DEVELOPMENTS IN PREDICTIVE DISPLAYS
FOR
DISCRETE AND CONTINUOUS TASKS

by

PETER JAMES GOILLAU BSc MSc

A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of the Loughborough University of Technology

August 1978

Supervisors: Professor B. Shackel MA FBPsS
Lampros Laios MSc PhD
Department of Human Sciences

"Man is not a mechanism:

mechanisms are extensions of man."

(Kelley, 1968)

ACKNOWLEDGEMENTS

It has been said that no research project can ever be the sole work of one person. This maxim seems particularly apt in the case of the present thesis, as it reflects the help and encouragement I have received from many friends and colleagues since my arrival in Loughborough in 1975.

I owe thanks to my supervisors, Dr. Lampros Laios and Professor Brian Shackel, for initially proposing this area of research and obtaining the S.R.C. CASE (Co-operative Awards in Science and Engineering) Research Studentship under which the material in this thesis was sponsored. I am particularly grateful to Dr. Laios for his assistance in the day-to-day planning and running of experiments, and to Professor Shackel for his overall direction and his comments on the structure and layout of the thesis.

At Warren Spring I am grateful to Mr. P.H. Hammond for agreeing to provide collaborative research facilities in connection with the CASE Studentship, to Dr. Peter King for overseeing my efforts, to Algis Cininas for supervising all and running some of the process operators, to John Hunt for his modifications to the Argus computer system, and to Dr. Ian Umbers for agreeing to read the draft thesis and for his useful comments and advice on protocol analysis.

At Loughborough, I have benefited much from discussion with my colleagues in the HUSAT (Human Sciences and Advanced Technology) Research Group. I should especially like to acknowledge the help of Dick Penn who wrote the simulation programs developed in Chapters 5 and 6, and of Tony Oliver and Tim Folkard for their programming assistance. Tim Moore also contributed much in the way of useful discussion on process operator models.

A special word of thanks must go to Sue Jennings who typed the draft manuscripts and for her encouragement and faith, without which the project might never have been completed. I am also grateful to Carol Dawson who typed the final draft. Above all I owe thanks to all those people who gave up their time to act as subjects in the experiments.

CONTENTS

CONTENTS

Page

PART I - PREFACE, THEORY AND APPLICATIONS 1

PREFACE 2

Ch. 1 BACKGROUND TO PREDICTIVE DISPLAY RESEARCH 4

1. INTRODUCTION 5

2. INDUSTRIAL PROCESS CONTROL 5

2.1 Background 5
2.2 Discrete and continuous tasks 8
2.3 Process operator skills 10
2.3.1 Role of the process controller 10
2.3.2 Dynamic manual control 11
2.3.3 Planning and scheduling 13
2.3.4 Supervisory control of automatic processes 14
2.4 Summary 15

3. PREDICTION 16

3.1 Human prediction 16
3.1.1 Experimental studies 17
3.1.2 Comparison of human and mathematical prediction 20
3.1.3 Summary 22
3.2 Mathematical prediction 22
3.2.1 Linear prediction theory 23
3.2.2 'Fast-time' approach 24
3.2.3 Choice of predictive control model 25
3.2.4 Why include the man? 25

4. METHODOLOGY 27

4.1 Laboratory, simulation and real-world studies 27
4.1.1 Advantages and disadvantages 27
4.1.2 The need for field validation 28
4.1.3 'Part-simulation' 29
4.2 Objective and subjective performance measures 30
4.2.1 Objective measures 30
4.2.2 Subjective measures 31
4.2.3 A compromise solution 33

5. APPROACHES TO MODELLING THE HUMAN OPERATOR 34

5.1 Mathematical approaches 34
5.1.1 Linear control theory 34
5.1.2 Statistical decision theory 37
5.2 Internal model concept 41
5.2.1 Early considerations 42
5.2.2 Kelley's model 44
5.2.3 Verbal protocol approaches 47
5.2.4 Bainbridge's model 47
5.2.5 Rasmussen's model 49
5.2.6 Advantages and disadvantages of internal model approach 53

CONTENTS (continued)

Page

Ch. 2 PREDICTIVE DISPLAYS 55

1. INTRODUCTION 56

1.1 Background and early attempts to assist the operator 56
1.2 Predictive display operation 60

2. APPLICATIONS OF PREDICTIVE DISPLAYS 65

2.1 Discrete applications 66
2.1.1 Stateboards 66
2.1.2 Production scheduling 68
2.1.3 Air traffic control 72
2.1.4 Soaking pit scheduling 73
2.1.5 Summary 79
2.2 Continuous control applications 80
2.2.1 Piloted aircraft 81
2.2.2 Spacecraft 84
2.2.3 Remote control systems 87
2.2.4 Ocean going vessels 88
2.2.5 Process control applications 90
2.2.6 Summary 92

3. FACTORS AFFECTING PREDICTIVE DISPLAY PERFORMANCE 94

3.1 Input uncertainty 96
3.2 Prediction span 98
3.3 Prediction model fidelity 101
3.4 Process dynamics/response characteristics 104
3.5 Repetition rate and frequency of updating 106
3.6 Modes of control 107

4. REVIEW 111

PART II - EXPERIMENTAL PROGRAMME 113

INTRODUCTION TO EXPERIMENTAL PROGRAMME 114

1. AIMS 114

2. DESIGN AND ANALYSIS 115

Ch. 3 AN EXPERIMENT TO EVALUATE THE EFFECT OF VARYING TASK CHARACTERISTICS AND PREDICTIVE DISPLAY PARAMETERS IN A SIMULATED SCHEDULING TASK 118

1. OBJECT OF THE EXPERIMENT 119

2. METHOD 119

2.1 Description of steel manufacturing process 120
2.2 Scheduling problem 121
2.3 The display 122
2.4 The controls 126
2.5 Predictive and system modes of operation 130
2.6 Experimental design 130
2.7 Procedure 132
2.8 Subjects 133
2.9 Data Collection 133

3. RESULTS AND STATISTICS 136

CONTENTS (continued)

