistake while they used the computer (p < 0.065).

**Commands**

Those subjects who reported being less afraid of looking silly if they made a mistake while using the computer attempted more 'psi' trials (p < 0.05). Closely linked to this was the correlation between subject who reported being less afraid if they were seen not to know how to use the computer seemed to attempt more psi trials (p < 0.065). Unlike the trials in the first series of experiments these were not covert.

Males tended to issue more commands than females (p < 0.063), and to issue their commands more quickly (p < 0.082). The number of commands correlated the use of a mainframe, such that those subjects who reported more frequently using a mainframe issued fewer commands (p < 0.05). Subjects who reported that they were not afraid of damaging the computer if they used it issued their commands faster (p < 0.05), as did subjects who reported being unafraid of looking silly if they made a mistake while they used the computer (p < 0.068). The same trend of issuing commands more quickly was also found to apply to those subjects who were not afraid of looking silly if they were seen not to know how the computer worked (p < 0.05).

**Function Linkage.**

Those subjects who frequently used a micro at home tended to psi miss (degrade system performance) (p < 0.01), as did those who used a micro at work (p < 0.05), and those who more frequently read books about computers (p < 0.05). Those subjects who reported being more satisfied with their lives also tended to degrade the SSM (psi miss) (p < 0.05).
10.2.1.8. Parapsychological Data.

As far as the psi data is concerned the experiment was a great disappointment. It revealed no evidence of psi, under what could be regarded as good test conditions for function linkage, and the SNS. The subject pool was composed of individuals who would be most likely to possess some form of function linkage, especially if one considers the position that most subjects had attained in their chosen profession.

The graph of psi trials shows a random pattern of hitting and missing, with an overall average hit rate of 2.15 (MCE=2.0) (p < -0.3).

The histogram showing the in-session SSM decision sources shows no obvious pattern. The only point of interest was that since the expert users did not repeat their commands the displacement source showed significant missing for the first time. This was taken as some confirmation that the effects seen previously in that source were in fact due to the artifact, and not any paranormal action.
10.2.1.9. Parapsychological Conclusions For Experiment 4.

These results were very disappointing from a parapsychological point of view. It had been hoped that this experiment would demonstrate some active use of the SNS. Based upon the parapsychological literature we could postulate some reasons for this finding, such as a fear of confronting an unknown force made the subjects 'disable' their psi (most were established in the 'current scientific paradigm'). The subjects could also have been forcing the system to produce a random result by using the potential optional stopping artifact within the design, although the author feels this is unlikely. Alternatively it could have been the experimenters fault, due to an unconscious fear of 'psi' on his behalf. Unfortunately, as we have already discussed, these sorts of concepts can never be falsified, so we will concentrate on more pragmatic possibilities. We have to consider the possibility that such machine influences cannot occur in the situations the experiment provided. It could be that 'psi' is not an intelligent effect (as would be required to conduct a search), but rather a blind force produced by an organism to act upon some aspect of its environment. Such a force may not be able to be controlled or to be made to react in precise, and intelligent ways. This explanation would explain the lack of success with the SNS, and would perhaps expect to find consistent deviations in the outputs of the REGs used in this study over all the subject groups. Unfortunately there is no empirical support for this hypothesis from these experiments. Finally we have to mention the possibility that the experiment was successful in removing artifacts, which might otherwise have masqueraded as an unknown form of influence, and have misled the experimenter into assuming that some form of paranormal action had occurred.

10.3. Overall Conclusions, And Possible Directions Of Future Research.

10.3.1. The HCI aspects of these studies and the possible future developments from this research.

The findings from this research cast series doubt on the 'lore' within 'heavy computing' user circles that:

WIMPs are for people with no brains, or who don’t know better (Comment made by an intensive computer user during an interview).
This is an important and intriguing point. Why do the more serious computer users have such a low regard for WIMP systems? Or is it that the anti-Wimp groups are more vocal, or in positions of greater perceived authority within computing? The apparent preference among intensive computer users for highly complex and esoteric command languages such as those found in Unix™ must have a rationale behind it. The author admits that before he conducted this investigation, he preferred to use such complex command line systems himself, but on the basis of his findings would now opt for a functionally identical Wimp interface. Part of the answer to the Wimp/Command line preferences may lie in the fact that very few Wimp alternatives provide a truly functionally equivalent system. Most of the available Wimp alternatives to the major commercial command line systems provide only a 'beginner's subset' of the commands available to the command line user. This is certainly the case with many Unix™ windowing systems, and with Wimp alternatives to MSDOS™, such as Gem™ and Windows™.

The first series of experiments provide some empirical evidence that a functionally identical Wimp interface is superior to it's command line alternative. As a result of this study, it can be argued that the pertinent HCI based question is no longer, which is better, a command line or Wimp interface. It is now why should intensive computer users prefer to use a less efficient method of controlling a system? The answer to this question would enable interface designers to make systems which were more attractive to this influential user community. The author suspects that part of the attraction in a system which is hard to master and use could be that it gives satisfaction at being able to control such a complex system. Alternatively it could be that it provides some form of security, or ego satisfaction. It may well be that in the same way that some groups prefer highly powerful and complicated audio equipment, or cameras, so some groups will be found to prefer more complex methods of using computer systems. If this is the case then manufacturers will have to decide the best balance between efficiency and attracting this particular user community. However the author considers that the case for it being 'a matter of taste' is far from proven, and it may well be that there are other factors at work.

The logical follow up to the HCI parts of the study described in this dissertation can be split into several separate investigations, although obviously there would be considerable overlap between them. The answers from well conducted studies into each of these areas would be

\[\text{This would be limited to within the task domain, and it should be understood that these findings may have limited applicability to other more naturalistic task domains.}\]
of considerable interest to manufacturers and suppliers of high performance computing systems. Indeed on the basis of the research described in this dissertation the author has been fortunate enough to have been given the opportunity to develop his computer interface research interests under the generous funding of the international computer manufacturer Digital Equipment Corporation. This will be in the interface design centre of DEC’s European external research programme at Portsmouth Polytechnic. This research will involve investigating the HCI aspects of integrated graphical software development tools. Since this research is being funded by a commercial body the author is restricted (by the commercial considerations involved in the sponsorship) in what he can discuss about his planned research. However the following discussion gives some of the areas of HCI research which are logical developments from the work described in this dissertation.

10.3.1.1. The effects of prolonged use of functionally identical interfaces.

This study would investigate the effects of prolonged use of the two different interfaces on different subject groups, computer naive, keyboard literate, and experts, in terms of direct in-session performance, and subjective measures similar to those used in the UEICS questionnaires. To make this project viable the software used would have to support the full range of functions which any reasonably prolonged use of that application would require. The KMDB system supported only a restricted range of functions (since competence had to be within one hour). However if an experiment demanded that subjects use an interface for extended periods (up to 6 months) it would need to be a full commercial implementation. It would also be essential that the two interfaces were functionally identical. Unless a suitable target system existed its construction would involve a considerable amount of effort. Fortunately DECwindows™ provides such a suitable environment, and can be contrasted to the VMS™ or ULTRIX™ command languages.

---

This is an advanced windowing interface based upon the X Window system developed under project Athena at the Massachusetts Institute of Technology. The X Window system and Microsoft’s Windows™ have been accepted as the OSF (open software foundation) standard interfaces for future computer systems (Manchester, 1989).
10.3.1.2. The effects of varying durations between use of the two interface
types to see the effects on performance, and preference.

This project would ideally form part of the first study (above), using the same subjects, who
could be randomly assigned to groups who would be asked to re-use the interface after set
periods of time had elapsed. These time periods could be varied over the groups to evaluate
the scale of any temporal effects on performance and preference. Using the same subjects
who had participated in the first study would enable measures of their decline in performance
to be compared to their learning curve. This would allow an accurate measure of any
temporal performance, and preference declines associated with the interfaces.

10.3.1.3. Compare intensive user’s preferences towards the two
alternative interaction methods.

This would be a longitudinal study, using computer science degree students as they
progressed from their first to fourth years of study and investigating how their preferences
with regard the two interaction methods changed over time. This study would investigate if
there was a considerable influence from the ‘culture’ of a computer science degree. The
initial questionnaires should try to isolate background factors which might bias the subject’s
preference, and then record their current feelings towards the two interfaces, asking them to
state why they have these preferences. At regular periods throughout the four year period
(perhaps 3 monthly) a brief questionnaire could record any changes in their feelings. Ideally
these results could simultaneously be compared to subject preferences in a non-computing
4 year degree, to try and control for any more widespread trends within the computing
market. It would be impractical to compare performance on two interfaces over the 4 year
period since the average ‘life’ of a commercial product is 6 months, and the 4 year old
interface would produce too many complicating factors in a preference evaluation.\footnote{17}

\footnote{17} This 6 month product life was also why the 6 month time period was specified in the first study.
10.3.1.4. Construction of a specific questionnaire to identify the relevant factors.

An assumed part of 3 (above) is the enhancement of the TAQ, and other evaluation questionnaires. The TAQ could also be developed as a predictive tool for computing use (or computing job selection). This development of a TAQ could be made quite rapid by producing an automated system in which when a student user first logs on to a system they have to complete a version of the TAQ. The subject’s computer usage would then be monitored and at the end of a pre-specified period the TAQ and computer usage data could be correlated to look for any predictive measures. This would allow relatively rapid empirical development of the TAQ into a useful predictive tool. If this process was spread over various faculties then sub-versions of the TAQ could be made highly specific to different vocations, or uses of the computer. An example of a simplified version of this idea was conducted in the Biology Questionnaire study. This study was conducted by a final year honors student under the supervision of Hamish Macleod, using version 3 of the TAQ developed in this dissertation. It was briefly described at the start of this chapter. Even if this method was not adopted the author still feels that the TAQ should be developed further, of particular interest is the area of kinesthetic sense. The author feels that there may be a link between ability with machines (or new technology), and physical dexterity. Obviously any further work with the TAQ should include a detailed investigation of the various factors within the TAQ data we have so far gathered. The author has made a start on this process (see appendix 3), but much more work needs to be done before the TAQ can really be used to its best effect.

10.3.1.5. Sociological investigation of the values of 'technocrat' and 'technophobe' groups.

This would involve the long term investigation of social groups of individuals who are either intensively involved in computing or who are intensively involved in avoiding technology. This study would attempt to isolate the factor(s) which form the shared group values at the core of each of these two groups. The results from such a study would largely depend upon the rationale behind conducting it, and could at least provide some information regarding the differences between the two groups. The author must however note a degree of discomfort in proposing this study, since both groups should have the right to maintain their
views. One ethical danger from such a study would be that the experimenter might attempt to use the information gained to change the 'technophobes' into computer users. There are some obvious ethical problems involved in such research, but they are beyond the scope of this discussion.

10.3.1.6. Investigation of the value of the various aspects of DM which are most error prone (in semantic and syntactic forms), and which produce the most productivity and satisfaction.

This research has established that Wimp interfaces are superior to their equivalent command line systems, within a simple application. Having established this, one obvious follow up (and the one the author will actually be involved in performing) is to investigate which aspects of the Wimp interface are more suitable for different tasks, and user groups. This can be done by varying the properties of the Wimp interface, such as the menus (pop up, or pull down), icons (the size, colour, shape, and representation), windows (size, and number), and mouse movements (sensitivity, movement and speed). These are only some of the aspects which could be varied. By having detailed recording processes within the interface the various aspects of the system (beyond the user interface) can also be evaluated and improved. It may be that applications held within windows which have not been accessed recently can be suspended, or iconised (thereby automatically tidying screen space). Alternatively the movements of the mouse towards certain applications can be taken as indications to activate those processes which lie in the intended direction of the mouse. This would be at the expense of those which might be found to be less likely to be selected next. The findings from such research might result in dramatic improvements in both system/user performance, and subjective satisfaction. Obviously such studies would involve considerable expertise, and the use of highly expensive computing resources.

10.3.1.7. Other beneficial uses of in-session recording systems.

One of the unintended benefits from the study described in this dissertation was the finding that the KMDB system identified subject/students who were having particular problems with their coursework, not just with computing content. Further the individual session logs and the questionnaires provided a detailed picture of the student which could be of considerable benefit in a teaching environment. One possible development of this research would be
within the area of computer aided learning (CAL) packages. By monitoring the in-session
details of a subject’s responses to the teaching package the system could identify those areas
in which the subject had most difficulty. This can be by the number of errors, or by the time
delays involved in answering the questions. The system can then configure it’s teaching
content to that student’s areas of difficulty. This would reduce the potential for wasted CAL
access teaching time (which is often highly restricted anyway), by concentrating only on the
areas which the student needed most. This also raises the possibility of the use of in-session
monitored computer counselling, where the patient’s sensitive areas may well be detectable
by the time response differences made on questions from a computerised interview. This
information could then be provided to the counselor and reduce the time involved in
providing effective help.

10.3.2. Future Parapsychological AHCI Studies.

The final section will briefly outline some more potential studies which could investigate
the area of AHCI\(^\text{18}\).

10.3.2.1. Use of live RNGs.

The study did not have the opportunity to use live random number generators as sources to
drive the SSM. The author feels it is unlikely that the use of such sources would have made
any difference to the results. However if one was being absolutely thorough in one’s
evaluation of the concept of smart noise then the use of electrical and atomic sources of
noise would be a valid follow up to these studies.

10.3.2.2. Use of biological systems as SNS (the ‘artificial nerve’ project).

One of the methods of noise generation which the author would have liked to investigate
was that of using a biological system as a source of randomness. The author would propose
such a source as being made up of some form of artificial electro-chemical (neural)
connection, the randomness in such a system being provided by the very slight variation in
conduction speeds, and electrical current through the connection.

\(^\text{18}\) A number of studies and potential applications were also given in chapter 5.
10.3.2.3. **Use of neural networks as part of SNS**

Closely tied to the 'artificial nerve' project (above) would be the use of smart noise to vary the signal connectivity strength within a neural network. Such a smart noise system could be tested for pattern recognition tasks against a normal neural network to get a control comparison. Such a system would obviously be of great value in such tasks as voice recognition, where the user would be trying to enhance the system's ability to recognise their voice to achieve some task objective.

10.3.2.4. **Look at mechanical breakdowns**

One of the most obvious areas of AHCI is that of machine breakdowns. This could be done by either looking at return rates\(^\text{19}\) (which is not recommended), or by constructing a test product with a normally distributed failure rate. The problems with these kinds of investigations are that no commercial company will want to be associated with such a study, since it implies some unreliability in their product. This is why manufacturing a product with a normally distributed failure rate might be the only method of doing this. One slight adaptation of this could be a computer simulation of the construction of these products. At the start of the experiment each simulated product could be assigned an operational life by an RNG. The assignment of these identical products could then be determined by subjects (who are blind to the RNG decisions) within some sort of computer game. Once all the products have been allocated between the subjects the computer could produce the failure rates for each subject. This could allow a cheap method of evaluating the feasibility of such a project before actually going to the expense of construction. This would be trivial to implement.

---

\(^{19}\) There are too many problems with investigating the return rates of products for it to be anything but anecdotal evidence of malfunction linkage. Examples of these complicating factors would be low income, types of use, origin of the basic components, the user's product expectation, environmental factors, the distribution of production runs, and the variables involved in the product's storage between manufacture and resale. Just to take low income as an example, poor quality mains supplies cause components to burn out more rapidly, and low quality products would break more frequently. The variations in demand and usage would make this a nightmare of uncontrollable variables. It is not recommended.
10.3.2.5. Computer Breakdowns.

It is possible to monitor computer 'down' times to see if there is any validity to the lore within academic circles that machines crash whenever the psychological pressures on the user is high. Again the problem with this is that computer managers may be reluctant to admit that their machines ever fail. This has certainly been the case whenever the author attempted to gain any data on the university's computer down time.

Concluding Summary.

Although the author's overall conclusions (above) were that these experiments provided little empirical support for either the smart noise, or FLP/MLP hypotheses, there are still large areas within each domain which have not been fully investigated. The area of AHCI is such a new and exciting experimental area for parapsychology to explore that it would be premature to regard the findings of these small number of experiments as anything but exploratory. It may be that the use of such a complex series of logical operations between the source and decision had some negative effect on the psi process. It would therefore be logical to investigate if the psi effect size found in these experiments was increased or decreased if a much simpler decision making system was used. This would involve the removal of the SSM, and investigating if one simple RNG was a better source of decision making.

Further investigations of the FLP/MLP categories should perhaps look at the effect of belief in both subjects, and experimenters. The KMDB system for the final series of experiments was designed so that only the computer 'knew' the decryption keys for the data. This was specifically designed to allow the investigation of multiple 'experimenters', each of whom could be selected on the basis of their degrees of belief in psi. With regard to subjects it would be valid to investigate the effect of introducing very strong psi believers and skeptics into a KMDB experimental series.

However the author would be unprofessional if he did not state one final word of caution to future experimenters in this area. The amount of effort and resources involved in conducting controlled experiments in AHCI is such that future research would be well advised to be done in collaboration with other research projects (as was done in this study). This would ensure that null results do not waste the field's limited resources. Fortunately HCI would
seem to be only one of many potential areas of overlap which AHCI should be able to develop with mainstream science. The introduction of the computer into most academic disciplines should be seen by parapsychologists who are interested in the area of human machine anomalies as a golden opportunity to investigate new and exciting areas of human machine interaction.
Anomalous Human-computer Interaction (A.H.C.I): Towards an understanding of what constitutes an anomaly (or how to make friends and influence computers).