		Page
	4. DISCUSSION	144
	4.1 Scheduling errors	144
	4.2 Predictive activity	146
	4.3 Decision horizons	149
	4.4 Search patterns	149
	4.5 Relationship between scheduling performance and predictive activity, scheduling performance and decision horizon	151
	4.6 Specialist vs. non-specialist users	154
	4.7 Subjective comments	155
	5. CONCLUSIONS	157
Ch. 4	AN EXPERIMENT TO VALIDATE THE EFFECT OF VARYING PREDICTION SPAN IN A SCHEDULING TASK USING TEST DATA FROM AN OPERATIONAL JOB-SHOP	159
	1. OBJECT OF THE EXPERIMENT	160
	2. METHOD	160
	2.1 Description of the problem environment	161
	2.2 The simulation	162
	2.3 The task	164
	2.4 Experimental design	164
	2.5 Procedure	165
	2.6 Subjects	165
	2.7 Data Collection	165
	3. RESULTS AND STATISTICS	166
	4. DISCUSSION	169
	4.1 Performance measures	169
	4.2 Subjective comments	171
	4.3 Uncertain environments	172
	5. CONCLUSIONS	173
Ch. 5	A PRELIMINARY INVESTIGATION OF PREDICTIVE TECHNIQUES IN CONTINUOUS CHEMICAL PROCESS CONTROL	174
	1. OBJECT OF THE EXPERIMENT	175
	2. METHOD	175
	2.1 The simulation	175
	2.2 The task	178
	2.3 Experimental design	180
	2.4 Procedure	182
	2.5 Subjects	183
	2.6 Data collection	183
	3. RESULTS AND STATISTICS	183

CONTENTS (continued)

Page

4. DISCUSSION 183

4.1 Total product manufactured 183
4.2 Subjective comments 188

5. CONCLUSIONS 189

Ch. 6 AN EXPERIMENT TO EVALUATE THE EFFECT OF VARYING TASK CHARACTERISTICS AND PREDICTIVE DISPLAY PARAMETERS IN A SIMULATED CONTINUOUS DUAL-METER MONITORING AND CONTROL TASK 190

1. OBJECT OF THE EXPERIMENT 191

2. METHOD 191

2.1 The simulation 192
2.2 The task 192
2.3 Experimental design 197
2.4 Procedure 198
2.5 Subjects 199
2.6 Data collection 199

3. RESULTS AND STATISTICS 200

4. DISCUSSION 209

4.1 Time outside limits data 209
4.2 Interpretation of Taylor series data 210
4.3 Interpretation of Perfect predictor data 211
4.4 Control histograms 214
4.5 Display switching 216
4.6 Subjective comments and questionnaire analysis 217

5. CONCLUSIONS 219

Ch. 7 AN EXPERIMENT TO VALIDATE THE USE OF PREDICTIVE TECHNIQUES IN THE CONTROL OF A PART-SIMULATED, SEMI-BATCH CHEMICAL REACTOR IN AN INDUSTRIAL SETTING 221

1. OBJECT OF THE EXPERIMENT 222

2. METHOD 222

2.1 The 'batch kettle' plant 223
2.2 Process operator research at Warren Spring 227
2.3 Multipen predictive recorder (m.p.r.) 229
2.4 Experimental design 231
2.5 Procedure 231
2.6 Subjects 232
2.7 Data collection 232

3. RESULTS AND STATISTICS 234

4. DISCUSSION 241

4.1 Percentage conversion 241
4.2 Standard deviation around the predicted variable 241
4.3 Calculated profit 242
4.4 Operator comments 243

5. CONCLUSIONS 244

CONTENTS (continued)

Page

PART III - RESULTS SUMMARY, DISCUSSION AND CONCLUSIONS 246

Ch. 8 SUMMARY OF CONCLUSIONS FROM EXPERIMENTAL PROGRAMME 247

1. EXPERIMENT 1 (Chapter 3) 248
2. EXPERIMENT 2 (Chapter 4) 249
3. EXPERIMENT 3 (Chapter 5) 250
4. EXPERIMENT 4 (Chapter 6) 252
5. EXPERIMENT 5 (Chapter 7) 253

Ch. 9 DISCUSSION 255

1. OVERVIEW 256

2. BENEFITS OF PREDICTIVE DISPLAYS 258

2.1 Objective performance measures 258
2.2 User acceptance 261

3. FACTORS AFFECTING PREDICTIVE DISPLAY PERFORMANCE 262

3.1 Input uncertainty 262
3.2 Prediction span 265
3.3 Prediction model fidelity 269
3.4 Process dynamics/response characteristics 271
3.5 Repetition rate/frequency of updating 272
3.6 Mode of control 273

4. ROLE OF PREDICTION IN HUMAN CONTROL AND DECISION-MAKING 274

5. VERBAL PROTOCOLS 275

5.1 Discrete tasks 276
5.2 Continuous tasks 282

6. DEVELOPMENT OF A HIERARCHICAL, PREDICTIVE MODEL OF THE HUMAN OPERATOR 289

6.1 Relation between internal and external predictive models 289
6.2 Necessary components of a human operator model 291
6.3 Implications of the proposed model 295

Ch.10 CONCLUSIONS 299

1. RECOMMENDATIONS 300
2. SUGGESTED AREAS OF FUTURE WORK 302
3. EPILOGUE 303

REFERENCES 304

APPENDICES 345

PART I

PREFACE, THEORY AND APPLICATIONS

PREFACE

How does one foretell the future? Since earliest times the existence of astrologers, fortune tellers and the like has borne witness to mankind's quest for information about the future. Their methods and the global nature of their predictions are, of course, open to question. But at a more mundane (though none the less important) level, the advent of fast, low-cost, high-capacity digital computers coupled with developments in the art of modelling has meant that computer-based mathematical models of relatively complex systems can be used to provide tentative predictions about the future. Weather forecasting is a case in point. It will be shown subsequently that Man is fundamentally a predictive, anticipatory creature and that provision of accurate information about the future is necessary for him to cope effectively with his environment. In the field of Engineering, there has been much evidence - admittedly largely from simulation studies of military and vehicular systems - that provision of predictive information in the form of a so-called 'predictive display' facilitates control.

However, in spite of the overwhelmingly favourable evidence, there have been few real-world applications, few studies to recommend particular designs of predictive display in specific settings, and few attempts to unify the experimental findings within a common theoretical framework. It is to this gap in our knowledge that the present series of experiments is addressed. The nature of the project has meant that the examples chosen have been restricted to the area of industrial process control. This is not in itself a bad thing: Bainbridge (1972) for example, recommends this field as a fruitful area of psychological study, having problems representative of many other real-world situations. Specifically, this thesis sets out to answer questions such as: 'How do predictive displays based on continuous information differ from those using discrete, non-continuous information?' 'How does varying the parameters of the predictive model affect performance?' 'Are the results of laboratory studies borne out in the real world?' 'Are predictive displays cost-effective?' 'How do operators make use of the displays?' 'What are the implications for the ways in which human beings themselves structure information and anticipate consequences?'

The plan of the thesis is as follows: The introductory chapters review the literature pertaining to human prediction and predictive control models (Chapter 1), and to engineering aspects of predictive displays (Chapter 2). Chapter 3 describes a fundamental study of predictive display parameters in a laboratory scheduling task, Chapter 4 attempts to verify these findings using test data from an actual job shop scheduling problem. Chapter 5 branches into the area of continuous control with a pilot study of predictive displays in a laboratory simulated continuous stirred-tank chemical reactor. Chapter 6 uses the experience gained in the pilot study as the basis for a comprehensive study of predictive display parameters in a further laboratory study of a simplified dual-meter monitoring and control task, and Chapter 7 attempts to test the optimal design in a part-simulated semi-batch chemical reactor using real plant and experienced operators in an industrial setting. The results of the experimental programme are summarised for convenience in Chapter 8. Chapter 9 draws together the threads from the various experiments and discusses the findings in terms of a general hierarchical model of an operator's control and monitoring behaviour. Finally, Chapter 10 presents conclusions and recommendations from the programme of research, together with suggestions for further work.

CHAPTER 1

BACKGROUND TO PREDICTIVE DISPLAY RESEARCH

1. INTRODUCTION

To introduce the thesis, this Chapter outlines the theoretical background necessary for research into the area of computer-based aids for industrial process control. The chapter is divided into four sections. The first section describes the nature of process control tasks and process operator skills. The second section describes the role of prediction, both human and mathematical, in control. The next section discusses some methodological issues. The final section in this chapter reviews some theoretical approaches to modelling the control and decision-making behaviour of the human operator.

2. INDUSTRIAL PROCESS CONTROL

2.1 Background

Process industries such as electricity, gas, steel, and petro-chemicals represent a high degree of task development and skill sophistication for the human in the system, as their operators do not manipulate the materials directly but rather exert control over the process machines to convert input materials to the desired finished product. During the last 20 years there has been a growing awareness of the changing role of the human operator in process control systems, culminating in the industrial survey by Edwards and Lees (1973). The principal cause behind the change in man's role has been the advent and subsequent rapid growth of first analogue and later digital computer technology. In the last few years, this trend has manifested itself with the introduction of dedicated minicomputers for process control, linked to low-cost visual display units through which information may be quickly and effectively communicated to the operator.

It now seems that the growth in microprocessor technology will further increase the scope for process automation.

Before the 1950's the process control industry was largely dependent on the skill and experience of its operators. As automatic control systems increased in scope whilst at the same time decreasing in price, the process operator found himself progressively relegated to a monitor or supervisor of the process he once controlled manually. Labour reductions also meant that instead of manually controlling a few control loops, he was now in supervisory control of many such loops. Traditional control skills have given way to a new requirement of cognitive skills.

The era of 'unmanned plant' is still some way off however (de Jong and Köster, 1971); a point is reached beyond which automation is no longer cost-effective, or even possible. Process engineers have in the past tended to automate as far down the system as possible, leaving man to perform those decision and other functions which they found were impossible to quantify. Although the human is capable of performing some of the remaining functions, albeit with questionable fidelity, this can be seen to be an unsatisfactory approach, originating from Birmingham and Taylor's (1954) philosophy to the effect that man "is best when doing least". Much has been learnt from attempts to 'automate' or 'computerise' control and decision functions. In many cases severe problems have arisen and it has proved necessary to bring the man back into the system (Lees, 1974). Control systems engineers are at last beginning to realise that they must accept the challenge of engineering both computer-compatible and operator-acceptable man-computer interfaces, if the timely and successful implementation of a process computer system is to be realised (Dupuis, 1975).

On practical grounds there are several good reasons for retaining an operator within the control system and not resorting to complete automation:

1) Cost - man is cheap to install and his overhead costs are relatively low, whereas a complex control system represents a substantial investment in capital resources and support staff.
2) Flexibility - man comes as a ready-packaged, general purpose controller and problem solver. To program a machine to cope with every plant eventuality would be prohibitively time-consuming and expensive.
3) Safety - man is subject to a 'graceful' breakdown in performance, whereas a machine will just stop functioning. Man can also make use of indirect, hard-to-quantify plant cues, e.g. noises, smells, the 'feel' of the plant.
4) Practice - in a supervisory control system, if there is even the remotest possibility that man may be required in a back-up capacity, his control skills must be retained at a meaningful level through adequate practice if they are not to be lost through neglect. This implies that he must actively interact with the process for part of its operating time.

For these reasons, the trend in recent years has been towards integrated man-computer control systems which aim to make full use of man's superior judgemental and pattern recognition abilities, coupled to the computer's facility for fast, accurate numerical computation and routine decisions. Given current technology, the most effective system in many applications is an operator-controlled, computer-supported system rather than a totally computerised one (Lees, 1974). Licklider's (1960) notion of 'man-computer symbiosis' and Jordan's (1963) belief that man and computer were complementary sowed the seeds of this movement, and it now seems likely that future systems will be judged by the degree to which a balance exists between man and machine.

The predictive display concept, which forms the subject matter of this thesis, is seen as an extension of this philosophy and an example of optimising the interface between the process controller and his instrumentation. In either manual or supervisory control applications, the concept can be used to compensate for the operator's relative inability to predict complex responses, whilst still leaving him with a meaningful role to perform in the system by providing him with accurate and pertinent information.

2.2 Discrete and continuous tasks

At first sight a distinction may conveniently be drawn between discrete and continuous processes. In the former, distinct items and stages are evident and some scheduling of resources is entailed, whereas the latter are characterised by material flows which are unbroken In the context of industrial process control, examples of discrete systems might be production scheduling through a machine shop or the control of ingots through a steel plant soaking pit complex. Examples of continuous systems might include chemical plant and some of the newer steel production techniques. The discrete-continuous dichotomy is a natural division to make, particularly as the theoretical backgrounds associated with each type of process have in the past tended to be separate (see section 5, this chapter). Decision theory has developed independently of manual control theory.

However, most process control applications include discrete and continuous components, the proportion depending on the nature of the process. For example, the operator's job in the steel or electricity industries often has a much higher discrete (scheduling) content than in the heavy chemical industries (Edwards and Lees, 1973). The same authors (Edwards and Lees, op. cit., Chapter 5) describe the broad range of process operator tasks. The experimental literature of recent years has also begun to reflect the combination of scheduling and control skills which occurs in practice. The maximum power demand task of Bainbridge et al. (1968), analysed in Bainbridge (1972, 1974), is a case in point. The controller's task was to allocate power between five separate steel-making furnaces, each of which went through a sequence of continuous manufacturing stages with a total cycle time of 5 hours. Other examples of discrete-continuous systems have been cited by Pew et al. (1966), Thomas (1973), Cohen and Ferrell (1969).