K. Morgan.

Department of Psychology, Edinburgh University.

With thanks to Prof. R.L. Morris, Dr. H. A. Macleod, Dr. J. Milton and Mrs. H. H. Sims, for their help and advice.

ABSTRACT.

The following is an attempt to clarify in what manner a genuine anomaly can be distinguished from an incident explicable by known physical means. It will also try to exemplify the various methods that could be used to simulate an Anomalous Human-Computer Interaction (A.H.C.I.). The paper will split "anomalies", and their possible known explanations into distinct groups. In many instances the known explanation will overlap into more than one group. The paper does not dwell in any more than a superficial manner upon the "psychology" involved in manipulating observers, to allow the described physical strategies to be carried out. This is a complementary area of study and would therefore demand a paper in its own right.

Introduction.

Part of the research being carried out at the Koestler Chair and other institutions is the investigation of Anomalous Human-Computer interaction. As with any area of parapsychological research there always exists the danger of the researcher mistaking a normally explicable phenomenon as an anomaly. This paper was written to try and help people who are confronted by an unusual happening on a computer to evaluate the situation, and to be aware of the possibility of there being normal methods of simulating the anomaly. Throughout this paper the following terms will apply, the term "anomalist" will refer to a person linked with an anomaly, this will have no presupposition as to the nature of the anomaly, be it fraudulently coordinated or not. The term "false anomalist" will be someone who produces or is associated with an anomaly which has a known explanation. Finally the term a "true anomalist" will refer to someone associated with an anomaly that does not appear to have a currently known orthodox explanation (Morris 1984). The paper will attempt to define the categories of A.H.C.I. phenomena that might be used to convince a computer user that an anomaly had taken place. Although each example will be given at the conceptual level, I will also try to give an illustration of a normal explanation in each of the categories.

1 In order to save space the figures and references have been removed from this paper.
1. HUMAN.

For the purposes of this paper the human element in an A.H.C.I. experience can be be split into those who
are influencing (whether willingly or not), and those who would like to be thought to be influencing the
equipment. It could be that an individual might be prepared to go to an extreme length to convince the
system user of the genuine nature of the anomaly, due to an attempt to increase the false anomalist’s
self-esteem, the propagation of the false anomalist’s belief system, or merely a financial reward. The
“tools” of the human element in a false “anomaly” could involve distraction ("Ilm-flam"), breaches in
computer/organizational security (prearranged), and breaches in computer/organizational security (sim-
ultaneously with anomaly), all of which will be elucidated in greater depth further on in this paper. Having
mentioned fraudulently created anomalies, it must also be emphasized that the majority of so called
“anomalies” probably fall in to the area of ignorance by the user of their own computer system. As a
general rule there are very few systems that don’t have some undisclosed feature or “bug”, the only factor
being the regularity and inconvenience of the “bug”. It is also true of humans (and that includes most users,
yes even computer scientists) that unless a feature is drawn to their attention, or is inconvenient, it tends
not to be noticed until it becomes of importance, or is drawn to their attention. In a great many cases the
system will always have had the potential for producing what will become an anomaly. The only missing
factor is the human capacity for “forcing” unconnected happenings into meaningful patterns. (Perhaps
there is room here for a synchronistic theory of A.H.C.I.) When a user’s micro picks up phantom messages
from beyond the grave (and beyond most people’s “boggle” point), the possibility of induced electrical
“noise” or stray computer telephone messages is sadly ignored. Alternative explanations are often very
obscure, and seemingly unconnected to what would otherwise have to be considered an “anomaly”. For
example the regular 5:30pm system crash that occurs every Tuesday is seldom connected to the heavy
goods van that passes by the computer building every week at that time, or to the radio controlled model
aircraft that always seems to be flying on a Tuesday. Instead users tend to find patterns within their mood
states, and inter-personal relationships that they assume have causality for an already existing, or explicable
“anomaly”.

2. SOFTWARE ANOMALY.

If a false anomalist required to make the target computer misbehave by affecting the software inside the
system, this could be achieved by an internal act of “manipulation”. Simulating such an anomaly in this
manner could be done by one of the following methods.

2.1 Replacement.

In this situation the false anomalist would gain a copy of the target program code and replace it with their
own “amended” version. This new version would produce the kind of results that the false anomalist
required. The whole process could be achieved by an infiltrating autonomous program that was “resident
in some part of the system. Such infiltrating programs can gain access to a system by being “squirited
down an external phone line by a remote user (such as the illegal hacker who logs on to the system in the
early hours of the morning, when security is low). With the move towards modular programs in
multi-tasking environments these invading programs can shuffle the code above and below the area of
memory they “want” to occupy, and can then “sit” undetected in memory waiting to spring into action
upon some predetermined event. After the intruding process has achieved its objectives it can be
programmed to delete all traces of its presence in the system, leaving an anomaly which is extremely hard
to explain.

2.2 Amending the target program.

The use of interpreted languages makes this easier, but it is also possible to perform upon precompiled
object code, using an on line debugging routine. This has the beauty of changing the executable file, while
leaving it with the same date/time stamp and size of file. Often this can cause even “experts” to be mystified
as to the cause of a fully tested programs failure. The only control against this is the use of a program to
compare the executable code with another working copy of the same code, byte by byte, listing all the
differences that are found. Another test is to copy the program that seems affected, and to try to run it on
another identical machine. In anomalous situations it would be ideal to have a means of recording all the
actions carried out on the target machine, in a manner that would allow a reconstruction of events under
control conditions. However even the discovery of a difference in the code of the misbehaving process
does not mean that the code has been changed fraudulently. The problem of how to distinguish “Psi” corrupted code from that changed by an editor or a malfunction has not yet been solved; however the final section of this paper will propose some methods which could be adopted to reduce the risk of fraudulently changed code being accepted as having been genuinely produced by an anomalous means.

2.3 Breaches in computer/organizational security - prearranged.

In the “internal” breach of security scenario the infiltration of the system would be in terms of a piece of software entering the computer system, and causing the anomaly. The “prearrangement” factor would be that the program that causes the anomaly would already be resident in the host computer system before the anomaly. If the false anomalist suspected that system security would be tighter before and during the planned time of the anomaly the code would have to “hide” in the target system for what might be some considerable time. The infiltrating program, once having gained access to the host, can hide from detection in a number of ways. All computer systems have a method of creating “invisible” files. These are usually used exclusively by the operating system and are unknown to the normal user. In a large system the invading program would simply add itself to the system clock queue, under a sound-alike system reserved file name (which avoids “nasty” data security personnel from spotting, and querying it). Then it would just write itself to the system disc, and wait to be awakened at the appropriate time by the computer’s clock queue. Once awakened it would probably be able to cause its prearranged anomaly, and delete all trace of its existence within a few millionths of a second. During its brief reawakening from backing storage, the invading process could hide among other programs very easily; for example if the process called itself an impressive enough name (which would be any name which contains the initials of the computer’s operating system, or the name of the system manager), it would be highly unlikely to be investigated by even the most sophisticated security program! Systems are so large and complex now that no one person knows all the parts of an operating system, and no security process is going to delete what might be part of the operating system.

2.4 Breaches in computer/organizational security simultaneously with anomaly.

In this scenario the false anomalist would break through system security at the time of the anomaly. This could be done by using an accomplice at a remote site, or by getting a remote computer to automatically break into the system at the prearranged time. The details would then be identical with the previous description once the infiltrating process has gained access. The only disadvantage with this method is the risk of detection during the log in process from the remote site.

3. HARDWARE ANOMALY.

These would necessitate the physical presence of the individual, or accomplice on the site of the computer equipment. In a traditional academic environment it is unfortunately very hard to keep very high standards of security. In a non-academic environment it should in theory be much easier to maintain tight security; however, a false anomalist need not lose heart, since it still might be quite easy to gain access. The false anomalist might have to get an attractive accomplice, but to someone who would be going to demonstrate “miracles” that should be a minor problem.

3.1 Replacement.

In principle the exchange of computing components would seem to pose very large problems in the way of a potential false anomalist. The serial numbers printed on to most computer units would seem a very good way of ensuring that they are not easily replaced, at least not without leaving a trace. In practice however there are very few users who check and record the serial numbers of their computer’s internal boards or chips, let alone think of checking if they had been exchanged when an anomaly occurs.

3.1.1 Peripherals.

An identical peripheral is prepared and is used to replace the original, before the actual test begins. The special effect may involve the printing of some special characters on a printer at some prearranged time. To achieve the effect of disrupting equipment permanently is much easier than achieving a temporary effect, since the latter will force the re-exchange of peripherals after the anomaly.
3.1.2 CPU/Internal boards.

In this scenario the change of internal boards could be to make the machine transmit all its actions to a remote site, or to simply make it crash. The replacement of a board or chip would require the dismantling of the target computer system. It might be necessary for a false anomalist to get an accomplice to come in to the building as a repair/maintenance engineer and effect the change that way. This would require considerable effort, risk, and expense. It would therefore be unlikely to be rated as being worthwhile, unless there was no alternative, and even then it might be better for a false anomalist to consider having a “psi missing” day.

3.1.3 Backing storage.

3.1.3.1 Discs.

In this case an identical floppy disc would be prepared as a copy of the target, and exchanged for the proposed target, when a suitable distraction arrived. This switch would enable new files to appear, or existing files to have their contents “amended”. Encryption of files would be an important guard against this possibility.

3.1.3.2 Printouts.

As for discs above, but with the added virtue, to a false anomalist, of being much more easily acceptable as evidence in a British court of law. (The British legal system is still grappling with the concept of what constitutes “evidence”, when it comes to information held or used by a computer.)

3.1.4 Communications.

The list of “tricks” that might be used by an unauthorized/undesired user in order to gain access to a system are numerous. These tricks usually fall in to one of two categories.

3.1.4.1. Impersonation.

This method requires learning or acquiring some privileged user’s identification. A good example of this is the “Trojan horse” method where the hacker holds on to an external line after having been refused access. The hacker then waits and “simulates” the target computer system's “logging in routine” to the next user who uses that line. In this manner the hacker gains that user’s identity and logs in to the system. In contrast the “backdoor” method involves the hacker knowing the privileged password used by the hardware/software engineers when they have to gain access to a machine, and using that password to gain a very high priority within the system. This makes the user potentially “above” any defenses the system may impose on users, such as invisible files and resident security monitoring processes. In fact it has been known for remote infiltrating users to set the system against legitimate users when they have been discovered! In such cases the only response is to close the whole system down. The scenario of a remote user/hacker pretending to be a privileged user is avoided by the use of fixed routines, such as one in which the computer phones the remote user back, after an initial attempt to log into the system has been accepted, but only to a known phone/modem number. This avoids the problem of impersonation, since the remote user must not only know the correct password, but also must be physically located at the correct remote location.

3.1.4.2. Penetration of system.

This involves an in depth working knowledge of the system, or using sheer power. In general it involves changing your own priority, or outsmarting poorly designed system defenses. An example would be where a potential infiltrator tries, and fails to guess a legitimate password. The hacker then asks the system for help in how to enter a password, and is shown an example which actually works. An example of sheer power would be a program which tried every password possible in succession. In general the more helpful, and user friendly a system is the less secure the system becomes.

3.2 Adjustment/destruction - “live”.

These would be much easier for the potential false anomalist to perform. They would not require the same amount of preparation as the replacement group, and they would be generally very effective/impressive.

3.2.1 Peripherals.

410
The majority of peripherals might respond in the manner a false anomalist required if some colourless and odourless liquid (yes water, wonderful stuff for causing short circuits) was spilt down the works. The use of "palmed" containers with "squirter" attachments would make this method very hard to detect. As an alternative, a quick random flip of the switches at the back of the unit could be quite effective in causing a large enough distraction, at least for any major system "adjustments" that a false anomalist wished to make for later in the "anomaly".

3.2.2 CPU/Internal boards.

3.2.2.1 Static

Static electricity could also be highly effective in the destruction of the target computer's internal workings. A false anomalist could easily build up enough static to blow most CMOS chip-based boards, simply by walking up and down on a man-made carpet. This could be done very quickly and without drawing attention. Then it would be simply a matter of getting access to any available printed circuit board, and brushing by one of the chips. It would usually be best to pick the biggest chip on the board, and to sneakily direct the static discharge to one of the small board/chip connection legs. If a false anomalist could not get access to a circuit board, then a computer's serial (RS232) port would be the next best thing.

3.2.3 Personal electronic "belongings".

In this group come the majority of "easy", and impressive anomalies that could be the stock-in-trade of any good prospective false anomalist. Restarting any battery-based electronic gadgets could simply be a matter of either heating the battery, and temporarily re-activating any residual chemical action in the almost exhausted power cells; shame the fix would not be permanent but the false anomalist could be long gone by the time the chemical action has died again) or of giving the object a "psi blast" (otherwise known as a damn good shake, which should put back any bad connections). In the case of a false anomalist being presented with anything with a printed circuit board, the "repair" could often be achieved by pressing the ends and sides of the unit (this would reform any broken solder connections on the board, but only so long as the pressure was maintained; still the false anomalist could claim that the "psi" only lasts that long anyway).

3.2.4 Backing storage.

3.2.4.1 Floppy Discs.

Another potential anomaly could be that of erasing the contents of a floppy disc. Unfortunately these magnetic storage devices are renowned for their environmental sensitivity. The following list of environmental factors are common enough causes of disc failures to cause disc manufacturers to print warnings about them on the dust covers of discs:

- weak magnetic fields.
- bending of the disc.
- exposure for long periods to strong sunlight.
- contact with sharp objects.
- humidity outside the range of 8% to 80%.
- temperatures outside the range 10 °C to 50 °C.
- contact with static electricity.
- any substances or object making contact with the disc's recording surface.

A would-be false anomalist would have merely to select the most convenient method for the situation; a "grubby" thumbprint over the recording surface (if physical contact is allowed) or a weak magnet (if no contact is allowed). The best place to aim the attack would be the outer edge, since that's where the majority of directories are stored. If the conditions got really well controlled, a good old fashioned coughing spasm over the recording surface might just do the trick. A false anomalist would also have the well-known advantage of being able to "Psi displace" (or kick the hard disc unit while everyone is watching the floppy disc).

3.2.4.2 Hard Discs.
These are also magnetic media, but they are accessed much more quickly and precisely than the floppy variety. Consequently they are very sensitive to changes in electrical mains supply. Changes in the mains tends to cause anything from a data corruption error to a read head crash. A would-be false anomalist might expect the user to protect the system against this by installing either a “pure” computer mains supply, or if the user is really expecting, trouble a battery supply during the anomaly. In the case of the former the attachment of a timer-based air fan heater on another pure mains outlet would do the trick. If the heater’s timer is set the false anomalist could time the “psi blast” just right. With a little luck the system users might even be monitoring the mains. In that case they might be amazed by the false anomalist’s ability to influence a mains supply. In the event of a battery-based power supply being used the false anomalist could leave an anonymous carrier bag near the battery. In many cases the battery supply is situated in a central unit in the building, which means the bag (which of course would contain 1-2 kilograms of “dry ice” (wonderful stuff for freezing the input pipes of water cooled mainframe computers, but that’s another story)) could be placed right on top of the battery. This method would have two advantages, and one disadvantage. The advantages would be that the battery supply would supply power for long enough for all the preliminary tests, and the second advantage would be that the dry ice would evaporate leaving no trace. The disadvantage would be that the vapor discharged by the dry ice might draw attention to the bag, but if the battery supply was in a central remote site this might not be a problem. If all these tricks were foreseen a false anomalist could resort to a cigarette (high tar of course), or even better, a cigar to blow into the disc’s air vent. The smoke would collect over the disc’s surface and cause a read head crash. All of this would assume that the system controllers are guarding against a well aimed kick/shake of the hard disc unit.

3.2.4.3 Printouts.

The use of a quickly evaporating dissolving spirit can be enough to remove some types of computer-produced printout, and is usually enough to remove characters from the carbon copies of printouts that are so popular in most large businesses. Pointing out the obvious gaps left by the dissolved text could be just the distraction for the well-prepared centre piece, be it a full hard disc head crash, or merely smearing a very thin layer of petroleum jelly mixed with ferrite dust onto the next floppy disc to be placed into disc drive. Not only would the disc fail, but “wonder of wonders” the read heads would fail too!! A false anomalist would of course be careful to avoid getting the spirit on anything other than the printouts since it tends to smell a bit.

3.2.5 Breaches in computer/organisational security - prearranged.

In the case of a false anomalist wishing to be thought responsible for mysterious, and untraceable messages coming into the computer system, the connecting lines between computers could have a radio receiver (with a powerful magnetic broadcaster) left in the vicinity of one of the computer system’s communication lines. This would swamp the normal messages being sent down the line to mystified users, who would now only receive the broadcast messages. The physical size of a receiver/transmitter can be very small indeed (perhaps the size of a box of matches). In the case of a “ring” or “ethernet” communication system, a line monitor could remove and reintroduce the packet that was being sent from user to user. This would also be done under radio control of the infiltrator’s remote computer system.

3.2.6 Breaches in computer/organisational security simultaneously with anomaly.

In the case of wishing to simulate a whole communications system failure as an anomaly, a powerful electromagnet could be left on the communication cables. The magnet could be timer-based so that the anomaly was short lived. This would enable a false anomalist to practice method acting and impress the system’s users as to the draining affect such powerful demonstrations have on the body. The brevity would also ensure that the more knowledgeable users would not be able to precisely locate the line fault, which would merely be a matter of checking the transmission potential periodically along the line.