Furthermore, a new theoretical framework, that of the 'internal model' (Kelley, 1968; Bainbridge, 1975a), equally applicable to discrete and continuous processes has recently found favour. It has long been established that the human operator functions as an intermittent correction servo (Craik, 1947, 1948). As Gregory (1970) has pointed out, from the human's viewpoint it is largely irrelevant whether the controlled process be discrete or continuous, since in either case he gathers information by sampling. Hence the operator's internal representation of the controlled process must be based on discontinuous data. So the distinction between discrete and continuous processes, which at first appeared so important, can now be seen to be less to. In subsequent chapters, an attempt will be made to investigate decision factors common to both areas and to develop a general 'internal model' theory of human decision-making and control behaviour.

2.3 Process operator skills

It is of interest to consider some of the recent studies of human operator skills in process control. As this area has been extensively reviewed by Edwards and Lees (1973), Bainbridge (1975a) and Umbers (1976) only the main findings will be summarised here. Many of the classic studies are reprinted collectively in Edwards and Lees (1974).

2.3.1 Role of the process controller

The process operator has a vital role to play in process control, since it is he who is responsible for achieving good control performance and meeting production targets. Bainbridge (1975a) has commented that the skills possessed by the experienced process operator are complex, encompassing a general knowledge of process behaviour and permitted control actions together with a structured overview of the current state and future behaviour of the process. Skill must be adaptable to different contexts. In a process control task, the operator has not only to identify the present state of the process, judge whether it is acceptable and if necessary adjust the control settings (perceptual and control skills); he also has to decide between different ways of scheduling resources and how to allocate his time (planning skills). The proportion of control to planning skills required will vary depending on, and during, the particular task. In many tasks dynamic control, whilst essential, is only a minor part of the operator's mental activity compared to the planning component (Beishon, 1969). Lees (1974) lists the functions an operator may be required to perform as goal formation, measurement, data processing and handling, monitoring, single variable control, sequential control, and other control, optimisation, communication, scheduling and manual operations, depending on the particular task. Basic cognitive operations can be summarised as perceptual judgements and dynamic control, plus the information processing and decision-making operations of calculation, judgement and prediction. These will be considered further.

2.3.2 Dynamic manual control

In dynamic manual control, perceptual skills are used to discriminate which aspects of the process output require attention, either directly from the plant using any one or a combination of the senses, or indirectly through remotely displayed information. Chart recorders, together with process meters, are currently the most common form of remotely displaying information, and seem useful for detecting plant malfunctions as well as for trend analysis (Anyakora and Lees, 1972; Attwood, 1970). It is known that process operators measure the process variables not in absolute but rather in relative terms (Crossman and Cooke, 1962), such as overlapping discrete categories with labels that include implications for action, e.g. 'on target', 'going outside limits', 'well above specification' etc,, (Bainbridge, 1971). Control skills have traditionally been studied through activity analysis - this method is however inadequate to reveal the mechanisms which determine the size and timing of control actions (this question is explored further in section 4.2). Laboratory studies of simple processes have in general shown that a naive operator's response to a step input can be divided into three phases; initial ballistic response, 'hunting' around the target value and maintenance on target. The second 'hunting' phase disappears with practice, indicating that the experienced controller does not control by simple feedback alone. For example, Crossman and Cooke (1962) found that an experienced controller in their water bath task could make accurate control adjustments in the absence of visual feedback.

The experienced operator knows the gains, lags and interactions of his process and so can make smooth changes in the process state, can maintain control in unusual conditions, or can diagnose and correct plant faults (Bainbridge, 1975a). The operator's so-called 'mental model' of his plant is built up over time through interacting with and controlling the process (Kelley, 1968). The more complex the plant characteristics, the longer it will take him to construct a representative mental model. Control skill can thus be seen to be related to the development of the mental model. An experienced operator can make use of his internal model to predict process behaviour, evaluate alternatives and take anticipatory action. The significance of controlling by this means rather than by direct feedback is that his uncertainty of the process state is reduced, he can sample less often, resulting in a lower workload and an increase in his available processing capacity. Crossman et al. (1964) showed that controllers sample less frequently with experience, they also sample less frequently when the process output is within its specified limits.

Umbers (1976) has commented that individual differences in control strategy are also present, even amongst experienced operators, and the effect of such differences is important in control. He attributes these to differences in personality, cognitive development and general level of skill between operators.

2.3.3 Planning and Scheduling

Considering the planning and organisation of control behaviour, Kelley (1968) has suggested a general framework of goal conception (prediction), goal selection (planning), sequence programming, and execution. The available evidence suggests that behaviour is organised at several levels of complexity, in some form of goal hierarchy. As goals can be achieved by several means, and any one routine can be adapted to serve in different contexts, it is unlikely to be a true hierarchy, in the 'tree structure' sense of the word, but rather a loosely structured, flexible organisation (Bainbridge, 1974). This notion will be considered in more detail in a later section. For the moment, the evidence from protocol studies suggests that the basic cognitive operations mentioned earlier are organised by the operator into pre-programmed routines within his mental model of the plant, which must also include an 'executive' routine to plan which routine is used in which context to form the 'overall' sequence (Beishon, 1969; Bainbridge, 1975a; Smith and Crabtree, 1975). A problem exists in that the executive routine must usually be inferred, since operators seem unaware of or unable to verbalise their reasons for choosing a particular strategy.

In a multivariable process the operator cannot check or control all the variables at once; instead it is necessary for him to divide his attention. From experience he knows the optimum timing not only for his control actions but also for checking the effects of such actions The consensus of opinion is that the experienced operator maintains a mental picture in his working memory of variables which are of interest to his control, storing ready-processed assessments of plant states rather than raw process data, and sampling to keep the picture up-to-date. Experimental data on short-term storage indicates that working capacity is limited to around 7 static items (Miller, 1956) or 2-3 running items (Yntema and Meuser, 1960; Olshavsky, 1971). Hence the advantage of a structured working memory is that more items can be remembered, and novel situations can be generalised within the existing classification.