4. EXTERNAL TO SYSTEM.

4.1 Natural “disasters”.

Quite often an unconnected event could be assumed to be a integral part of an anomaly. Central electrical supply fluctuations could be the cause of many so-called amazing anomalies. A change in mains alternation speed or current might produce anything from a speeding up of the computer’s clock, to a burnt out circuit board. Such anomalies are unlikely to be replicated. If they coincided with a meaningful event - e.g. an
argument among users during the mains surge, it may cause local lore to be established. “Every time Julie and Konrad have a blazing row the machine seems to......”

5. METHODS OF AVOIDING COMPUTER BASED FRAUDULENT ANOMALIES.

If this seems to be painting a hopeless situation, hope can be derived from some new developments in computing technology. These make many of the older problems which would have been involved in detecting and preventing any computer-based psi fraud much easier.

5.1 WORMS (write once read many times) optical discs.

These are non-editable, and like audio compact discs are completely immune to “grubby” thumbs, magnetic fields, and static. A false anomalist would have to either switch discs, or physically scratch the surface, if they wanted to be seeming to “erase” them. (About to gain greater prominence by being included in the specification of the new IBM micro computer systems.-Models 60-(Intel 80286 @ 10MHZ) & 80-(Intel 80386 @@@ 16 & 20MHZ) have options for a 200Mbyte optical drive.)

5.2 Optical fibre cables.

Prevents electromagnetic monitoring of data passing down a computer line (previously achieved by recording the magnetic field around the cable); instead a very complex, and obvious line monitor would be needed.

5.3 Automated technical advisers for computer based security.

Z is a technique for describing systems at their design stage, to highlight gaps in security which can then be amended to avoid the errors highlighted. Developed at the University of Oxford, it is based upon set theory, and Typed first order logic.

5.4 Gypsy verification Environments.

Based on specification language and proof of code based on predicate calculus and Hoare logic. They usually include a run time verification of code. In the sense of this paper’s application these environments would be used to check that a program had not had its function amended.

5.5 Program analysis tools.

- SPADE - Developed for Pascal programs by the University of Southampton.
- Malpus - Developed for Coral66 programs by RSRE.

Both products analyse program code to check control flow, dataflow and information flow. “Control flow graphs” are reduced by eliminating straight line code. The result shows the potential areas that could be used by “infiltrating” programs, while “hiding” in the system. This would allow the program to be amended to try and reduce the possibility of this occurring.

5.6 Cryptographic methods.

In most cases experts agree (British Telecom 1987) that with a sufficiently long encryption key, this can be an effective method against illegal access to data. Its only weakness in the present context is the possibility of the encrypted file being deleted by remote users (“Psi hackers”), who may be involved in the malicious destruction of their own poor “psi” results, however when encryption is combined with WORMS technology a much more satisfactory situation is reached. As long as the physical storage device was secure, the experimenter could be fairly happy at the integrity of the data being collected and stored.
by remote users. The use of encryption within the "message packets" being sent between machines would also reduce the risk of second parties gaining copies of internally transmitted information.

5.7 Shielding

Simple shielding of vdu screen emission can eliminate the chance of a computer screen being reconstituted outside the system confines.

6. CONCLUSION.

This paper has tried to exemplify the various scenarios that could be interpreted as "Anomalous" Human-Computer Interaction (A.H.C.I). It has also tried to show that there are conceptual patterns that allow A.H.C.I anomalies to be categorized, along with their possible fraudulent explanations. It is hoped that armed with such a method of categorization experimenters may be able better to record and evaluate the intriguing field of A.H.C.I. Finally, with regard to the suggested "solutions" provided in many of these "attack" scenarios, it may be more cost effective to create a means of detecting that a fraudulent anomaly has occurred, rather than to try and proof your system against every possible threat. It is certain that a false anomalist with a large financial and time resource would be able to create fraudulent anomalies, regardless of the tightest precautions. The only hope that the system user can hope for is to make the false anomalist "work" for every anomaly produced. Experimenters should adopt the same attitude but couple it with an experimental condition where no one star "makes or breaks" the results. By using large "anonymous" source groups the incentive for any one individual to create false anomalies is greatly reduced.
Appendix 2. An Abridged Listing Of An Example Session Report Produced By The Kmdb Analysis System. Samples From Each Of The Report’s Sections Have Been Provided.

Database Analysis Program

Copyright (c) 1987
Version 1.02
K.morgan.
Dept. Psychology,
University Edinburgh.
All Rights Reserved.

Subject Number : - 1
Run Number : - 1
Interface Type : - Command Line

Details From Run Commencing 1988 : 10 : 4: 10: 57: 6

As 183 Starting Seed Values Were

Ix = 17995
Iy = 4540
Iz = 26536

Error Report

First Error.

Time Of Error : 10:57: 8
Time Since Last Error : 0: 0: 0
Number Of Errors Recorded So Far : 1
Length Of Session : 0: 0: 0
Current Average Time Between Errors : 0: 0: 0
Error Number 72 Message Displayed Was :
“Assessment Record File Does Not Exist”
Generated From Within Program Module Assessme.pas

Next Error.

Time Of Error : 10:57: 8
Time Since Last Error : 0: 0: 0
Number Of Errors Recorded So Far : 2
Length Of Session : 0: 0: 0
Current Average Time Between Errors : 0: 0: 0
Error Number 73 Message Displayed Was:
"Records Must Be Created Before Any Query"
Generated From Within Program Module Assessme.pas

Next Error.

Time Of Error : 11: 2:44
Time Since Last Error : 0: 5:36
Number Of Errors Recorded So Far : 3
Length Of Session : 0: 5:36
Current Average Time Between Errors : 0: 1:52
Error Number 18 Message Displayed Was:
"First Verb Unrecognised"
Generated From Within Program Module Command.pas

Totals Per Error Number This Session...

Total Of Error Number 4 = 1
Total Of Error Number 6 = 4
Total Of Error Number 7 = 8
Total Of Error Number 8 = 11
Total Of Error Number 11 = 1
Total Of Error Number 18 = 3
Total Of Error Number 38 = 17
Total Of Error Number 39 = 1
Total Of Error Number 55 = 1
Total Of Error Number 72 = 1
Total Of Error Number 73 = 1
Total Of Error Number 82 = 1

Issued Commands Report

Details Of Command Entered To System At 10:57:40

Command Type = Help
Help Was Accessed On The Following Topic
Basic Help, No Subtopic
Record Type Command Requested For:
All - General Type.
Condition Requested (ne Is Default) : All
Length Of Time Since Last Command = 0: 0: 0
Number Of Commands Issued So Far : 1
Length Of Session : 0: 0: 0
Current Average Time Between Commands : 0: 0: 0

Details Of Command Entered To System At 10:58:12

Command Type = Select
Nature Of Select Can Be Seen From Record Type(s)
Record Type Command Requested For:
Department Record.
Condition Requested (ne Is Default) : All
Length Of Time Since Last Command = 0: 0:32
Number Of Commands Issued So Far : 2
Length Of Session: 0: 0:32
Current Average Time Between Commands: 0: 0:16

**Details Of Command Entered To System At 10:59:54**

Command Type = Select
Nature Of Select Can Be Seen From Record Type(s)
Record Type Command Requested For:
Student Record.
Condition Requested (ne Is Default): Ne
Length Of Time Since Last Command = 0: 1:42
Number Of Commands Issued So Far: 3
Length Of Session: 0: 2:14
Current Average Time Between Commands: 0: 0:44

**Total Number Of Commands Issued Per Command Type:**

Number Of Selects Issued = 18
Number Of Helps Issued = 7
Number Of Modifys Issued = 2
Number Of Adds Issued = 6
Number Of Deletes Issued = 0
Number Of Exits Issued = 1
Number Of Quits Issued = 0
Number Of Saves Issued = 0

**Report On Random Sources.**

**Long Period Pseudo Rng Distribution Was:**

Total Produced = 435
Total Taken As Right = 215
Total Taken As Left = 220
Z Score For Right = -0.240
Z Score For Left = 0.240

**Short Period Pseudo Rng Distribution Was:**

Total Produced = 7
Total Taken As Right = 2
Total Taken As Left = 5
Z Score For Right = -1.134
Z Score For Left = 1.134

**Interuption Of Fast Clock Distribution Was:**

Total Produced = 28
Total Taken As Right = 13
Total Taken As Left = 15
Z Score For Right = -0.378
Z Score For Left = 0.378
Report On Smart Noise Trials

Smart Noise Under Found Condition

Request No : 1

Time Of This Request : 11:1:31
No Of Trials : 4
No Of Hits : 3
No Of Misses : 1

Cumulative Totals So Far For This Request :
Trials = 4
Hits = 3 Mce Would Be = 2
Misses = 1 Mce Would Be = 2
Ratio Hits To Misses = 3

Request No : 2

Time Of This Request : 11:41:59
No Of Trials : 1
No Of Hits : 1
No Of Misses : 0

Cumulative Totals So Far For This Request :
Trials = 5
Hits = 4 Mce Would Be = 2
Misses = 1 Mce Would Be = 2
Ratio Hits To Misses = 4

Report On Smart Noise Trials

Smart Noise Under Not Found Condition

Request No : 1

Time Of This Request : 11:49:16
No Of Trials : 1
No Of Hits : 1
No Of Misses : 0

Cumulative Totals So Far For This Request :
Trials = 1
Hits = 1 Mce Would Be = 0
Misses = 0 Mce Would Be = 0

Cumulative Totals For Requests (under Both Conditions):
Trials = 7
Hits = 6 Mce Would Be = 3
Misses = 1 Mce Would Be = 3
Z Scores For Hits = 1.890
Z Scores For Misses = -1.890
Ratio Hits To Misses = 6

Report On Psi Based Records

Results From Trial Number = 5

<table>
<thead>
<tr>
<th>Source Name</th>
<th>Direction Selected by Source</th>
<th>Direction Selected by SSM Overall Vote.</th>
<th>Actual Direction.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coin</td>
<td>Left</td>
<td>Left</td>
<td>Left</td>
</tr>
<tr>
<td>Unigram</td>
<td>Left</td>
<td>Left</td>
<td>Left</td>
</tr>
<tr>
<td>Quick Stab</td>
<td>Left</td>
<td>Left</td>
<td>Left</td>
</tr>
<tr>
<td>Quick Coin</td>
<td>Left</td>
<td>Left</td>
<td>Left</td>
</tr>
<tr>
<td>Regular Wave</td>
<td>Right</td>
<td>Left</td>
<td>Left</td>
</tr>
<tr>
<td>Processor Speed</td>
<td>Left</td>
<td>Left</td>
<td>Left</td>
</tr>
<tr>
<td>Displacement</td>
<td>Left</td>
<td>Left</td>
<td>Left</td>
</tr>
<tr>
<td>Exact Missing</td>
<td>Left</td>
<td>Left</td>
<td>Left</td>
</tr>
</tbody>
</table>

Cumulative Totals Per Noise Source:

Mce Would Be: 0.5
Coin Hits Predicted: 1
Unigram Hits Predicted: 1
Quick Stab Hits Predicted: 1
Quick Coin Hits Predicted: 1
Reg. Wave Hits Predicted: 0
Proc. Speed Hits Predicted: 1
Displacement Hits Predicted: 1 Number Times Used = 1
E. Missing Hits Predicted: 1 Number Times Used = 1

Combined Psi Choice Total So Far ..

Mce Would Be: 0.50
Psi Hits Predicted: 1; z = 1.000
End Of Report
Appendix 3. Reports on the 6 main factors which emerged from a factor analysis of the TAQ3.

Results of the factor analysis on the Biology 1st year TAQ3 questionnaire and recorded Emas usage data (N = 64).

Factor 1 - 'The Function Linked Person'.

Factor 1 had an eigenvalue of 13.15, and a percentage of total variance figure of 13.8.

Positive factor matrix data (n = 64, number of items = 29, Cronbach's reliability coefficient = 0.75).

People in this type had a large number of sessions, high overall time on the machine and issued a large number of commands to the operating system. They used electronic mail, listed and edited a large number of files. They subscribed to bionews, and used the help and 'finger' commands. From their TAQ responses it seems that they had a tendency to use a programmable calculator, play video games, use a microcomputer at home, and at college, and to wear multi-function digital watches. They read books about computers beyond their course needs, and seldom found it difficult to relax around computers and new technology. They and their friends regularly subscribed to computer and new technology magazines. They stated their main use of the computer was for programming.

Negative factor matrix data. (n = 64, number of items = 8, Cronbach's reliability coefficient = 0.68).

This type were predominately male, and were taking the course to get a good job. They enjoyed working with, and repairing machines. They also preferred using a machine themselves rather than letting someone else use it for them. They reported feeling that computers were fun to use. Members of this group rarely reported getting the feeling that they could not do as well as others (high self esteem).

Comment.

This group would form the ideal computer user and worker for any general company selection scheme.

Factor 2 - 'The Malfunction linked person'.

Factor 2 had an eigenvalue of 6.91, and a percentage of total variance figure of 7.3.

Positive factor matrix data. (n = 64, number of items = 15, Cronbach's reliability coefficient = 0.33).

People in this type completed a large number of sessions, and issued a large number of commands. However they did not enjoy working with tools, equipment, or machines, in fact they liked someone else using the machine for them. This type felt that computers were not fun to use, and that computers did not

1 The finger command allows the user to check if another user is logged on to the system.
help them work better. They had fairly low self-esteem (they frequently felt others did better than themselves), and they felt that they let people use them. They were unsure of how they would cope with new situations, and were prone to stress in their daily work. This group also reported being frightened by some kind of equipment when they were children.

Negative factor matrix data. (n = 64, number of items = 14, Cronbach’s reliability coefficient = 0.76).

Looking at the negative factor matrix scores for this type we found that they felt that computers were too complicated for them to understand. They reported being bothered by the impersonal nature of computers, and seldom using a microcomputer at home. They reported having frequent problems with mechanical and electric breakdowns in their home environment. This group also reported that they were dissatisfied with their life, and that they frequently avoided social situations where there was a chance of criticism or rejection. They reported that people they knew frequently ignored them, and that it was hard for them to relax around computers and new technology. They reported that they avoided machines because they always seemed to break down when they used them, members of this group were never encouraged to explore how machines worked when they were children.

Comment.

This group of the population are evidently experiencing quite severe problems in their daily lives when they try to use and interact with machines.

Factor 3.

Factor 3 had an eigenvalue of 5.64, and a percentage of total variance figure of 5.9.

Positive factor matrix data. (n = 64, number of items = 10, Cronbach’s reliability coefficient = 0.75).

Factor 3 types issued few (if any), commands on the computer, and did not enjoy working with machines. However they reported having few problems with machines in their working environment (perhaps because they did not use them). They predominately adopted a strategy of self-teaching when they had to learn how to use a computer system, and liked to be shown and then be supervised in general learning situations. This group did not report that they regularly participated in any recreational activity which involved physical dexterity and precision.

Negative factor matrix data. (n = 64, number of items = 5, Cronbach’s reliability coefficient = 0.55).

This group had a high degree of self-esteem, and did not let other people ‘use’ them. They reported seldom helping their family to repair equipment in their family home (when they were young), and they also reported that they were not allowed to use dangerous equipment. This type were not encouraged to explore how machines worked by their parents.

Comment.

People who fall into this group seem to seldom use the computer (of their own free will). Indeed there is little evidence from the questionnaire data that this group voluntarily engage in any activity at all.

Factor 4.

Factor 4 had an eigenvalue of 4.94, and a percentage of total variance figure of 5.2.

Positive factor matrix data. (n = 64, number of items = 9, Cronbach’s reliability coefficient = 0.41).

People in the factor 4 type issued a high number of quit commands, and frequently used the editor. This type reported that they did not enjoy doing things in a specific order, or following instructions. They reported having few problems with their electrical ‘home entertainment’ systems. They felt that they
were quick to adapt to changes in their life, and they felt that people seldom if ever ignored them. This group preferred to learn new things by being shown, and then being left to experiment by themselves.

**Negative factor matrix data. (n = 64, number of items = 5, Cronbach’s reliability coefficient = 0.2).**

People in this group seldom used the finger command (used to check if other computer users are logged on to the system). They felt that computers were too complicated for them to understand, and seldom read books about computers in their own time. However surprisingly they did state that some of their immediate family’s preferred occupations involved new technology. This type disliked being strictly supervised.

Comments.

In contrast to factor 3 types, factor 4 types actually do some work on the computer of their own free will, but they do not like using machines. However their dislike of doing things in strict order and being supervised does raise some problems in the majority of positions available in business computing.

**Factor 5.**

**Factor 5 had an eigenvalue of 4.15, and a percentage of total variance figure of 4.4.**

**Positive factor matrix data. (n = 64, number of items = 8, Cronbach’s reliability coefficient = 0.8).**

Factor 5 types performed 'list', 'ecce' and 'files' commands on the machine, indicating that they were viewing pre-existing material, editing it, and creating new material. They claimed to be using computers for statistical analysis (this element was felt to be 'noise').

**Negative factor matrix data. (n = 64, number of items = 6, Cronbach’s reliability coefficient = 0.63).**

This type had a low number of logins, and quits indicating that they were not intensive computer users. However they tended to carry out a large number of commands when they logged on each time. This suggests they had created a ‘plan’ of the actions they required and they stuck to it during each computer session. This group seldom used a microcomputer at home.

Comments.

This factor’s members seem to be methodical in their use if the computer, but do not use it for any form of enjoyment. This group would seem to be 'users', rather than 'creators' when it comes to machine use.

**Factor 6.**

**Factor 6 had an eigenvalue of 3.77, and a percentage of total variance figure of 4.0.**

**Positive factor matrix data. (n = 64, number of items = 7, Cronbach’s reliability coefficient = 0.41).**

People in this factor used the 'stop' and 'finger' commands quite often. The finger command is regarded a social command, so it is possible that this type viewed using the computer as chance for social interaction. This group was predominately female, and they did not feel that computers could help them to work better (probably because they were using the computer as a means of social interaction, and not as a tool). However they did read about computers beyond their course requirements. They preferred to learn how to use the computer either in a workshop, or from another person (this would tie in with their need for social interaction).

---

2. This is open to dispute since the reliability coefficient is quite high.
Negative factor matrix data. \( n = 64, \) number of items = 2, Cronbach's reliability coefficient = 0.3).

Members of this group did not issue many quit commands, instead they preferred using the 'stop' command. It is possible that the connotation with 'giving up' was not popular with them. They did not use the password setting/resetting command very frequently.
Appendix 4. Full Colour Plates (used to provide digitised images in Chapter 6).

Figure 128. Plate 1 - Napier Polytechnic of Edinburgh.
Plate 2. View of the Amstrad Machines.

Plate 3. View Of The Experimental Room.
Plate 4. Another View Of The Amstrad Machines.

Plate 5. View From The Experimental Room.
Appendix 5. Sample Error Categories and Types.

Please note that these are sample error descriptions, for a description of all the error messages please see the technical documentation.

The error numbers form distinct groupings which identify both the type and origin of the error. This allows a histogram to be produced for all the errors from each experiment, and this in turn allows the error making behaviour of each group to be compared.

Error Types over each interface

![Error Number Histogram](image)

Figure 133. Example Error Number Histogram.

As an example these are the error number histograms from the first experiment.

The error numbers are used to produce the breakdown of errors into syntax and semantic categories.

**Example Error Details.**

The following errors are generated in the program module command.pas

**Error number = 2**

*Message = * Age exceeds human limits! *

*Category of this error = ((Syntactic)(Semantic((object)(action))))

*Most likely category = ((Syntactic))

*Description = This is found in the function valid_age, and is issued if the age provided is greater than 140 years. Syntactically this could be due to typing a wrong number at the keyboard. Semantically it could be misunderstanding of the action (adding a students age), or the object of task they were undertaking namely age.

This error is possible on both systems.
Error number = 4
Message = * invalid number entered! *
Category of this error = ((Syntactic)(Semantic((object)(action))))
Most likely category = ((Syntactic))

Description = This is found in the function value_of_string, and is issued if the string provided to the function was non-numeric, or negative. It is called whenever an ASCII string has been read into the system, and has be converted into a numeric format. Syntactically this could be due to typing a wrong or invalid number at the keyboard. Semantically it could be a misunderstanding of the action (adding a number), or the object of task (a number) they were undertaking.

Possible in both systems

The following errors are generated in the program module common.pas

Error number = 72
Message = * assessment record file does not exist *
Category of this error = ((System(external file)))
Most likely category = ((System(external file)))

Description = This error is produced by procedure fill_assessment_tree when it searches the disc and finds that no assessment file exists. This message is displayed by the system automatically on start up of a new experimental session, since no assessment files exists.

Possible in both systems

Error number = 76
Message = * "key=" missing in modify command *
Category of this error = ((Syntactic)(Semantic((object)(action))))
Most likely category = ((Syntactic)(Semantic((action))))

Description = This is produced by procedure process_modify when the syntax element "key=" is missing during the modify command. This message is not possible on the Gem system since Key= is supplied by the system.

The following errors are generated in the program module command.pas

Error number = 78
Message = * badly formed modify statement *
Category of this error = ((Syntactic)(Semantic((object)(action))))
Most likely category = ((Syntactic))

Description = This error is produced by procedure process_modify when either the number of words input in the modify command is incorrect, or none of the words are recognised as being valid. This error is possible in both systems.
Error number = 79
Message = * Department Key already used - Modify intended? *
Category of this error = ((Syntactic)(Semantic((object)(action))))
Most likely category = ((Syntactic)(Semantic((object))))
Description = This error is produced by procedure analyse_add when the user has requested to add a record key that already exists. This could be a semantic action error if the user did not realise that the add command required a unique key for a new record, this could also be object error if the user did not understand what a key was. This error is possible on both systems.

Error number = 86
Message = * Dept Record Key NOT yet created - Add intended? *
Category of this error = ((Syntactic)(Semantic((object)(action))))
Most likely category = ((Syntactic)(Semantic((object))))
Description = This error is produced by procedure modify_query when a user has requested to modify a record they have not yet created. This could be a semantic action error if the user did not realise an already existing record has to specified, or it could be an object error if the user does not understand keys and records.
This error is possible on both systems.

Technology Attitude Questionnaire.
Prof. R. Morris, Mrs. S. Gibbs, Konrad Morgan (c) 1988.

Form Number: _______________________

Please complete the following information by filling out the required details or circling the appropriate answer:

Today's date: __________

Your course name: ______________________

Your age: ___ Your sex: (Male) / (Female)

Are you left or right handed? (Left) / (Right) / (Both)

________________________

1. Please indicate how you prefer to learn new things (tick one of the following boxes):

   a) [ ] Being shown and then experimenting by myself.
   b) [ ] Being shown and then trying while supervised.
   c) [ ] Working it out by myself.
When you answer the following series of questions please circle the numbers that most appropriately reflect your feelings.

2 - I enjoy figuring out how things work

True False
---1---2---3---4---5---6---7---8---

3 - I enjoy hobbies or activities which involve:

a. Doing things in a specific order

True False
---1---2---3---4---5---6---7---8---

b. Following instructions/recipes

True False
---1---2---3---4---5---6---7---8---

c. Working with tools or equipment

True False
---1---2---3---4---5---6---7---8---
4 - I am taking this course primarily because I hope it will get me a good job

True False
---1---2---3---4---5---6---7---8---

5 - I like to work with machines (e.g. cars, sewing machines, stereos)

True False
---1---2---3---4---5---6---7---8---

6 - I like to repair machines

True False
---1---2---3---4---5---6---7---8---

7 - I think computers are too complicated for me to understand.

Strongly agree Strongly disagree
---1---2---3---4---5---6---7---8---

8 - I want to be good at whatever I do

True False
---1---2---3---4---5---6---7---8---

9 - I prefer to use computers myself rather than have someone else use them for me

Strongly agree Strongly disagree
---1---2---3---4---5---6---7---8---

10 - I am afraid of damaging the computer if I use it.

True False
---1---2---3---4---5---6---7---8---

11 - Computers are more dependable than a human being.

Strongly agree Strongly disagree
---1---2---3---4---5---6---7---8---

432
12 - I am afraid of looking silly if I make a mistake while using the computer.

True False
---1---2---3---4---5---6---7---8---

13 - I am afraid of looking silly if I am seen by others not to know how to use the computer.

True False
---1---2---3---4---5---6---7---8---

14 - The impersonal nature of computers bothers me.

True False
---1---2---3---4---5---6---7---8---

15 - Computers are fun to use

True False
---1---2---3---4---5---6---7---8---

16 - I think computers can help me to work better

Strongly agree Strongly disagree
---1---2---3---4---5---6---7---8---
Please circle the most appropriate numbers.

17 - I use/have used the following devices:

a. Programmable Calculator
   Never       Daily
   ---1---2---3---4---5---6---7---8---

b. Automatic bank teller
   Never       Daily
   ---1---2---3---4---5---6---7---8---

c. Multi-function digital wrist watch
   Never       Daily
   ---1---2---3---4---5---6---7---8---

d. Video games
   Never       Daily
   ---1---2---3---4---5---6---7---8---

e. Microcomputer (at home)
   Never       Daily
   ---1---2---3---4---5---6---7---8---

f. Microcomputer (college)
   Never       Daily
   ---1---2---3---4---5---6---7---8---

g. Terminal/Mainframe (college)
   Never       Daily
   ---1---2---3---4---5---6---7---8---
18 - If you own or use one of the following items, please circle the option that states how often the item has acted as though it required repair or replacement in the last year.

a. Car, Motor Cycle, or Bicycle.

<table>
<thead>
<tr>
<th>Daily</th>
<th>Weekly</th>
<th>Monthly</th>
<th>6 Monthly</th>
<th>Yearly</th>
</tr>
</thead>
</table>

b. Personal electrical appliances. {Electric Shaver, Hair Drier, Wrist watch etc.}

<table>
<thead>
<tr>
<th>Daily</th>
<th>Weekly</th>
<th>Monthly</th>
<th>6 Monthly</th>
<th>Yearly</th>
</tr>
</thead>
</table>

c. Camera & Film.

<table>
<thead>
<tr>
<th>Daily</th>
<th>Weekly</th>
<th>Monthly</th>
<th>6 Monthly</th>
<th>Yearly</th>
</tr>
</thead>
</table>

d. T.V., Video Recorder, or Personal Computer.

<table>
<thead>
<tr>
<th>Daily</th>
<th>Weekly</th>
<th>Monthly</th>
<th>6 Monthly</th>
<th>Yearly</th>
</tr>
</thead>
</table>

e. Hi-Fi, Compact Disc Player, Radio, Tapes, Compact discs, and Records.

<table>
<thead>
<tr>
<th>Daily</th>
<th>Weekly</th>
<th>Monthly</th>
<th>6 Monthly</th>
<th>Yearly</th>
</tr>
</thead>
</table>

19 - I read books/articles about computers - beyond those required for my course.

Never       Daily
---1---2---3---4---5---6---7---8---
In the following section please tick the most appropriate answer or answers:

20 - I initially learned to use a computer through

a: [ ] using a manual
b: [ ] using a self taught course
c: [ ] attending a workshop or course
d: [ ] learning from another person
e: [ ] Other :- please specify

21 - I usually use a computer for:

a: [ ] Programming
b: [ ] word-processing
c: [ ] financial processing
d: [ ] statistical analysis
e: [ ] data base development
f: [ ] graphics
g: [ ] Other, please specify

436
Early Attitudes and experiences.

Please circle the numbers that most appropriately reflect your attitudes and experiences as you were growing up.

22 - I used to help my family repair equipment

Never 1 2 3 4 5 6 7 8

23 - I was allowed to use dangerous equipment

Never 1 2 3 4 5 6 7 8

24 - My family made sure broken equipment was repaired promptly

Seldom 1 2 3 4 5 6 7 8

25 - I was at least mildly injured by a machine

Never 1 2 3 4 5 6 7 8

26 - I was encouraged to explore how machines worked

Never 1 2 3 4 5 6 7 8

27 - My parents repaired equipment themselves

Never 1 2 3 4 5 6 7 8
28 - At our house machines seemed to break down.

Seldom          Often
---1---2---3---4---5---6---7---8---

29 - I was frightened by some kind of equipment

Never           Often
---1---2---3---4---5---6---7---8---

Thank you for filling this out.

Please add your comments about the questionnaire:

Were any questions hard to understand? If so which one(s), and why?

Did any questions seem to be prompting for particular answers? If so which one(s), and why?

General comments (please use the other side of this sheet if required):
Form Number: ____________________________

Todays date: ____________________________

Interface type: COMMAND

Your course name: _______________________

Your age: ___

Please circle the appropriate answer.

Your sex is: (Male) / (Female)

Are you left or right handed? (Left) / (Right) / (Both)

Please circle the numbers that most appropriately reflect your impressions about using this computer system. Please add your written comments about any item at the end of this questionnaire.

In all the following questions please circle "N/A" if the question does not apply.

1 - Have you used a command driven computer system before?

(Yes) / (No)

2 - How did you find the use of the keyboard?

very easy          very hard

---1---2---3---4---5---6---7---8---

N/A

3 - How closely did the words used as commands match the action which the computer then carried out?

poorly           closely

---1---2---3---4---5---6---7---8---

N/A.
4 - When you tried a new command did you find that your experience from previous commands helped you get it correct?

never       always

---1---2---3---4---5---6---7---8---
N/A.

5 - How did you find the instructions which described the tasks you had to perform?

confusing  clear

---1---2---3---4---5---6---7---8---
N/A.

6 - What did you think of the number of things the system required you to do before it completed an action?

too many       too few

---1---2---3---4---5---6---7---8---
N/A.

7 - Did the system prevent mistakes?

never       always

---1---2---3---4---5---6---7---8---
N/A.

8 - When the system displayed information about how each command had worked was the information sufficient?

never       always

---1---2---3---4---5---6---7---8---
N/A.

9 - Was the work simplified by the way the system displayed its information?

never       always

---1---2---3---4---5---6---7---8---
N/A.
10 - How did you find the systems speed?

too slow                too fast

---1---2---3---4---5---6---7---8---

N/A.

11 - Were the error messages helpful?

never                  always

---1---2---3---4---5---6---7---8---

N/A.

12 - Did the error messages indicate the corrective actions to be taken?

never                  always

---1---2---3---4---5---6---7---8---

N/A.

13 - How was the error correction?

confusing              clear

---1---2---3---4---5---6---7---8---

N/A.

14 - How was the on-line help?

confusing              clear

---1---2---3---4---5---6---7---8---

N/A.

15 - What was it like learning to use the system?

difficult              easy

---1---2---3---4---5---6---7---8---

N/A.
16 - How was the information you needed to complete each task presented?

needed to be memorized accessible on screen
---1---2---3---4---5---6---7---8---
N/A.

17 - What were your overall reactions to the system?

a) bad good
---1---2---3---4---5---6---7---8---
N/A.

b) frustrating satisfying
---1---2---3---4---5---6---7---8---
N/A.

c) dull stimulating
---1---2---3---4---5---6---7---8---
N/A.

d) difficult easy
---1---2---3---4---5---6---7---8---
N/A.

e) inadequate power adequate power
---1---2---3---4---5---6---7---8---
N/A.
Thank you for filling this out.

Please add your comments about the questionnaire:

Were any questions hard to understand? If so which one(s), and why?

Did any questions seem to be prompting for particular answers? If so which one(s), and why?

General comments (please use the other side of this sheet if required):

User evaluation of interactive computer systems
Konrad Morgan (c) 1988.

Form Number:_____________________

Today's date:_____________________

Interface Type: GEM

Your course name:_________________

Your age:___

Please circle the appropriate answer

Your sex:  (Male) / (Female)

Are you left or right handed?  (Left) / (Right) / (Both)

Please circle the numbers that most appropriately reflect your impressions about using this computer system. Please add your written comments about any item at the end of the questionnaire.

In all the following questions please circle "N/A" if the question does not apply.

1 - Have you used a 'mouse' controlled computer system before?

(Yes) / (No)

2 - Did you find the 'mouse' easy to control?

very easy  very hard

---1---2---3---4---5---6---7---8---

N/A.

3 - How did you find the concept of 'pull down menus'?

very easy  very hard

---1---2---3---4---5---6---7---8---

N/A.

444
4 - How did you find using the mouse and the keyboard at the same time?

   Very easy         Very hard
   ___1___2___3___4___5___6___7___8___
   N/A

5 - How closely did the words which the computer recognised as commands match the action it carried out?

   poorly            closely
   ___1___2___3___4___5___6___7___8___
   N/A.

6 - When you tried a new command did you find that your experience from previous commands helped you get it correct?

   never             always
   ___1___2___3___4___5___6___7___8___
   N/A.

7 - How did you find the instructions which described the tasks you had to perform?

   confusing         clear
   ___1___2___3___4___5___6___7___8___
   N/A.

8 - What did you think of the number of things the system required you to do before it completed an action?

   too many          too few
   ___1___2___3___4___5___6___7___8___
   N/A.

445
9 - Did the system prevent mistakes?

never always
---1---2---3---4---5---6---7---8---
N/A.

10 - When the system displayed information about how each command had worked, was the information sufficient?

never always
---1---2---3---4---5---6---7---8---
N/A.

11 - Was the work simplified by the way the system displayed its information?

never always
---1---2---3---4---5---6---7---8---
N/A.

12 - How did you find the systems speed?

too slow too fast
---1---2---3---4---5---6---7---8---
N/A.

13 - Were the error messages helpful?

never always
---1---2---3---4---5---6---7---8---
N/A.

14 - Did the error messages indicate the corrective actions to be taken?

never always
---1---2---3---4---5---6---7---8---
N/A.
15 - How was the error correction?

confusing        clear

---1---2---3---4---5---6---7---8---

N/A.

16 - How was the on-line help?

confusing        clear

---1---2---3---4---5---6---7---8---

N/A.

17 - What was it like learning to use the system?

difficult        easy

---1---2---3---4---5---6---7---8---

N/A.

18 - How was the information you needed to complete each task presented?

needed to be memorized         accessible on screen

---1---2---3---4---5---6---7---8---

N/A.

19 - What were your overall reactions to the system?

a)  bad          good

---1---2---3---4---5---6---7---8---

N/A.

b)  frustrating   satisfying

---1---2---3---4---5---6---7---8---

N/A.