2.3.4 Supervisory control of automatic processes

A high degree of skill, process knowledge and flexibility is also required by the supervisory controller who may be expected to adjust set points on an automatic controller, monitor its performance, and take over on the rare occasions when it develops a fault. The importance of this type of monitoring behaviour or supervisory control was recognised by a recent NATO Symposium on the subject (Sheridan and Johannsen, 1976). Many automated plants are in fact monitored by operators who once controlled the process manually, but who now act as machine minders and rarely interact with the plant except under emergency conditions, with a concomitant deterioration in their skill level. One of the operator's main monitoring functions on a plant is to prevent the development of situations which may become serious incidents, e.g. breakdowns or accidents. 'Incident avoidance' by the operator is considerably more effective than automatic protection systems whose main action is to shut down the plant when danger threatens and thereby incur expensive down time.

The task of the monitor or supervisory controller in responding to process faults can further be broken down into three functions, fault administration comprising failure detection, failure identification and corrective action (Curry and Gai, 1976). It is ironic that in most jobs the supervisory controller has greatly reduced opportunities to develop and maintain his control skills and his mental model, even though the crucial nature of his task demands greater proficiency in these very areas. Operators cannot retain control skill by watching an automatic controller at work (Brigham and Laios, 1975), so it seems that if the human is to play a meaningful supervisory role he should have the opportunity to control the process manually or should at least have access to a simulation of the process on which to practice. Kelley and Prosin (1972) have suggested that predictive displays may assist in the detection and correction of faults.

2.4 Summary

In summary, "the main feature of human behaviour in a control task is open loop control" (Umbers, 1976) for which prediction by the operator is a vital component. Kelley (1968) has commented that predictive information is so important to the operator that he uses every conceivable means for obtaining it. However, as Flowers (1978) points out, engineering models of human performance frequently overlook the importance of prediction. The ability to predict or anticipate plant response is an important skill distinguishing experienced from trainee operators. As van Heusden (1977) states, a control room operator usually practices 'management by expectations'. He does not change anything in the process control system unless one or more of the controlled variables are moving away from their target values. A skilled operator tries to avoid situations in which process variables become off-normal: he tries to predict the new situation in the case of no further actions, and reacts to diminish any undesirable deviation in process output. Bainbridge (1975a) has commented, however, that we know very little about controllers' predictive abilities: she suggests specialist tests involving covering the displays yet asking the controller to continue control or to assess the process state over time. In fact, a body of evidence does exist concerning prediction in applications other than process control, and this will be reviewed in the following section.

3. PREDICTION

3.1 Human prediction

It is now widely recognised that knowledge of future events and demands is a vital component of a wide range of human control and decision-making skills: driving a car, piloting an aircraft, catching a ball or crossing the road all depend to a greater or lesser extent on prediction. Visual events are seldom seen continuously - consider the scene from a moving car where extrapolation from one such event to another is a common task. And when supervising the operation of automated processes, such as those commonly found in the chemical industry, it is important to anticipate unsatisfactory conditions before they occur. The fact that the observer is frequently unaware of such processes implies that their locus is subconscious. There is even some conjecture that predictor circuits may exist at the physiological level (Milner, 1971). Only through anticipation of future consequences can naive, closed-loop 'feedback' control, where the human acts as an error detector to reduce differences between input and output in a servomechanism analogy, be transferred into skilled, open-loop 'feed-forward' control. As mentioned in the previous section, several workers (Crossman and Cooke, 1962; Cooke, 1965; Beishon, 1969; Kragt and Landeweerd, 1974; Brigham and Laios, 1975; Bainbridge, 1975a; Umbers, 1976) have commented on the substitution of open-loop for closed-loop control as skill develops. It is thought that this phenomenon corresponds to the learning of 'action sequences' which can be run directly from long-term memory, without the need for continual checking of both the process output and the operator's own behaviour (Bainbridge, 1978).

3.1.1 Experimental Studies

It is of interest to review some of the experimental work concerned with human prediction. One of the earliest recorded studies is thought to be that of Wundt (1874), who investigated how accurately a pendulum's future position could be judged. Human prediction was studied extensively during the 1950's and early 60's, mainly due to the research efforts of Gottsdanker (1952, 1955, 1956) and Poulton (1952, 1957, 1964). Poulton (1974) provides a general review of the early tracking literature pointing to the need for advance information or 'preview'. Poulton (1957) has also drawn attention to the difference between 'receptor anticipation' and 'perceptual anticipation'. In the first, input information is received in advance of the operator's motor response, as in tracking with direct preview where the subject is able to look ahead along the course. In perceptual anticipation however, direct preview is not possible, but the operator can use his knowledge of the course characteristics to predict future demands. An internal, cognitive model is therein implied. Perceptual anticipation can be in the short-term, involving the extrapolation of on-going events a short time ahead (speed anticipation); or in the longer-term, such as the anticipation of track reversals (course anticipation). Perceptual anticipation is particularly important in the control of systems with lags and complex dynamics, such as vehicles and process plant. Because the results of operator actions may only become evident over a period of time, there can be no direct preview. However, the complexities of the systems and the interaction with his own control actions frequently make the act of mental prediction very difficult. Poulton (1964) in a study of simple or complex sine wave tracking found that post-view (historical) information helped only when tracking the more complex input without any form of preview, whereas preview (future) information reduced tracking error in all conditions but especially where the amount of track ahead visible was sufficient for the subject to see up to the next track reversal. Similar effects of advanced and delayed information were reported by Gifford and Lyman (1967). Gottsdanker, Frick and Lockard (1961) further note that operators have difficulty in detecting gradual changes of velocity. They suggest that accelerated motion is perceived by comparing early and late velocities rather than by direct sensing, which would account for man's inability to judge motion near to its time of onset.

Davis and Behan (1960) in a study of visual prediction of a radar blip position, and Foot (1969) in a task requiring subjects to predict the point of coincidence of two pointers moving at different rates, both found estimation error to increase over the dead-reckoning period, whereas viewing period had only a slight effect. The latter finding was confirmed by Wiener (1962). An analogy can be drawn with throwing a ball at a target. As the time during which the ball is in the air increases, the likelihood of missing the target also goes up. A common strategy is therefore to throw the ball hard at the target so reducing the required prediction period. It seems that accuracy of prediction is likely to be highly task dependent. Compare, for example, throwing a ball with predicting the effect today's technology will have in the next century.

When the target is obscured as part of a tracking task, the subject's response deteriorates during the obscured interval but rapidly recovers on reappearance of the target (Hammerton and Tickner, 1970). McLeod (1972) notes that during the obscured period subjects continue to pursue the target at its mean velocity when visible, a strategy obtained by sampling from previous behaviour and again implying some form of internal representation of the track characteristics. Flowers (1978), however, found that subjective response during the obscured period was more varied and depended on individual subjects, the point on the track at which vision was lost, as well as the characteristics of the track itself. The path predicted by subjects was often incorrect. When sight of the target is lost momentarily, as during blinking, no deterioration in performance takes place if blinking is unintentional unless the tracked signal is very complicated. This is probably because a person blinks mostly when he does not need to look. However, if the man is forced to blink at a set rate then performance does deteriorate, as instructed blinking is not only an additional task to perform but the man may also be obliged to blink at a critical point in the track (Poulton and Gregory, 1952).

Kahneman and Tversky (1973) present one of the purer studies on the psychology of category prediction and numerical extrapolation. They conclude that people ignore all statistical considerations when predicting the outcome of a particular situation, and instead choose the outcomes which to them appear most representative of the evidence. Hence in contrast to what one might be led to believe from statistical decision theory, intuitive predictions are insensitive both to the reliability of evidence and to the prior probability of the outcome. People will wrongly predict rare events or biassed, extreme values if these seem to be representative of the situation, and furthermore will have an unjustifiably high level of confidence in their judgements. The implications for process control are far-reaching. Spencer (1961) points out that in order to build up a mental picture of the current plant state, a process operator must estimate average values from the continuously varying plant readings. Spencer found large individual differences between operators when a 'rogue' value was introduced, though usually the rogue was underweighted. Although operators were on the whole quite good at estimating average values, errors of judgement increased with the amount of material on which the estimate was based and with its scatter.

3.1.2 Comparison of human and mathematical prediction

Several workers have followed Sheridan's (1966) philosophy of comparing operator performance to that of a reference mathematical model. (See also section 5.1, this chapter.) Rouse (1973 a,b) devised a serial cognitive prediction task which required subjects to estimate the relative position of the next point in a discrete time series. Supervisors of computerised control systems are likely to be concerned with such sampled data. The task was slow enough to eliminate the effects of reaction time and neuromuscular lags, so that any suboptimality was due solely to cognitive limitations. Human performance on the task in terms of prediction error was approximated by a linear regression model with limited memory (which effectively 'forgot' old data) and observation noise (equivalent to an operator's short-term memory limitations and his inability to perfectly estimate the magnitude of a stimulus). Models using linear or quadratic extrapolation from past data points yielded far higher prediction errors, whereas models with perfect memory or learning both did much better than the subjects. From this evidence, human prediction strategies are based on more than just simple extrapolation, yet are still suboptimal compared to the best mathematical model. Sheridan and Rouse (1971) had found prediction error to increase progressively if more than one stage ahead was estimated, as the human has difficulty in determining the amount of signal history due to noise level. It seems that physiological factors are the main constraints on a human's ability to predict for short prediction times (in the order of seconds) whereas cognitive factors predominate for long prediction times (in the order of minutes and above). Van Heusden (1977) repeated Rouse's experimental task for simpler mathematical models and confirmed that human prediction of a sampled first-order system, disturbed by pseudo-predictable random noise, is suboptimal. A 'black box' auto-regressive model was found to give a significantly better fit to the experimental data than a limited memory model, although the fidelity of the latter increased monotonically as memory length increased. Toutenhoofd (1974) found that high order 'black box' models gave no better results than lower order models. Tainsh (1977) also compared subjects' predictions of the standard deviations of current and future process parameters with estimates obtained from a mathematical model, this time a Kalman filter. No correlation was found, though a relationship did exist between subjective estimates and the standard deviations calculated by linear regression. Tainsh concludes that the estimates of inexperienced operators are more likely to be related to the spread of the displayed points rather than to statistical considerations, a view tending to support that of Kahneman and Tversky discussed earlier. It is interesting to note that the noise level inherent in the track had little effect on subjects' judgements.

3.1.3 Summary

In summary it would appear that although unconscious anticipation through some form of internal model is an important feature of control and decision-making behaviour which humans are quite able to manage in their everyday life, human predictive abilities are not all they might be when using abstract displays. Operators have difficulty in predicting the response of systems involving complex dynamics and time lags, and they cannot accurately perceive first or higher order derivative information or detect gradual changes in velocity. In the absence of updated information the accuracy of prediction gets worse over time. People ignore statistical considerations in their estimates, with the result that their predictions are suboptimal.

3.2 Mathematical prediction

The problem of predicting the mathematical response of time-dependent systems has received considerable attention in statistical theory, control engineering and economic forecasting. Comprehensive reviews of forecasting techniques are given in Gilchrist (1976), or Montgomery and Johnson (1976). Basically, a number of approaches to forecasting can be distinguished: intuitive, causal and extrapolation techniques. Intuitive methods are the classical methods of forecasting, and are based essentially on an individual's feeling for the situation. Causal methods try to forecast effects on the basis of knowledge of their causes: since many causes are economic in nature, causal methods find wide applications in economic theory. Extrapolative methods are based on the extrapolation into the future of features shown by relevant data in the past, usually through the construction of a mathematical or statistical model of the data, smoothed if necessary to eliminate the effect of purely random variations. Thus the components of forecasting by extrapolative techniques are always data collection, data reduction, model identification and parameter estimation, leading to the extrapolation of future values.

3.2.1 Linear prediction theory

The classic early work in this area was carried out by Norbert Wiener (1949) in his 'Extrapolation, Interpolation and Smoothing of Stationary Time Series'. Wiener was motivated by the anti-aircraft gunnery fire control problem which became apparent during World War II, and when his aircraft predictor was incorporated into radar trackers the number of hits per thousand shells went up considerably. Wiener's basic premise was to assume that the statistical properties of the plane's time history would remain stationary over a limited period of time, and could thus be used as a basis for extrapolation. Wiener's tome became known as the 'yellow peril' due to the formidable mathematics which it entailed, and it was left to Bode and Shannon (1950) to provide a translation of Wiener's linear, least-squares predictor into terms that engineers and laymen could more readily interpret. A further refinement to linear smoothing and prediction theory was developed by Kalman (1960) for use in control tasks where the basic information is uncertain. In Kalman filtering a number of measures from the process to be estimated are input to the filter and combined with a statistical estimator (the covariance matrix) calculated on the basis of prior information. The combination of new state information plus old covariance information leads to better estimates of future process variable means and standard deviations. The precision of the estimates increases as further information is input to the model, though there is the disadvantage that initial values are required by the filter before it can begin its iterative procedure, so there are necessarily considerable fluctuations in the initial stages of the filter output.