447
c) dull                  stimulating  ---1---2---3---4---5---6---7---8---
                         N/A.

d) difficult              easy            ---1---2---3---4---5---6---7---8---
                                     N/A.

e) inadequate power      adequate power  ---1---2---3---4---5---6---7---8---
                                       N/A.

Thank you for filling this out.

Please add your comments about the questionnaire:

Were any questions hard to understand? If so which one(s), and why?

Did any questions seem to be prompting for particular answers? If so which one(s), and why?

General comments (please use the other side of this sheet if required):
Appendix 9. TAQ Version 3.

Technology Attitudes Questionnaire V3.0
Psychology Department,
Edinburgh University.

Form Number: __________________________

Please complete the following information by filling out the required details or circling the appropriate answer:

Today's date: ______________

Your course name: __________________________

Your age: ___  Your sex: (Male) / (Female)
Are you left or right handed? (Left) / (Right) / (Both)

1 - Please indicate how you prefer to learn new things (tick one of the following boxes):—

a) [ ] Being shown and then experimenting by myself.
b) [ ] Being shown and then trying while supervised.
c) [ ] Working it out by myself.
When you answer the following series of questions please circle the numbers that most appropriately reflect your feelings.

2 - I enjoy figuring out how things work

True 1 2 3 4 5 6 7 8 False

3 - I enjoy hobbies or activities which involve:

a. Doing things in a specific order

True 1 2 3 4 5 6 7 8 False

b. Following instructions/recipes

True 1 2 3 4 5 6 7 8 False

c. Working with tools or equipment

True 1 2 3 4 5 6 7 8 False

4 - I am taking this course primarily because I hope it will get me a good job

True 1 2 3 4 5 6 7 8 False
5 - I like to work with machines (e.g. cars, sewing machines, stereos)

   True   False
   ---1---2---3---4---5---6---7---8---

6 - I like to repair machines

   True   False
   ---1---2---3---4---5---6---7---8---

7 - I think computers are too complicated for me to understand.

   Strongly agree   Strongly disagree
   ---1---2---3---4---5---6---7---8---

8 - I want to be good at whatever I do

   True   False
   ---1---2---3---4---5---6---7---8---

9 - I prefer to use computers myself rather than have someone else use them for me

   Strongly agree   Strongly disagree
   ---1---2---3---4---5---6---7---8---

10 - I am afraid of damaging the computer if I use it.

   True   False
   ---1---2---3---4---5---6---7---8---

11 - Computers are more dependable than a human being.

   Strongly agree   Strongly disagree
   ---1---2---3---4---5---6---7---8---

12 - I am afraid of looking silly if I make a mistake while using the computer.

   True   False
   ---1---2---3---4---5---6---7---8---
13 - I am afraid of looking silly if I am seen by others not to know how to use the computer.

True False
---1---2---3---4---5---6---7---8---

14 - The impersonal nature of computers bothers me.

True False
---1---2---3---4---5---6---7---8---

15 - Computers are fun to use

True False
---1---2---3---4---5---6---7---8---

16 - I think computers can help me to work better

Strongly agree Strongly disagree
---1---2---3---4---5---6---7---8---

Please circle the most appropriate numbers.

17 - I use/have used the following devices:

a. Programmable Calculator
Never Daily
---1---2---3---4---5---6---7---8---

b. Automatic bank teller
Never Daily
---1---2---3---4---5---6---7---8---

c. Multi-function digital wrist watch

Never Daily
---1---2---3---4---5---6---7---8---
d. Video games

Never       Daily
---1---2---3---4---5---6---7---8---

e. Microcomputer (at home)

Never       Daily
---1---2---3---4---5---6---7---8---

f. Microcomputer (college)

Never       Daily
---1---2---3---4---5---6---7---8---

g. Terminal/Mainframe (college)

Never       Daily
---1---2---3---4---5---6---7---8---

18 - If you own or use one of the following items, please circle the option that states how often the item has acted as though it required repair or replacement in the last year.

a. Car, Motor Cycle, or Bicycle.

<table>
<thead>
<tr>
<th>Daily</th>
<th>Weekly</th>
<th>Monthly</th>
<th>6 Monthly</th>
<th>Yearly</th>
</tr>
</thead>
</table>

b. Personal electrical appliances. {Electric Shaver, Hair Drier, Wrist watch etc.}

<table>
<thead>
<tr>
<th>Daily</th>
<th>Weekly</th>
<th>Monthly</th>
<th>6 Monthly</th>
<th>Yearly</th>
</tr>
</thead>
</table>
c. Camera & Film.

<table>
<thead>
<tr>
<th>Daily</th>
<th>Weekly</th>
<th>Monthly</th>
<th>6 Monthly</th>
<th>Yearly</th>
</tr>
</thead>
</table>

d. T.V., Video Recorder, or Personal Computer.

<table>
<thead>
<tr>
<th>Daily</th>
<th>Weekly</th>
<th>Monthly</th>
<th>6 Monthly</th>
<th>Yearly</th>
</tr>
</thead>
</table>

e. Hi-Fi, Compact Disc Player, Radio, Tapes, Compact discs, and Records.

<table>
<thead>
<tr>
<th>Daily</th>
<th>Weekly</th>
<th>Monthly</th>
<th>6 Monthly</th>
<th>Yearly</th>
</tr>
</thead>
</table>

19 - I read books/articles about computers - beyond those required for my course.

Never

--- Daily

---1---2---3---4---5---6---7---8---

In the following section please tick the most appropriate answer or answers :

20 - I initially learned to use a computer through

a: [ ] using a manual
b: [ ] using a self taught course
c: [ ] attending a workshop or course
d: [ ] learning from another person

---
e: [ ] Other :- please specify


21 - I usually use a computer for:

a: [ ] Programming
b: [ ] word-processing
c: [ ] financial processing
d: [ ] statistical analysis
e: [ ] data base development
f: [ ] graphics
g: [ ] Other, please specify


455
Current attitudes and experiences.

Please circle the numbers that most appropriately reflect your current attitudes and experiences.

22 - How happy are you with your life?

Dissatisfied
---1---2---3---4---5---6---7---8---
Satisfied

23 - How often do you get the feeling that no matter how hard you try you can never do as well as others?

Never
---1---2---3---4---5---6---7---8---
Daily

24 - Do you feel that you give too much and sometimes let people use you because you can’t imagine why else they would hang around you?

Never
---1---2---3---4---5---6---7---8---
Frequently

25 - Do you get very angry or depressed when you are criticized?

Frequently
---1---2---3---4---5---6---7---8---
Never

26 - Do you ever avoid social situations where there is a chance of criticism or rejection?

Frequently
---1---2---3---4---5---6---7---8---
Never

27 - How quickly do you think you adjust to change in your life?

Gradually
---1---2---3---4---5---6---7---8---
456
Quickly

28 - How sure are you that you will be able to cope with new or uncertain situations?

Very Sure
---1---2---3---4---5---6---7---8---

Unsure

29 - Do you or have you ever participated in any recreational activity that involve any degree of physical dexterity and precision?

Frequently
---1---2---3---4---5---6---7---8---

Never

30 - If you close your eyes for 20 to 30 seconds, how accurately can your sense the exact position of your fingers and hands?

Exactly
---1---2---3---4---5---6---7---8---

Poorly

31 - How often do people you know ignore you?

Frequently
---1---2---3---4---5---6---7---8---

Never

32 - Do you often have to turn down social engagements because you already have previous commitments?

Never
---1---2---3---4---5---6---7---8---

Frequently

33 - How often do you get so pressured with your work that you feel you just can’t cope?

Never
---1---2---3---4---5---6---7---8---

Daily

34 - Do you find it hard to relax when you are around computers and new technology?

Frequently
---1---2---3---4---5---6---7---8---

Never
35 - I don’t like using machines because they always seem to break down when I use them.

Frequently  Never
---1---2---3---4---5---6---7---8---

36 - How does your immediate family view gadgets and new technology?

Uses them frequently  Avoids them
---1---2---3---4---5---6---7---8---

37 - Do any of your immediate family’s preferred occupations involve new technology?

Totally  Not at all
---1---2---3---4---5---6---7---8---

38 - Do you or your friends regularly subscribe to computer or new technology based magazines?

Never  Regularly
---1---2---3---4---5---6---7---8---

Early Attitudes and experiences.
Please circle the numbers that most appropriately reflect your attitudes and experiences as you were growing up.

39 - I used to help my family repair equipment

Never  Often
---1---2---3---4---5---6---7---8---
40 - I was allowed to use dangerous equipment

Never                                      Often
---1---2---3---4---5---6---7---8---

41 - My family made sure broken equipment was repaired promptly

Seldom                                      Often
---1---2---3---4---5---6---7---8---

42 - I was at least mildly injured by a machine

Never                                      Often
---1---2---3---4---5---6---7---8---

43 - I was encouraged to explore how machines worked

Never                                      Often
---1---2---3---4---5---6---7---8---

44 - My parents repaired equipment themselves

Never                                      Often
---1---2---3---4---5---6---7---8---

45 - At our house machines seemed to break down.

Seldom                                      Often
---1---2---3---4---5---6---7---8---

46 - I was frightened by some kind of equipment

Never                                      Often
---1---2---3---4---5---6---7---8---

DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD
Thank you for filling this out.

Please add your comments about the questionnaire:

Were any questions hard to understand? If so which one(s), and why?

Did any questions seem to be prompting for particular answers? If so which one(s), and why?

General comments (please use the other side of this sheet if required) : 

460
Appendix 10. Pre-Session Questionnaire (Study 4).

Pre-session questionnaire.

Thank you for participating in this study. As your first task could you please complete the following brief list of questions.

1. Do you think it is likely that 'good' computer users have some form of unknown influence over the functioning of the equipment?

   Impossible  Certain
   --1--2--3--4--5--6--7--8--

2. Have you ever directly experienced or witnessed any inexplicable form of function linkage? (For example: Systems suddenly working or ceasing to work).

   Yes/No

3. How many hours per working day (on average) do you spend working with computer systems?

   ______ Hrs

4. Do you work consistently with one system or type of system?

   Yes/No

5. Please specify the type of system you spend most of your time working with.

   PCs-----Minis-----Mainframes

6. Please name the type and make of this system.

   ___________________________________________________________

   ___________________________________________________________

   461
7. Did you ever have an 'instinct' or 'hunch' that turned out to be correct about the source of problems (Hardware or Software) with regard to that system?

   Yes/No

8. It has been said that "computing can be a lonely, and antisocial business". Do you agree?

   Agree/Disagree/Agree (but don't mind it).

9. Do you ever use a computer system just for the enjoyment/’feel’ of it (not work based use)?

   Never           Often
   --1--2--3--4--5--6--7--8--

-----------------------------------

Thank you for completing this questionnaire. If you have any comments about this questionnaire please add them here. After you have completed this please start the technology attitudes questionnaire.
Appendix 11. Post-session Questionnaire (Study 4).

Post session questionnaire.

Number: __

Thank you for completing the session. As one final task could you please complete the following brief list of questions.

1. How much did you enjoy the session?

   Very Much    Not At All
   --1--2--3--4--5--6--7--8--

2. Has it changed your views on the likelihood of 'function linked' human-machine interaction?

   Not At All    Totally
   --1--2--3--4--5--6--7--8--

3. How often do you use a graphical/mouse based system?

   Constantly--Daily--Weekly--Monthly--Never

4. How do you view graphical interfaces for 'intensive computer users' like yourself?

   Useless    Useful
   --1--2--3--4--5--6--7--8--

5. Did you adopt any special strategy to try to increase your 'function linkage' during the session?

   Yes/No

6. If you answered Yes to 5 (above) please state the strategy you adopted - even if you feel it was very unsuccessful or unimportant.
7. If you answered Yes to 5 (2 questions above) then do you feel the strategy was successful?

Yes/No

8. Do you have any feelings or suggestions about how you might improve upon the strategy/ies you adopted?

9. Do you feel the experimenter (Konrad) influenced your performance in any way?

Unfavorable Influence  Favorable Influence

--1--2--3--4--5--6--7--8

Thank you for completing this questionnaire. If you have any comments about the study please add them here.
Appendix 12. Session Log.

Experimental Session Report.
Konrad Morgan (c) 1988.

Date of experimental session : / / 19

Time experiment started : HRS (GMT/BST)

Physical location of experiment :

Name of Experimenters :-
1) Mood :
2) Mood :
3) Mood :
4) Mood :

Expected number of subjects :
Actual number of subjects :
Number of Gem subjects :
Number of Com subjects :

Barometric pressure on day of experiment :
Weather conditions :

Phase of Moon (during experiment):
Position of Moon (during experiment):
Position of Sun (during experiment):

Comments on session (cont. on next leaf) :

The call to the function psi_choice generates the next left or right tree path decision. This is the result of some highly complex decision making systems. We will briefly discuss each of them in turn, remembering that these are covered in much greater depth in the technical reference to the KMDB system.

Exact Missing

The system held records of the overall left/right decisions made by the decision making system over the last 16 occasions, and the correct left/right targets which matched those guesses as series of 16 bit binary numbers. In the exact missing source these two 16 bit words were compared for any regularity which was exactly missing. If any such regularity was detected the Exact Missing process made its result the opposite of the result which had been produced by the other decision making systems in the current trial. If no exact missing was detected then no result was produced. As with all the decision making systems a weighting of previous success in predicting the correct direction decided how much influence the source made on the current trial. The nature of the b-tree task made both this and the following source vulnerable to an artifact, which will be discussed fully after we have outlined the method used by these two sources.

Displacement

The displacement system also used the two 16 bit result words to check for any displaced regularity between them. Displacement is believed by some parapsychologists to be a method used by subjects to avoid psi being observed, or as a natural error in psi processing.

The B-Tree Artifact.

The nature of the b-tree record retrieval task is such that it produces an artifact in both the exact missing and displacement checking systems if the user repeats the same trial multiple times in succession. This artifact results since it is possible for both these sources to 'learn' from the repeated patterns that emerge from two 16 bit result words if they are repeatedly given the same trial sequences as data. This artifact results in the exact missing source missing, and the displacement source hitting. This could have been avoided by erasing the past trial memory of the system. However this would result in the two sources (exact missing and displacement) only being able to check for displacement or exact missing over the past three to four trials. Using the adopted method these sources checked for displacement over the last 16 trials (as the whole process was being done 'on the fly' in real time it was impossible to check for any positive displacement). Since the study was covert subjects did not know that commands actioned trials, and only the author knew about the artifact it was felt to be safe to use the full 16 bit method. However the control trials which were run on the system included several duplicated trials to establish a realistic base line. The graphs of the decision making sources shown in chapters seven, eight, and nine confirm the accuracy of these control trials.

---

1 This reduced ram requirements and made PK corruption more difficult, since the most significant digit (MSD) of the word was swapped and copied to a register during the checking process. This was direction swapping was done because otherwise any bit shifting operations would cause arithmetic overflow to be generated.

2 Unless one used a pre-generated sequence, and assumed some form of retroactive influence. However the author would prefer to falsify all the real time alternatives before he is forced to accept the existence of transtemporal effects.
Quick Coin.

This was designed to produce a very fast binary result. It is handed the fast clock digits from the clock when the trial was initiated. This is used as the seed to a short period PRNG. It was hoped to be able to compare the complexity of each decision making source against its performance to see which was best for these kinds of smart noise applications. This source is in the low to middle range of complexity. The short period PRNG used in this source was chosen because of its simplicity. The mathematical behaviour of the system can be very easily determined, and remembered. This is in contrast to the long period PRNG, which is quite hard to comprehend. It would take a sophisticated mathematical ability to determine the outcome of the long period PRNG, in real time given the seeds.

Quick Stab.

This is perhaps the simplest of the decision making systems. It is based on the fast clock digits being in the range 00 through to 99. It was hypothesized (by the author), that by interrupting these digits and applying a modular two operation on the result one would get an odd or even result which had an equal probability. When this system was designed the author had assumed that these fast clock digits would behave as one would expect, that is, giving a uniform distribution between odd and even digits. Tests on the development machines seemed to confirm this assumption. However analysis of the data files produced by these sources during the actual experiments indicated a significant bias towards an even result. Detailed investigations revealed that this was unlikely to have been due to a paranormal influence. It is more likely to have been due to an exhausted real time clock battery. Fortunately this has no effect upon the integrity of the SSM. However it could have produced 'false PK' in a study that used fast clock seeding. This is one of the many problems which are discussed in the chapter on computer security.

Coin.

This uses the output of the long period PRNG AS183 which has a rectangular distribution in the range of 0.00 through to 0.99. It uses this PRNG as the basis of a binary random number source. This is implemented by splitting the output at the mid point in the distribution. This produces a very good binary random sequence. Empirical tests of over 40 million digits showed that it performed within chance expectation.

Unigram.

This is based on the method used by the I Ching, (see Rubin and Honorton, 1971). It is one of the most difficult sources to predict, since it uses the coin function (see above) three times, with each 'throw' being taken as either a three or a 2. The total produced over three calls or 'throws' of the coin function produces the outcome of the process, in the range six to 9. Even outcomes are taken as being left decisions, while odd outcomes are taken as right decisions. An added feature of this source is that a degree of certainty is also built into the result produced by the procedure. This degree of certainty has only two possible values, these being one and five respectively. This 'certainty value' is used with the success weighting to calculate the degree of influence for the unigram source. Great care was taken to ensure that the implementation of this source was in total accord with the method promoted by the ancient chinese system of divination.

Regular Wave.

This was based upon the extensive work the author did with random number generators in the first year of the study. In brief summary he found that most of these RNGs had sub-repeating patterns within them. Since many researchers had reported finding significance using these types of RNGS it was thought to

---

3 Trials indicated that the Quick Coin source produces over 8,000 decisions a second (over a million in two minutes).

4 AS183 has a period of 6.95 x 10^{12} (Wichmann & Hill, 1987).

5 As a note the Unigram source did no better than any of the other sources, all performed within MCE.
be reasonable to assume that the subjects might be using these repeating sequences to achieve their results. This would be so the subject's timed their responses to coincide with a favorable subpattern. The basis of regular wave is a 20 digit cycling sequence, the digits zero through nine are taken as left, and 10 through 19 are taken as being right. Entry to the wave sequence is determined by the fast clock digits, based upon the time the user commenced the operation. Since these fast clock digits are in the range 00 to 99, they are first corrected to be in the range zero to 20 by a modulus 20 operation. This result is then taken to be the pointer to the part of the regular cycle the system will use to make its decision. In the same way that the Unigram method (above) had a 'certainty' attached to its decision, the regular wave assumes a certainty by determining the extent to which the 'pointer' is currently into a left/right cycle.

Processor Speed.

This is one of the sections of the system that was the greatest disappointment to the author, and would not be implemented on any subsequent SSMs. The idea behind the processor speed decision source was that it is possible to run a section of code and time its execution. It is then possible to run the same section of code again and check the timing differences. On a true universal machine they should always be the same. However, even with the interrupts disabled there are always slight differences in these execution times. The processor speed source therefore records the fluctuations in the performance of the operating system. In MS DOS, which was the operating system used in this study, these fluctuations can be as large as 10% of the execution time (Souter & Davis, 1988). This decision making source was designed to investigate the possibility of Ss using a paranormal ability to control the timing of their trials to coincide with these fluctuations in the operating system's performance. The problems which these variations in performance could pose for computer based parapsychological experiments are discussed in the chapter on computer security.

The Majority Vote Of The Sources And The Role Of The Weighting.

The logic in the procedure psi_choice has two counters called l_count and r_count, standing for left count and right count respectively. These are set to zero at the start of each call of psi_choice, and each decision making source has its decision added to the appropriate counter value. As we have already discussed, each decision making source has a weighting attached to it. These operate so the more hits a source makes the higher the source's weighting becomes. All the weights start with values of 10. The 'voting' process involves adding the source's weighting (multiplied by 1) to the directional counter selected by the source for the current trial. This is except for the Unigram, and Regular wave sources, where the weighting is multiplied by the 'certainty' value. After the accuracy of the vote has been determined by the searching logic, the weights are updated for each source. Sources which were correct get their weighting incremented by 100, and those that were unsuccessful get their weight decremented by 100. This means that the weights are varying on a trial by trial basis, and have a direct influence on the overall outcome of the SSM. It would be interesting to have some throttle which directly controlled the size of the increments and decrements to these weightings. However since this was the first attempt at a SSM system the author felt it was probably best not to vary these weighting increment rates. For future research it might even be a possibility to allow the users to calibrate the throttle themselves. This would mean that the user should be able to see the success rates live at run time. It might be possible under these circumstances to have a small window on the screen which graphically displayed the success rate of the various sources, and allowed the user to vary the weightings and throttle.

Deadlock.

It will have occurred to the more discerning reader that it is possible for the SSM to come to a complete stalemate. Obviously there are two possible solutions, either to try again (with the possibility that it would reach stalemate again), or to have some means of resolving the conflict. It was decided that a quick and random decision was probably best if deadlock occurred, so in this event the coin function (see above) is used to resolve the problem.

A Possible Problem With RNGs.

No discussion of the use of psi would be complete without mention of the possible effects of what the author came to regard as one of the major problems with using RNGs in parapsychological experimental

---

6 This would still involve a paranormal ability, unless the sequence was very poor.
work. It is a possibility that psi could operate in the SSM we have discussed, but never be recorded. For example a subject could examine the system with ESP and use PK to maintain the system to within chance behaviour. An example will suffice to explain this concept. Suppose we have the perfect RNG, and it is producing a random sequence for a subject, but that the RNG would, under non-psi conditions, be producing a non random 'glitch' as all good RNGs occasionally should. The subject could be influencing the RNG to remove the glitch, and if it was a negative 'glitch' the subject could be producing very strong positive FAA indeed. Unfortunately this would never be recorded. In the SSM database system this would require the subject knowing the precise location of all the records in all the b-trees, and the exact outcomes of all the decision making sources for the session duration. The subject would then have to choose the timing of an operation so they exactly cancelled the decision making sources positive or negative FAA.
Appendix 14. Example of the Processing limits with a real time biological mechanism.

If it is assumed that the mind has a finite processing capacity (which is a reasonable biological assumption), even allowing for a massive amount of parallel processing, there is a minimum time resolution at which it will be unable to detect and react to activity. This is a similar problem to that faced by computer systems which have to monitor real time events. In such real time computer systems the minimum sensing resolution determines the theoretical limit of the 'snapshots' of activity which the processor(s) can be aware. In human biological systems ESP derived information could be instantly available to the processors (neurons), and reactions to the physical environment (PK) could also be made the instant this processing was complete. Therefore the minimum determining factor for such a system (in a real time model) would be its processing speed, which we have seen is around 2 milliseconds. These timings represent the minimum human inter-nerve cell communication times. These specifications will be taken to approximate the minimum times in which it is assumed that biologically based cognitive awareness can take place. This discussion assumes that the S will not be able to monitor and read the contents of the electrical signals, and addresses travelling to and from the processor, since to do so would involve 'sensing' the content of a parallel stream of electrons travelling at speeds approaching those of light (300,000,000 m/s).

Example of biological limits.

It is possible to determine the likely real time processing capacity of such a real time biological system by using a simple example. Let us assume that a human biological processor's task is to monitor the activity taking place in the computers 1 megabyte of RAM (1024K bytes). Let us also assume that the bit states within the RAM are capable of changing once every 500 ns (as is the case on a slow machine like the Amstrad 1512). If we use a slow machine as an example then any effect should also apply to the more high performance micros or mainframes. Now assuming the brain has approximately $10^{10}$ neurons (Michie, Turing & Good, 1948 in Hodges, 1983 p387) and a processing speed for each processor of 1 millisecond (which is half the speed usually quoted (Ganong, 1973)), this should give a generous over estimate of currently estimated neural processing capacity, since the discharge latency (in-activity) will overlap. If we presumed each brain cell to be a processor (which is unlikely), we can calculate the uppermost biological processing limits that are possible for a real time system that is trying to monitor the 1 megabyte of RAM, using instantaneous ESP and PK.

First we can calculate the number of states to be monitored by the processors. There are 1024,000 bytes, each byte of which contains 8 bits of information. These are assumed to be able to change every 500ns. Note that all 1024,000 locations are updated on every memory refresh cycle, otherwise the electrical potential held in that location decays, and is lost.

Target = 8,192,000 (1024K * 8) monitored states every 500ns.

8,192,000 states * $5^{-7}$ = $4.096 \times 10^{14}$ bit changes per second.

1 It is pointless to construct models of infinitely powerful nonphysical information processing systems which have transtemporal, and acausal capabilities. Postulating such systems cannot help advance our understanding, since such a system would have infinite capabilities.

2 Assuming no transtemporal activity takes place.

3 Current biological theory is that small networks of cells form the nearest approximation to term 'processor' as used in this context. This theory assumes these processors are capable of massive amounts of parrellellism, so our calculations must take account of this.

470
Next we can calculate the processing capacity (assuming instant information input). Assuming $10^{10}$ processors with a speed of 1 millisecond that gives

$10^{10}$ processors with capacity of 1 millisecond = $1^{14}$ actions per second

The following equation can be derived from standard information theory as giving the utilization of the total processing capacity of a system

$$\frac{A}{P_c} = U_p$$

where $A$ is the average activity, $P_c$ is the processing capacity, and $U_p$ gives the percentage utilization this task will cause the processor(s).

average activity / processing capacity = utilization of processor

As an example, if a system had to monitor 500 events a second and was capable of processing 1000 then this would give us:

$\frac{500}{1000} = .5$ or 50% capacity.

In the case of the biological system this becomes

$\frac{10^{14}}{1^{14}} = 10.$

or 100% over maximum capacity.

Therefore we can see that just the task of monitoring the bit changes in a relatively slow 1 megabyte RAM board there is an order of magnitude between the biological systems capabilities, and those that would be required to achieve successful monitoring of the bits in the RAM.

---

4 It will in fact be less due to overlapping discharges, and recovery time taken before a cell can re-discharge, but we can use it as a theoretical upper limit.
Appendix 15. Diagnosis of Potential Anomalies.

Before starting the experiments it was necessary to create a classification system for the possible malfunctions or events that could occur when investigating machine malfunctions. This allows the events to be analysed at some later point in time. This classification system complements the AHCI classification system outlined in chapter 4. It splits events into categories of Software, Operator, Environmental and Hardware respectively. This will be found to cope with any AHCI event which occurs in an experimental situation. Each sub-category is further split into degrees of severity, the first being totally explained, and the last being totally unexplained. If any events fall into this last category it merely means the user of the category system has insufficient information upon which to base a judgement, and should not be taken as being proof that the event was paranormal 1.

Software Based

An IRSM {Identifiable Repeatable Software Malfunction}.

These are easily spotted software bugs, their causality can be determined, and replicated.

An INSM {Identifiable Non-repeatable Software Malfunction}.

These are often caused by an interaction between two events, which rarely coincide. Steve Richards's (see chapter 1) anecdote, where he and one other user were the victims of a computer failure on an '1108' system fell into this category. Often these are caused by some form of deadlock2 between two executing programs. Since the events which caused the anomaly rarely coincide, it is often difficult to replicate such an effect.

An URSM {Un-identifiable Repeating Software Malfunction}.

These events are fortunately rare, since in computing terms they are among the most infuriating for the user. That the causality is unidentifiable probably means that the anomaly is tied into an interaction between the operating system and the application being run.

An UNSM {Un-identifiable Non-repeating Software Malfunction}.

These are probably the cause of many false attributions of a paranormal machine anomaly. Since the event only occurred once (it is non-repeatable), the task of attributing causality to it is harder.

---

1 In connection with this comment Macleod (personal communication, 1989) has raised the important issue of what would be acceptable as proof of a paranormal happening. The answer to this depends upon the reader's belief system, and criteria of judgement. The author would probably define it as a reliably replicatable effect which defied explanation within current scientific knowledge.

2 Known in software engineering as a 'deadly embrace'.

472
**Sub-Categories of Human Errors - 'Slips'**

<table>
<thead>
<tr>
<th>Operator/Subject Based</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Capture Error</strong>: When a more powerful motor program switches into the running of another.</td>
</tr>
<tr>
<td><strong>Description Error</strong>: Perform the correct actions with the wrong objects.</td>
</tr>
<tr>
<td><strong>Data Driven Error</strong>: When an automatic action gets fed an on-going action sequence, causing behaviour that was not intended.</td>
</tr>
<tr>
<td><strong>Associative Activation Error</strong>: Where associations among thoughts and ideas take place (Freudian Slip).</td>
</tr>
<tr>
<td><strong>Loss of Activation Error</strong>: Forgetting in mid action why one is performing the action.</td>
</tr>
<tr>
<td><strong>Mode Error</strong>: Occurs when devices have different modes and actions which are appropriate in one and have different meanings in others.</td>
</tr>
</tbody>
</table>

Figure 134. Types of Human Error.

---

**An IROE (Identifiable Repeatable Operator Error).**

This category is where the operator is making a consistent error in their actions, which they are unable to discriminate themselves. In the first experimental subject population one of the subjects repeatedly claimed the system jammed whenever the author was not nearby. On investigation it was found that the subject was pressing the PrtSc button (which causes the operating system to print the screen to its printer), instead of the return button. However the subject was unaware that she was doing this since she was a touch typist, and paid more attention to her actions than an experimenter was in her immediate vicinity.

**An INOE (Identifiable Non-repeatable Operator Error).**

This class of operator error is where the user has obviously performed the wrong actions\(^3\), which either they do not recall, or they deny performing. The latter was suspected to be most frequent with first time computer users when they thought they had made 'an expensive' mistake.

**An UROE (Un-identifiable Repeating Operator Error).**

Fortunately this is a very rare category, and would probably be an indication that the experimenter was unfamiliar with the operation of the system which the subjects were using. Fortunately this would have been unlikely in the study described in this dissertation since the experimenter had sufficient knowledge of the target computer system.

**An UNOE (Un-identifiable Non-repeating Operator Error).**

This would be where both the experimenter and subject are unaware of the cause of an error.

**Environmental Based**

These effects will be tied to at least one of the other sub-categories. Environmental effects are only manifested by their effects upon some other sub-category of the Associated Implicit Environment (AIE) (see chapter 3).

---

\(^3\) An example would be where the user has obviously issued the 'quit' command (since the system log shows this), but the subject denies having done so.
An IREA {Identifiable Repeatable Environmental Affect}.

This would be where some aspect of the AIE (see chapter 4) had an effect upon the performance of the system. Excessive degrees of heat or cold would be good examples of these kinds of effects.

An INEA {Identifiable Non-repeatable Environmental Affect}.

This kind of effect would be most likely to be associated with some unique feature(s), or combinations of environmental factors, which could not be duplicated exactly.

An UREA {Un-identifiable Repeating Environmental Affect}.

Events which fall within this category are extremely frustrating, and are likely to be associated with an environmental effect which has been overlooked. Many anomalies may well be covered by this category.

An UNEA {Un-identifiable Non-repeating Environmental Affect}.

Events within these last two categories would indicate that some aspect of the experiment was uncontrolled. If experimenters repeatedly found effects within this category they should consider reevaluating their experimental method.

Hardware Based

An IRHM {Identifiable Repeatable Hardware Malfunction}.

Events within this category are typically called 'hard' faults. They usually mean that some component(s) require replacement or repair.

An INHM {Identifiable Non-repeatable Hardware Malfunction}.

These events are usually caused by the in-harmonious interactions of multiple hardware components. As with the non-repeatable software errors (above) these form some of the most irritating aspects of computing.

An URHM {Un-identifiable Repeating Hardware Malfunction}.

These are caused by poor hardware design, construction, or assembly.

An UNHM {Un-identifiable Non-repeating Hardware Malfunction}.

These are surprisingly frequent in computing. They are associated with the UM problems discussed in chapter 4.

The experimental protocol demanded that if a hardware anomaly occurred on a system during an experimental session, the experimenter would dismantle the machine until he found the fault. On the one occasion this became necessary, the fault was not detected, and was found to have vanished when the

---

4 The computer system (affectionately known as 'Aristotle') which was constructed by the author to write up this dissertation has certain of these features, making the system interesting to use. The author has noted that such unpredictability can cause computer systems to appear to have a personality.
machine was reassembled and tested. It was assumed that some dust lodged in the circuits had caused this effect.
Appendix 16. Implementation problems encountered before the pilot experiment.

The experimental system described in chapters 2, 4 and 5 was implemented on the IBM XT286 computer systems at the psychology department in Edinburgh University. However the three main studies described within this dissertation used the Amstrad 1512 PC as their host system. Both computers are what is termed IBM compatible, that is they conformed to the then industry standard operating system and hardware conventions. The author had selected Digital Research Inc’s (DRI) Graphical Environment Manager (GEM) to run the WIMP version of the interface. This was because GEM runs on both the IBM and Amstrad systems, and the author assumed that they would be compatible enough to run the same applications. Indeed to try to ensure the transfer from the IBM system to the Amstrad had the minimum number of problems the Amstrad version of the GEM programmer’s tool kit had been used in the construction of the system. However the GEM version on the Amstrad 1512 machines ran under DRI’s proprietary operating system called DOS+. Unfortunately DOS+ is a CPM media compatible operating system, this meant that the KMDB system which had been developed and compiled to run under MSDOS 3.2 crashed every time it made a MSDOS function call. This made the running the KMDB system on the 1512s a rather brief experience. The obvious solution to this problem was to run MSDOS 3.2 on the Amstrad systems and use the MSDOS version of GEM. Unfortunately the Amstrad systems use a proprietary (non-standard) mouse driving system, which had been implicitly handled by DOS+, but was unrecognized by MSDOS 3.2. This meant that the GEM version of the operating system would run, but without the mouse which was unacceptable. After detailed inquiries directed to Amstrad, DRI, and Prospero Software, an Amstrad MSDOS 3.2 mouse was found, and the GEM application successively ran on the target system. The command line system, which did not have to rely upon the software support of a windowing environment, had no implementation problems since it could directly address the computer’s hardware to perform any actions it required. The independence of the windowing support also made the command line system take only one disc (in comparison to the GEM version) and take 2 and be half the size of the GEM version. It would have been considerable faster as well, but this was controlled by certain ProPascal compiler option settings, so both interfaces performed at the same rate. This shows that the system overhead involved in making a computer provide a graphical interface are far greater than those required for a comparable command line system. However it is a basic premise of computer science that anything which can be implemented in software can also be implemented in hardware, so it would be possible to reduce these overheads by the use of specific hardware support for windowing environments.