3.2.2 'Fast-time' approach

Ziebolz and Paynter (1954) first conceived the method of running a simulation model of the controlled process repetitively in 'fast time' to provide phase advance information within the control system. Computers are particularly suited to the execution of such techniques due to their capacity for rapid, accurate computation and logical sequencing. Discrepancies between predicted and desired pre-computed future status could be repetitively computed and fed to an automatic controller which acted to eliminate the discrepancies. Chestnut, Sollecito and Troutman (1961) discuss applications of predictive control systems actuated by estimates of future error signals, in the areas of space navigation/rendezvous missions, chemical process control, and the landing of aircraft along prescribed paths. The authors note that it is not always necessary to construct an exact, fast-time model of the system to be controlled for the purpose of extrapolation, and that a simpler second- or third-order model using equivalent time constants can be used with considerable success.

The ability of predictive control systems to operate with inexact models is particularly important for their use in dynamic chemical control processes whose characteristic changes cannot always be easily measured. Further applications of predictive control systems to chemical process control have been cited by Adams and Schooley (1968), Burghart and Lefkowitz (1969), and Klubnikin (1966). Predictive techniques have now become an accepted part of modern control theory and practice, though it should be noted that in general the industrial application of advanced control methods is much less than the importance accorded to them in the literature would suggest (Edwards and Lees, 1973).

3.2.3 Choice of predictive control model

The choice of an appropriate predictive control model is largely an engineering problem, and in practice depends on the computational facilities available as well as the requirements of the system. Of course no model can hope to predict catastrophic system failures, though some indication that all is not well might be revealed in the lead-up stages. At one extreme of sophistication a simple least-squares model may be used to obtain a 'best fit' line through a set of recent data points. The Taylor series extrapolation technique is an extension of this model, comprising a polynomial whose number of derivative terms may be varied to give a response which varies from a simple straight line to an n-th order curve. At the opposite extreme of complexity some extremely sophisticated prediction models are available, including multivariate and adaptive techniques. In the latter, statistical parameters of the smoothing model are updated over time, as in Trigg and Leach's (1967) 'adaptive tracking filter'. The design of adaptive servosystems is discussed by Tomizuka (1975); however a detailed review of such techniques is beyond the scope of this thesis.

3.2.4 Why include the man?

It must be stressed that in all applications so far discussed mathematical prediction has been incorporated as an integral part of the control system itself. If such techniques are feasible, one may reasonably ask, why then include the human operator in the control system at all? The answer, that of a human's lower capital outlay, coupled with his flexibility and safety, was mentioned in an earlier section. But in addition many of the mathematical optimisation techniques are too unwieldy and exacting for real-world applications, since they assume at the very least a problem environment which can be defined in some way. Most real-world problems are, however, extremely ill-defined. The utility of predictive control is therefore limited to those processes in which trajectories can be computed for all possible contingencies. Where it may be imperative to deviate from the prescribed program some form of manual over-ride is necessary. Recent excursions by control engineers into the area of 'fuzzy' control, automating the type of conditional logic verbalised by process operators, suggest that existing fully automated control techniques still leave much to be desired (King and Mamdani, 1975; Gaines, 1976; Tong, 1977). A compromise solution might be to retain the operator in the control loop but display to him directly the prediction model's forecast of future system behaviour, so compensating for human weaknesses in predictive ability but retaining his many advantages. This was the approach adopted by Kelley (1958) in his development of the 'predictive display' concept, and forms the underlying concept for the approach adopted in subsequent chapters.

4. METHODOLOGY

This section considers two issues of concern when researching into aids for the control of industrial processes: the problem of laboratory versus real-world studies, and the methods used to collect performance data.

4.1 Laboratory, simulation and real-world studies

The question of whether or not to simulate is an old and vexed issue. (For a cross-section of the many different viewpoints the term 'simulation' evokes, see the review of a recent symposium on the subject by Whitfield and Goillau, 1978.) Certainly the process control literature seems quite sharply divided between rigorous laboratory experiments and cruder, real-world 'case studies' (Drury and Baum, 1976).

4.1.1 Advantages and Disadvantages

At a theoretical level, laboratory studies have the distinct advantage that the experimenter may isolate the variables he is interested in and determine their precise effects and interactions, having first excluded by careful design those factors which do not interest him. At a practical level, laboratory studies are considerably simpler and cheaper to carry out than their real-world counterparts. Simulation as a technique can be more realistic than other types of laboratory study but does not involve the difficulties of experimenting on an operational process. Bainbridge (1975a) has commented that all the systematic studies of process control skill have been performed in the laboratory.

There are dangers, however, in the ivory tower approach. Chapanis (1967) has heavily criticised the bulk of laboratory techniques for their lack of relevance to the real world. By their very nature laboratory experiments are at best only rough and approximate models of any real-life situation. Bainbridge (1975a) again notes that laboratory processes are usually much simpler and easier to control than in a real context. An experimenter exercises bias both in the independent variables he selects to test, and in the dependent variables he chooses to measure with. As a result, hidden or unsuspected interactions in real-life can easily nullify, or even reverse, conclusions arrived at in the laboratory. In addition, variables often change when brought into the laboratory. The effect of controlling extraneous or irrelevant variables is to increase the precision of an experiment, but at the risk of discovering effects so small that they are of no practical importance. Coupled with the unrealistic methods of presenting variables in many laboratory simulations, Chapanis feels that one should only generalise with extreme caution from the results of laboratory research to the solution of practical problems. Poulton (1972) echoes Chapanis' disquiet, adding that it is never possible to simulate the general activity, stresses and bustle of the shop floor in the laboratory - McEwing's (1977) 'organisational' as opposed to 'technical' realism.

4.1.2 The need for field validation

For the results of ergonomic research to be of value they must clearly be validated through field experiments (Hartnett and Murrell, 1973). However, field research is not without its problems, necessitating somewhat different techniques, procedures and research strategies (Johnson and Baker, 1974). Few would deny that carrying out field studies is considerably more difficult than laboratory experimentation. Given that most ergonomic field research involves an industrial setting, one must contend with (a) management who may agree to provide research facilities so long as their production schedules are not disrupted, (b) a workforce who may co-operate providing their earnings or job status are not adversely affected, and (c) numerous quirks, such as outmoded industrial practices and shift systems, which are part-and-parcel of the factory environment.

An ideal situation would perhaps include the best of both worlds: a representative laboratory simulation study to iron out potential faults in the main system and to optimise the operator-instrument interface, followed by field validation trials. Given that the latter is not always feasible, an iterative approach between field and laboratory may be called for, with an initial task analysis determining which factors should be included in the laboratory simulation. At the very least, representative test data from an operational environment should be used in any laboratory simulation.