---

1. An application is the name given to GEM programs.
2. Amstrad had made a commercial agreement with DRI to try to boost DOS+ sales by bundling it with the Amstrad 1512. In return DRI had supplied GEM to Amstrad at a reduced rate.
3. The mouse interface is patched under low level software control from the mouse through the Non-Volatile-RAM (NVR) (see chapter 4) so that it appears as input from the keyboard. The Amstrad mouse is supposed to be Microsoft Bus mouse™ compatible, but this is only when used with a special software driver. Both Amstrad and DRI were reluctant to provide details of this method, since it would mean 3rd parties could run non DOS+ versions of GEM on the Amstrad 1512, invalidating the existing trade agreement between them.
Appendix 17. Development of a possible theoretical background to function enhancement of computer systems.

During the process of this dissertation informal models and observations were developed which formed the rationale behind much of the design of the final experiment. They are also included in the hope that they might be of value to future researchers in this area, even if they serve only as warnings against such speculation! These speculations are all based upon the assumption that 'psi' exists. This was done to make the discussion easier to follow, and less time consuming.

MLPs and FLPs.

From talks to the teaching staff at Napier Polytechnic about those few subjects who displayed apparent MLP 'abilities' (particularly the 'star' MLP in experiment one) one consistent factor emerged. This was that such subjects tended to have poor concentration on whatever tasks they were doing. For example the 'star' MLP in experiment I spent most of her time complaining about the task and staring into space. Her senses were not being applied to the computer based operations she was undertaking, instead she was directing her mind and imagination to anything but the task at hand. Having poor concentration would make it likely that the person concerned would perform badly at most activities requiring coordination, and this might result in them having a poor self image with regard to machine interactions. This would not be due to any paranormal action, but merely because they would not be paying full attention to what they were doing. However it is not the case that paying full attention makes you a 'psi' FLP, because those subjects who paid full attention to what they were doing produced results within MCE. Instead it could be that, if psi exists, being engrossed in some deeply physical/sensory aspect of activity (involving the established senses and mind) reduces the possibility of psi activity. This would mean that most people would not be using the faculty of 'psi' during the course of daily life. If this was true one would expect most 'psi' to appear spontaneously when the percipient was not engaged in some activity that required concentration or sensory monitoring/coordination. The apparent lack of coordination in MLPs which was found in the GEM UEICS data could be due to the lack of concentration and monitoring of sensory feedback. Using this model of 'psi' functioning one would expect that people who predominately used imagination, or lead a mental life which was not task oriented (day dreamers, even some forms of psychotics/neurotics) would have a lot of 'psychic experiences'. Also one would expect to find that traditional methods of using psi would require some abstract/random form to use the imagination upon and some very fixed task or ritual. The use of a ritual where all the actions are fixed, and memorized might be to free the mind from the task of sensory monitoring. In the field of parapsychology we would expect experiments that freed the subject from their normal monitoring and concentration, to work best (for example the Ganzfeld, OBE, and Non-striving).

An example of such a computer assisted study to explore such ideas might be as follows. One could create a computer art package, and introduce artists (amateurs, professionals and a control group) to the possibility of art on a computer. The noise (paint brush) movements could have 'smart noise' built in to try to anticipate the artists next required movement. This noise could be scaled down to a few pixels and therefore be just noticeable to the subject. Alternatively it could also be made so small that the subject did not notice it, in order to see if this made any difference. The amount of noise involved in the drawing process could be varied by the use of a SSM. Using the mental faculty model of 'psi' (outlined above), one would expect beginners at most tasks to not use 'psi' at all due to the concentration and physical coordination required (see the later MLP/FLP comment for anticipated exclusions to this prediction). In contrast competent users of the arts package would be expected to be able to enter an internal 'free mental state' to make use of the available FAA. It might be possible to get subject/operators to go into this FMS with a little training. This model presumes that 'psi' is very closely linked to the mental faculties of

1 This would apply to most tasks except those which involved no coordination, and lots of imagination. An example of these kinds of tasks would be gambling, or Parapsychological studies. In these cases beginners would be expected to do better than those with experience, due to the freedom from mental restriction accompanying their inexperience. More experience would tend to make subjects concentrate on replicating their internal states, with a subsequent increase in sensory monitoring.
imagination, belief, and need. Exercises which developed imagination and the will power would help the emergence of psi under FMS such as daydreaming, or meditation.

The communality between MLPs and FLPs.

Within this model both MLPs and FLPs share some common characteristics, the only difference is they way in which they are applied. This model proposes that the communality between the two types of function alteration is that the subject is not fully concentrating upon the task (s)he is undertaking. They are intermittently entering some form of FMS which is not directly associated with the task they are undertaking. The difference is in the form which the associated mental activity accompanying the FMS takes. In this model one would predict that the MLP is worrying (maybe unconsciously), and becoming stressed about the 'bad' things they associate with their interactions with machines and computers. In contrast the FLP may be associating positive thought with their interactions with computers. In both cases the subjects are not fully directing their mental faculties to the task at hand².

However, as we have already noted there are some normal artifacts at work here, which complicate the investigators task. The lack of concentration of the MLP would be evident as lack of coordination (see the results from experiment 1), and this would make them more error prone, and less rapid to learn in new situations that involve machines/computers. In contrast the FLP would already know about the system(s) being used and would have committed all the details of the machine's use to long term memory.

This could be an important factor which allows the FLP to enter a FMS, and use the faculty of 'psi' to enhance the apparent performance of the system. It is very important to realize that, within this model, the FLP is unlikely to use 'psi' on a new or unfamiliar piece of equipment due to the amount of task based concentration required. Only after the FLP has committed the system details to memory, so they can enter a FMS without affecting their performance, will function linked activity (FLA) have a chance to occur. In contrast a MLP will show most malfunction linked activity (MLA) when first using a new system. This is because their anxiety will be strongest on the first use of some equipment. This model would predict that unless the MLP has an extremely bad session (thus increasing session anxiety) the subsequent uses of the system would show less MLA. An exception would be of course under conditions of extreme stress, or anxiety when the MLP will be most likely to cause extreme MLA. In this model even the FLP (if stressed enough) will cause MLA, since imagination would play a large role in these events. This would be how this model explained the Murphy's law of frequent breakdowns when ostensible FLPs are under stress. The model describes these stressful and non-stressful periods under the terms of Function and Malfunction Linked Time Periods (FLTP & MLTP).

Definitions developed from these speculations on Function Alteration Activity.

Activity which alters performance or function (which I have termed Function Alteration Activity) can be taken up as Function Linked Activity (FLA) Malfunction Linked Activity (MLA) or Baseline Chance Activity (BCA), respectively. Some contemporary parapsychologists (Braud 1980) propose that the total Function Alteration Potential (FAP) (which is the lability or 'slop' in the system), is the area most likely to be influenced by FAA.

---

2 This would tie in with the 'release of effort effect' often associated with PK work. This is where the subject switches to a FMS after establishing a need (willing the outcome), and using the imagination to direct PK to the target.

3 The speed with which a subject learns or adapts to new situations may be an indication of these FLP/MLP tendencies. It was for this reason that questions relating to the ease with which subjects adapt to new circumstances were included in version 3 of the TAQ.
### Index

#### A

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abler &amp; Sedlacek</td>
<td>179</td>
</tr>
<tr>
<td>Abler, &amp; Sedlacek</td>
<td>183</td>
</tr>
<tr>
<td>Access Control Facility</td>
<td>2, 154</td>
</tr>
<tr>
<td>AHCI classification system</td>
<td>472</td>
</tr>
<tr>
<td>Ahl</td>
<td>197</td>
</tr>
<tr>
<td>AI</td>
<td>197</td>
</tr>
<tr>
<td>Alcock</td>
<td>6, 8, 98</td>
</tr>
<tr>
<td>ALGOL</td>
<td>188</td>
</tr>
<tr>
<td>Altered States of Consciousness</td>
<td>6</td>
</tr>
<tr>
<td>Alvey Programme</td>
<td>209</td>
</tr>
<tr>
<td>Andre</td>
<td>46, 69</td>
</tr>
<tr>
<td>ANDREW</td>
<td>69</td>
</tr>
<tr>
<td>Anomalous Human Computer Interaction</td>
<td>31</td>
</tr>
<tr>
<td>Apple Computer Inc</td>
<td>135</td>
</tr>
<tr>
<td>Apple Computer Inc</td>
<td>135</td>
</tr>
<tr>
<td>Apple Macintosh</td>
<td>176</td>
</tr>
<tr>
<td>Apple-Adept</td>
<td>45</td>
</tr>
<tr>
<td>ARC</td>
<td>113</td>
</tr>
<tr>
<td>Arndt, Cleveenger, &amp; Meiskey</td>
<td>205</td>
</tr>
<tr>
<td>Arthur's seat</td>
<td>255</td>
</tr>
<tr>
<td>ASCII</td>
<td>111</td>
</tr>
<tr>
<td>Ashcroft-Nowicki</td>
<td>22, 23</td>
</tr>
<tr>
<td>Associated Implicit Environment</td>
<td>134</td>
</tr>
<tr>
<td>Associated Implicit Software Environment</td>
<td>160</td>
</tr>
<tr>
<td>Associative Activation Errors</td>
<td>473</td>
</tr>
<tr>
<td>ATTCQ</td>
<td>202</td>
</tr>
<tr>
<td>ATitudes Towards Computers Questionaire</td>
<td>202</td>
</tr>
<tr>
<td>Austin</td>
<td>185</td>
</tr>
<tr>
<td>Ayeroff &amp; Abelson</td>
<td>14</td>
</tr>
</tbody>
</table>

#### B

<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babbage's Analytic Engine</td>
<td>172</td>
</tr>
<tr>
<td>Babu</td>
<td>49</td>
</tr>
<tr>
<td>Bacon</td>
<td>6</td>
</tr>
<tr>
<td>Bakan</td>
<td>11, 12</td>
</tr>
<tr>
<td>Baker, Holding, &amp; Loeb</td>
<td>180</td>
</tr>
<tr>
<td>Barber</td>
<td>103</td>
</tr>
<tr>
<td>Barnard</td>
<td>173</td>
</tr>
<tr>
<td>Barnard et al</td>
<td>300</td>
</tr>
<tr>
<td>Barnard et al,</td>
<td>190</td>
</tr>
<tr>
<td>Baseline Chance Activity</td>
<td>478</td>
</tr>
<tr>
<td>BASIC</td>
<td>64</td>
</tr>
<tr>
<td>Batcheldor</td>
<td>18, 25, 380</td>
</tr>
<tr>
<td>Bath</td>
<td>11</td>
</tr>
<tr>
<td>Beller</td>
<td>221</td>
</tr>
<tr>
<td>BCA</td>
<td>478</td>
</tr>
<tr>
<td>Beginners All-Purpose Symbolic Instruction Code</td>
<td>51</td>
</tr>
<tr>
<td>Beginners All Purpose Symbolic Instruction Code</td>
<td>51</td>
</tr>
<tr>
<td>Beloff</td>
<td>6, 7, 227</td>
</tr>
<tr>
<td>Beloff &amp; Evans</td>
<td>26</td>
</tr>
<tr>
<td>Beloff &amp; Regan</td>
<td>36, 44</td>
</tr>
<tr>
<td>Beloff &amp; Evans,</td>
<td>54</td>
</tr>
<tr>
<td>Benassi, Sweeney, &amp; Dreyer</td>
<td>15</td>
</tr>
<tr>
<td>Benyon</td>
<td>177</td>
</tr>
<tr>
<td>Benyon, Innocent &amp; Murray</td>
<td>179</td>
</tr>
<tr>
<td>Berger &amp; Honorton</td>
<td>38, 64</td>
</tr>
<tr>
<td>Berger, Schechter, &amp; Honorton</td>
<td>64</td>
</tr>
<tr>
<td>Beringer, Peterson</td>
<td>177</td>
</tr>
<tr>
<td>Bertini, Lewis, &amp; Witkin</td>
<td>235</td>
</tr>
<tr>
<td>Bevan et al</td>
<td>173</td>
</tr>
<tr>
<td>Bierman,</td>
<td>65, 66</td>
</tr>
<tr>
<td>Bierman &amp; Houtkooper</td>
<td>56</td>
</tr>
<tr>
<td>Bierman &amp; Wout</td>
<td>59</td>
</tr>
<tr>
<td>Binary search</td>
<td>221</td>
</tr>
<tr>
<td>Binary tree</td>
<td>221</td>
</tr>
<tr>
<td>Bird &amp; Chung</td>
<td>200</td>
</tr>
<tr>
<td>Black &amp; Sebrechts</td>
<td>173</td>
</tr>
<tr>
<td>Blackmore,</td>
<td>65</td>
</tr>
<tr>
<td>Blackmore &amp; Trosclianko</td>
<td>10</td>
</tr>
<tr>
<td>Blackmore, Trosclianko</td>
<td>16</td>
</tr>
<tr>
<td>Block</td>
<td>179</td>
</tr>
<tr>
<td>Bloom</td>
<td>199</td>
</tr>
<tr>
<td>Bouchard, &amp; McGeoo</td>
<td>180</td>
</tr>
<tr>
<td>BRAIN</td>
<td>237</td>
</tr>
<tr>
<td>Brain Wave Driven Computer Interface</td>
<td>178</td>
</tr>
<tr>
<td>Braun,</td>
<td>82, 86, 87, 231, 235, 478</td>
</tr>
<tr>
<td>Braun &amp; Braud</td>
<td>63, 64</td>
</tr>
<tr>
<td>Braun &amp; Jackson</td>
<td>229</td>
</tr>
<tr>
<td>Braun &amp; Kirk</td>
<td>82</td>
</tr>
<tr>
<td>Braun &amp; Schultz</td>
<td>229</td>
</tr>
<tr>
<td>Braun &amp; Schroeter</td>
<td>67</td>
</tr>
<tr>
<td>Braun, Wood &amp; Braud</td>
<td>235</td>
</tr>
<tr>
<td>Braude</td>
<td>88</td>
</tr>
<tr>
<td>Braun, Goupil, Giroux, &amp; Chagnon</td>
<td>184</td>
</tr>
<tr>
<td>Braun, Goupil, Giroux, &amp; Chagnon</td>
<td>181</td>
</tr>
<tr>
<td>Braun, Goupil, Giroux, &amp; Chagnon,</td>
<td>202</td>
</tr>
<tr>
<td>Breederveld &amp; Jacobs</td>
<td>227</td>
</tr>
<tr>
<td>Breederveld &amp; Jacobs,</td>
<td>227</td>
</tr>
<tr>
<td>Brer &amp; Tyninski</td>
<td>216, 217</td>
</tr>
<tr>
<td>Brier &amp; Tyninski</td>
<td>219</td>
</tr>
<tr>
<td>British Standards Institute</td>
<td>153</td>
</tr>
<tr>
<td>British Standards Institute,</td>
<td>116</td>
</tr>
<tr>
<td>Bronowski,</td>
<td>5</td>
</tr>
<tr>
<td>Broughton,</td>
<td>24, 26, 32, 65, 78, 105</td>
</tr>
<tr>
<td>Broughton &amp; Millar</td>
<td>76</td>
</tr>
<tr>
<td>Broughton &amp; Perlstrom</td>
<td>72, 73</td>
</tr>
<tr>
<td>Broughton, Millar &amp; Johnson</td>
<td>79</td>
</tr>
<tr>
<td>Broverman, Klaiber, Kobayashi, &amp; Vogel</td>
<td>180</td>
</tr>
<tr>
<td>Buffery, &amp; Grey</td>
<td>180</td>
</tr>
</tbody>
</table>

#### C

<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cadoret</td>
<td>215</td>
</tr>
<tr>
<td>CAIN</td>
<td>205</td>
</tr>
<tr>
<td>CAL packages</td>
<td>403</td>
</tr>
<tr>
<td>Campbell,</td>
<td>113</td>
</tr>
<tr>
<td>Campbell, &amp; Dobson</td>
<td>206</td>
</tr>
<tr>
<td>Camstra</td>
<td>78</td>
</tr>
<tr>
<td>Canadian Broadcasting Corporation</td>
<td>41</td>
</tr>
<tr>
<td>Caplan</td>
<td>179</td>
</tr>
<tr>
<td>Caplan, Macpherson, &amp; Tobin</td>
<td>179</td>
</tr>
<tr>
<td>Capture errors</td>
<td>473</td>
</tr>
<tr>
<td>CARAQ</td>
<td>206</td>
</tr>
<tr>
<td>Card et al</td>
<td>170, 171</td>
</tr>
<tr>
<td>Card, English, &amp; Burr</td>
<td>177</td>
</tr>
<tr>
<td>Carpenter,</td>
<td>46, 217, 218, 236</td>
</tr>
<tr>
<td>Carroll</td>
<td>190</td>
</tr>
<tr>
<td>CARS</td>
<td>207</td>
</tr>
<tr>
<td>CAS</td>
<td>204</td>
</tr>
<tr>
<td>Chari</td>
<td>88</td>
</tr>
<tr>
<td>Chauvin, &amp; Genthon</td>
<td>54</td>
</tr>
<tr>
<td>Chen</td>
<td>183, 200</td>
</tr>
<tr>
<td>Cheney, &amp; Dickson</td>
<td>202</td>
</tr>
<tr>
<td>Chevako</td>
<td>149</td>
</tr>
<tr>
<td>Circles</td>
<td>39</td>
</tr>
<tr>
<td>CIS</td>
<td>229</td>
</tr>
<tr>
<td>Clark</td>
<td>216</td>
</tr>
<tr>
<td>Cleve Baxter</td>
<td>82</td>
</tr>
<tr>
<td>CMOS</td>
<td>382</td>
</tr>
<tr>
<td>COBOL</td>
<td>188</td>
</tr>
<tr>
<td>Coffey</td>
<td>155</td>
</tr>
<tr>
<td>Coleman</td>
<td>200</td>
</tr>
<tr>
<td>Collis &amp; Ollila,</td>
<td>202</td>
</tr>
<tr>
<td>Collis, &amp; Ollila,</td>
<td>185</td>
</tr>
<tr>
<td>Collis, &amp; Williams</td>
<td>206</td>
</tr>
<tr>
<td>Command Line Macho</td>
<td>189</td>
</tr>
</tbody>
</table>
Heuristic Search Hypothesis, 220.
Hewlett Packard, 34, 218.
HI-FLP, 243.
HI-MLP, 243.
Hiscock, 179.
Hit-Rate, 247.
Hodges, 32, 470.
Hoeckstra, 19.
Honorton, 71.
Honorton & Barksdale, 75.
Honorton & Harper, 235.
Honorton & May, 52.
Honorton & Tremmel, 73.
Honorton, Barker, & Sondow, 64.
Hoppe, 54.
Houtkooper, 88.
Howard, & Smith, 204.
Hubbard, Bentley, & Pastucale, 8.
Huber, 185.
Hudgens, & Billingsley, 180.
Human Computer Interaction, 20, 30.
Hutchins, Hollan, & Norman, 280.
Hutt, 180.
Hypnosis, 226, 227.
IBM, 135.
IBM 3081 mainframe, 58.
IBM International test scoring machine, 41.
IBM Porta-Punch, 48.
IBM Yorktown Heights, 176.
Identifiable Non-repeatable Environmental Affect, 474.
Identifiable Non-repeatable Hardware Malfunction, 474.
Identifiable Non-repeatable Operator Error, 473.
Identifiable Non-repeatable Software Malfunction, 472.
Identifiable Repeatable Environmental Affect, 474.
Identifiable Repeatable Hardware Malfunction, 474.
Identifiable Repeatable Operator Error, 473.
Identifiable Repeatable Software Malfunction, 472.
IDS, 25, 94.
IDT, 123.
Index of preference, 215.
Information Theory, 213.
Intelligence Quotient, 12.
Interference, 213.
Interpersonal Trust Scale, 203.
Interrupt Descriptor Table, 123.
Intuitive Data Sorting, 93.
Irrational Beliefs Test, 17.
ISA, 54.
Isaacs, 8, 18, 25, 80.
Iseki & Sheideman, 191.
Iseki & Sheideman, 190.
ITS, 203.
Jacobs, 49, 51, 88.
Jahn, 233.
Jahn & Dunne, 24.
Jahn, Nelson & Dunne, 26, 58.
Jarke, et al., 191.
Johnson, 231.
Johnson & Hartwell, 87.
Johnston et al., 180.
Jones, Russell, & Nickell, 12.
Jordan & Stroup, 198.
Joseph, 121.
Kahn, 41.
Kahn & Neisser, 41.
Kahneaman, 199.
Kanathamani & Kelly, 213.
Kaplin, & Sedney 1980, 179.
Kautz, 235.
Kelly & Locke, 235.
Kendall & Babington-Smith, 24.
Kendall & Babington-Smith, 34, 40, 99.
Kennedy, 186, 219, 236.
Kerber, 203.
Kestenbaum, & Weinstein, 181, 202.
Klip, 26, 236.
Klonoff & Cambell, 202.
KMDB, 240, 241.
Knuth, 34, 221.
Koester, 23, 88.
Koslowski, Lazar, & Hoffman, 206.
Krause, 19.
Kraut et al, 173.
Kraut, Hanson & Faber, 190.
Kreitler, & Kreitler, 13.
Kruglanski & Ajzen, 12.
Kupka, 186.
Kurtz, 5.
Lability, 67.
Lafontaine, & Tredeau, 184.
LAN, 142.
Langer, 13, 16.
Langer & Roth, 14.
Leder et al., 173.
Ledgard, Whiteside, Singer, & Seymour, 190.
Lee, 196.
LeShan, 6.
Levy & Andre, 84.
Levy & Davis, 81.
Levy & Terry, 81.
Levy and Andre, 81.
Levy, Davis & Mayo, 81.
Lewis, 94.
Lewis & Payne, 34.
Lewis, & Cooney, 181.
Liebelt, McDonald Stone & Karat, 190.
Linear search, 221.
Linn, 184.
Locke & Shontz, 17.
Lockheed, & Frakt, 183.
Locus Of Control, 15.
Lorimer et al., 19.
Loss of Activation Errors, 473.
Lorwy, 51.
Loyd, Loyd & Gressard, 182.
Loyd, Loyd & Gressard, 204.
Lu, 8, 177.
Luchins, & Luchins, 180.
Lusk & Kernick, 185.
Meier, 199.
Messick, 185.
Mice, 177.
Michie, Turing & Good, 470.
MicroEmacs, 111.
Microsoft, 124.
Mihalasky, 42, 43.
Millar, 26, 32, 44, 65, 88, 105, 229, 236.
Millar & Broughton, 57.
Milton, 213.
Minnesota Satisfaction Questionnaire, 202.
Minsky, 197.
Mischo, 18.
MLA, 478.
MLE, 243.
MLP, 243.
MLTP, 478.
Mode Errors, 473.
Modular Communications Testing system, 37.
Moore, 206.
Morgan, 24, 105, 135, 189.
Morris, 8, 24, 29, 33, 50, 60, 63, 110, 150, 198, 210, 228, 229, 230, 232, 375.
Morris & Reilly, 74.
Morris, Garcia-Noriega, 59.
Morrison and Davis, 77, 91.
Morrison et al, 173.
Morrow, Prell, & McElroy, 200.
MSQ, 202.
Mulholland, 98.
Multiple Guessing Technique, 214.
Murray et al, 173.
Myers Briggs, 39.
Myoelectric Channel Driven Computer Interface, 178.

N
Nanko, 16, 17.
NASA, 178.
Nash, 25.
National Physical Laboratory, 382.
Natural Language Interfaces, 173.
Near Death Experiences, 17.
Nelson, 176, 280.
Nelson & Radin, 32.
Nelson, Dunne, & Jahn, 58.
Neppe, 20.
Newell & Simon) in the following way, 220.
Nickell, & Pinto, 205.
Nicol, 34, 40.
NLI, 173, 191.
Norman, 18, 21, 187, 473.
Norris, 19.
NPL, 382.
NSA, 144.
NUL: device, 268.

O
O'Connor, 5.
O'Quin, Kinsey, & Beery, 200.
Object errors, 247.
Ogden & Boyle, 191.
Omerod, 183.
OOBE, 477.
Oswald, 215.
Out of the Body Experience, 17.
Owen & Sparrow, 18.

P
PA, 243.
PA-FLP, 243.
PA-MLP, 243.
PA-MLP, 243.
PA-MLP, 243.
PA-MLP, 243.
PA-MLP, 243.
PA-MLP, 243.
Paci Man, 18.
Palmer, 6, 13, 103, 104.
Palmer & Kramer, 37.
Palmer & Weiner, 34, 40.
Parker, 44.
Parkson, 24, 41.
Pask & Scott, 183.
Paul Kammerer, 88.
Payne, 185.
PEAR, 58, 94.
Performance Altering Persons, 243.
Perlmutter & Montymane, 19.
PESP, 79.
Pestonjek et al, 19.
Peteruzzell, 146.
Pictu re World, 176.
PK, 242.
PK with Plants, 83.
PKARC, 113.
Placer, Morris, & Phillips, 37, 59, 64.
Plant Sheep-Goat Affect, 83.
Playfair, 18.
PMIR, 231.
Pratt, 6, 43.
Price, 88.
Princeton Engineering Anomalies Research group, 58.
PRNG, 218.
ProPascal, 476.
Pseudo Random Number Generator, 34.
Pseudo Random Sequence, 156.
Psi Ball, 61.
Psi Invaders, 39, 64.
PsiBall, 64.
PsiFi, 37, 51, 62, 71.
PsiLab, 38.
Psychophysical Research Laboratories, 38.
Psychosis, 235.
Pulos & Fisher, 182.
Puthoff, 49, 79, 218.
Puthoff & Targ, 24, 232.

R
Radiation Intelligence, 151.
Radin & Boswell, 54.
Radin, May & Thomson, 32, 93.
RAND, 46, 67, 99.
Random Event Generator, 24, 40.
Random Number Generator, 7, 25.
Raub, 202, 203.
Rayleigh wave, 138.
Reason, 473.
Reason & Mycielski, 473.
Redmayne, 22, 40.
Reichbart, 234.
Reiser, Ludwig, Saxe, & Wagner, 233.
Resnik & Lammers, 185.
Resource Authorisation and Control Faculty, 154.
Retrocausation, 88.
Reznikoff et al, 202.
Rezni koff, Holland, & Stroebel, 179.
Rhine, 25, 41, 103, 104, 208, 212, 213.
Rich, 179.
Richards, 22.
RIELCS, 203.
RNG, 25.
RNG-PK Control Trials, 240.
Roberts, 191.
Rohrer & Simonson, 198.
Rolex Oyster, 382.
Rosen, Sears, & Weil, 207.
Rosenthal, 19.
Ross, 11.
Roter's Internal-External Locus of Control Scale, 203.
Roulette, 216.
Royal Statistical Society, 148.
Rubin & Honorton, 228.
Rubin & Honorton, 467.
Rush, 32, 33, 88, 89, 105.
Rutkowski, 176, 280.
RV,231.
Ryzl,218,226,227,236.

Safety checks,269.
Sanford and Davis,66.
Sargent,34,40.
Satler,19.
Scapin,173.
Schechter, Barker, & Varvoglis,61.
Schechter, Barker, & Varvoglis,61.
Schild et al,176.
Schlitz & Gruber,232.
Schlitz & Haight,232.
Schmieder,18, 43, 216, 229,230.
Schmieder & McConnell,18.
Schmieder & Borchardt,49,51.
Schmieder & McConnell,18,83.
Schmidt,26,42,45,46,49,50,55,56,62,67,84,85,89,91,235,236.
Schmidt & Pantes,95.
Schmidt & Terry,70.
Scott,215.
Scott and Koohang,201.
Scriing,235.
SEA,113.
Sem-Jacobsen,178.
Semantic errors,247.
Shafer,49,52.
Shallis,7,22,100.
Shannon,202.
Sharpe,180.
Sheep,15.
Sheingold, Kane, & Endreweit,183.
Sheweder,10.
Shneiderman & Mayer,171.
Shneiderman & Mayer,170,277.
Siann, Durndell, Macleod, & Glissov,184.
Singer & Benassi,15.
Singer & Benassi,12.
Skeptical Software Monitor,220,223.
Slak & Shaffer,12.
SM,240.
Small & Weldon,191.
Smart Noise Applications,237.
Smart Noise Programming,224.
Smith,377.
Smith et al,176.

SNS,228,240,243.
Socioeconomic Class,181.
Soloway et al,170,171.
Souter & Davis,468.
Souter & Davis,153.
Spencer Brown,34,40,99,100.
Spotteswoode,232.
Sproull, Zubrow, Kiesler,200.
SRI,60,94,231.
SSM,243,247,250.
SSM FAA Measures,264.
Staidford,180.
Standardized Test of Computer Literacy,205.
Stanford,67.
Stanford,6,24,93.
Stanford & Fox,76.
Stanford & Kottoor,76.
Stanford, Zenhausern, Taylor & Dwyer,93.
Stats' Stats+,258.
Staub,19.
STCL,205.
Steingner & Colsher,13.
Step search,221.
Stevenson, Oram, & Markwick,110.
Stuart,215.
Studt,19.
Subjective Paranormal Experience,20.
Swann,37.
Symantec's Q&A,258.
Sympathetic nervous system,229.
Syntax errors,246.

T
Taddonio,18.
Taetsch,216,218.
Taetsch,216.
Talbert and Debes,63,92.
TAQ,239.
TAQ3,385.
Tag,233.
Targ & Hastings,20.
Targ & Hurt,48.
Targ & Tag,47.
Targ & Tart,47.
Targ and Puthoff,232.
Targ, Puthoff & May,232.
Tart,34,45,79,80.
Tart & Labeore,360.
Technology Attitudes Questionnaire,210.
Temperature,216,229.
Terry & Schmidt,90.
Thalbourne,20.
The B-Tree Artifact,466.
The computer game Poink!,72.
The Linear Congruential Method,34.
The Power Method,34.
Thouless,19,48,110,215.
Timm,104.
Tobacyk & Mitchell,17.
Transparency,176.
Treisman & Faulkner,13.
Troselj,6.
Turing,32.
Tversky & Kahneman,13.
Tyrell,24,40,97.

U
UEICS,239.
UEICS2,249.
UM,119.
Un-identifiable Non-repeating Environmental Affect,474.
Un-identifiable Non-repeating Hardware Malfunction,474.
Un-identifiable Non-repeating Operator Error,473.
Un-identifiable Non-repeating Software Malfunction,472.
Un-identifiable Repeating Environmental Affect,474.
Un-identifiable Repeating Hardware Malfunction,474.
Un-identifiable Repeating Operator Error,473.
Un-identifiable Repeating Software Malfunction,472.
Unfunction Alteration Activity,478.
Universal Machine,32,119.
University of Virginia,43.
User Based System Design,199.

V
Vallee,233.
Van der Veer, Tauber, Waern, & Van Muylwyk,185.
Van-Eekhout, & Rouse,18.
Vassy,49.
Very Large Scale Integration,130.
Vilenskaya,234.
Virtuality,176.
Vitulli,61.
Volition,39,64.
Von Lucadou,380.
Von Lucadou,26,213,236,360.

W
Wagman, 205.
Walker, 26, 28, 236.
WAN, 142.
Ward, 100.
Weinberg, 197, 199.
Weinberg & English, 197.
Weinberg & Fuerst, 197.
Weiner, 72, 73, 77.
Weiss, 11, 12.
West, 25, 98, 103, 213.
Whitney-roper, 201.
Wichmann & Hill, 467.
Wicker, Richardson, & Lambert, 17.
Wilder, Mackie, & Cooper, 182.
Wilder, Mackie, & Cooper, 202.
Wilson, 25, 32, 44, 96, 97, 105.
Wise, Robinson-Staveley, & Nelson, 21, 182.
Witken & Goodenough, 186.
Witken, Dyke, Paterson, Goodenough, & Karp, 186.
Witken, Dyke, & Oltman, 186.
Wong, 106.
WORMS, 110.
Write Once Read Many Times, 145.
WYSIWYG, 191.

X
Xerox Parc, 176.
Xerox Star, 176.

Y
Yu & Yu, 110.
Yulish, 32.

Z
Zahn, 11, 12.
Note to the reader.

Due to the large number of literature reviews required in this dissertation the references have been condensed into two columns per page. The references are supplied in the form suggested in the Edinburgh University guide 'Dissertation and report writing: Guidance notes on the use and acknowledgement of sources' (Library Guide 42). This is non-standard and takes form: Author, Title, Edition, No Vols., Publication details, Editors or Joint Authors, Name of Journal, Date, Page / Volume details, and Publication series. As recommended authors names have been left in the original forename/initial format as found in the source. Readers should also note that the index (see previous page) does not include authors who have been sited in footnotes.


2 Indentified as Bierman (1985a).
3 Indentified as Bierman (1985b).


Brier, R. M., And Tyminski, W. V. "Ps Application: Part II. The Majority Vote Technique Analyses And Observations."


Dubey, R. S. "Persistence, Sex Differences And Educational Performance." Asian Journal Of Psychology And Education. 1982 Vol 9(3) 24 29.


Fertsch, Danna Elizabeth. "Gender Differences In Computer Attitudes Among Middle School Age Students (equity)." Dissertation Abstracts (1985) Number = Aae8517700.


Giesler, Patric V. "Parapsychological Anthropology: II. A Multi Method Study Of Psi And Pk Related Processes In The Umbanda Ritual Trance Consultation." Journal Of The American Society For Psychical Research. (1985a) Apr Vol 79(2) 113 166.


Halpern12, D. “A Different Answer To The Question Do Sex Related Differences In Spatial Abilities Really Exist.”. American Psychologist, (1986b) 41, 1014 1015.


---

12 Identified as Halpern (1986b).
13 Identified as Halpern (1986a).


May", E. C. "Psi / A Physiology Coupled, Noise Driven Random Generator To Extend Pk Studies." NJ: Metuchen, NJ:


Mikov, D. "Personal Communication." 1988


Morgan, K. "Do We Need Rngs?" Un-published Paper Presented To The European Pa Conference In Holland. (1988).


Targ, E. Personal Communication. 1987


Thouless, R. H. "Additional Note On 'a Test Of Survival.'" 1st Proceedings Of The Society For Psychical Research, (1946) 9, 48, 253 263.


Van Muylwijk, Bert, Van Der Veer, Gerrit, And Waem, Yvonne. "On The Implications Of User Visibility In Open Systems: An Overview Of The Little We Know And Of The Lot We Have To Find Out." Behaviour And Information Technology. (1983) Oct Dec Vol 2(4) 315 326.

Varoglis, M. P., And Mccarthy, D. "Conscious Purposive Focus And PK: Ring Activity In Relation To Awareness, Task Orientation, And Feedback." The Journal Of The American Society For Psychical Research. (1986), 80, 1 29