4.1.3 'Part-simulation'

An interesting 'part-simulation' technique has recently been developed at Warren Spring Laboratory, whose facilities were employed for the experiment described in Chapter 7. Safety and economic considerations precluded the testing of novel operator aids on a conventional chemical plant. A computer was therefore used to model the equations of the chemical reaction, to accumulate readings from and transmit data to the plant instrumentation, and to log the state of the system (King and Ray, 1972; Cininas, 1975a; King, 1975). The computer was, however, interfaced to actual operating plant and instrumentation, and process operatives were trained to control the process from its commissioning. This is of particular importance, as the skills of controlling complex plant may only develop over many years. Whilst aware that a computer was part of the process, a reasonable assumption in view of the acknowledged trend towards computer control systems, the operators treated the plant as though it were a conventional, fully productive process. By this deception, valid data on process operator's behaviour could be obtained without the risk factor and high cost associated with conventional plant.

A principal disadvantage of the 'part-simulation' technique is the high cost in time and money of setting up and conducting authentic simulation trials, though this may be justified if the confidence one may place in any conclusions reached is enhanced.

4.2 Objective and subjective performance measures

A second thorny problem concerns the choice of objective or subjective measures of performance. Techniques available to the experimenter include individual or system performance scores on the one hand, and questionnaire or subjective report (verbal protocol) analyses on the other. Their respective merits will be considered more closely.

4.2.1 Objective measures

Objective performance measures have in the past been the conventional, even automatic, choice. This may be traced back to the early tracking literature (Poulton, 1974), where single or multiple axis amplitude error scores, frequency counts of control effort, special-purpose engineering measurements or more recently adaptive measurement techniques have been commonly employed (Kelley, 1969). In addition, the option of individual-centred measures (error scores, control effort) or more global system-centred measures (amount produced, profit) must be specified. Whilst objective performance measures are convenient to implement and give a good general picture of the phenomena under study, they are open to bias (Poulton, 1969, 1973) and often fail to give much indication of an operator's underlying strategies or thought processes. Observing an operator's control room behaviour and carrying out an activity analysis to some degree overcomes this limitation, but by no means gives a complete picture. Measuring eye movements does not reveal whether an operator perceives what is in his gaze, and blanking off instruments cannot let us know which items of information have been lost. The latter approach also assumes that an operator responds to a single instrument reading, rather than to the pattern produced by a block of instruments (Shepherd, 1977).

4.2.2 <u>Subjective measures</u>

Questionnaire and subjective report (verbal protocol) techniques have found favour in recent years, as psychological fashion has again made introspective methods acceptable. Questionnaire design for attitude measurement surveys is an established field (Oppenheim, 1966), and there is at least one example in the process control sector of the use of questionnaires to determine an operator's manual control strategy (King and Cininas, 1976). These workers found that off-the-job questionnaires yielded considerable general information about control strategy, though more detailed questions on aspects of control skill gave inconsistent information. It should be remembered that questionnaire methods also suffer from inherent limitations: the operator may be unable to describe co-ordinated, sequential control actions or automatic, unconscious skills when answering questions, his answers may not reflect changes in strategy when under pressure, and he may tend to describe the official rather than the actual procedure for control. In short, the framework of the questionnaire imposes restrictions on subjects' responses.

The remaining method of inferring operators' cognitive processes is from their subjective reports, the so-called 'verbal protocol' technique, stemming from the classic work of Newell and Simon (1972) in the area of human problem solving. Subjects are requested to think aloud whilst performing the task as they would do normally. This approach has been successfully employed by Rasmussen and Jensen (1974), Bainbridge (1968, 1972, 1974) and Umbers (1976). Less successful applications include those of Crossman and Cooke (1962) who found little conscious decision-making associated with control changes in their water bath task, and Brigham and Laios (1975) who found that their subjects were unable to give detailed descriptions of how they used information. The last two studies highlight some of the problems inherent in the protocol approach. A one-to-one mapping must first be assumed between subjects' verbal reports and their internal thought processes. This may not be the case, or the relationship may change with time. Bainbridge (1972) found a close agreement between a subject's protocol data and the corresponding computer log of his actions. Cininas (1976), however, found that an operator's reported chart recorder observations did not match the computer log (obtained by fitting flaps over the front of the recorders) in the initial stages of controlling the reaction, when the operator was overloaded and could not report all his actions. Further problems are that the act of verbalising itself tends to interfere with thought processes and depress cognitive performance (Henderson, 1975). It also constrains the subject to reporting his behaviour serially, so that parallel processing can only be inferred. Often in skilled behaviour several things are done together in a co-ordinated way which cannot be represented adequately by a serial description. Operators are unable to report decisions which they perform or things that they notice automatically at an unconscious level. They may only report easier tasks or give the official version of their operating procedure. A common finding is that subjects make little direct comment about their strategy, rather they simply report actions. Indirect methods must therefore be used to find what determines the sequence of activity.

4.2.3 A compromise solution

From the above it is clear that protocol techniques are still at the experimental stage of development, and much work still needs to be done before they can be regarded as a standard technique. Given that both objective and subjective methods have their relative advantages, it would seem sensible in practice to employ a combination of protocol and questionnaire information to supplement more conventional performance measures and activity analyses. Since its aims are somewhat different, the protocol analyses need not be as extensive as those of workers who have used protocols as their only source of data. In this way, a complete picture of the operator's control actions and strategy may be constructed.

5. APPROACHES TO MODELLING THE HUMAN OPERATOR

There have been three broad approaches to modelling the human operator's behaviour; two essentially mathematical approaches namely linear control theory and statistical decision theory which both aim to model and so predict an operator's control and decision-making behaviour, and the 'internal model' concept. The first two will be considered briefly and the latter in more detail.

5.1 Mathematical approaches

5.1.1 Linear control theory

With the development of a mathematical theory of linear servo-mechanisms (see for example Atkinson, 1968) came an interest in modelling the human controller. Early attempts by engineers to model an operator's control processes visualised him as an element in the control loop whose sole function was to close the loop as best he could. The servo-mechanism analogy was supported by studies of psychomotor skills and the subsequent discovery of a 'psychological refractory period' (Telford, 1931), suggesting that the human operator behaved broadly as a ballistic, intermittent correction servo (Craik, 1947). It was also paralleled by the development of Cybernetics - the science of control and communication in the animal and the machine (Wiener, 1948).

During the 1960's and early 70's much effort was spent on attempts to model the human operator. Many of the findings were published in the American journal IEEE Transactions on Human Factors in Electronics (later Man-Machine Systems), and in the Annual NASA University Conferences on Manual Control. The work of this era has been summarised by McRuer and Jex (1967), Kelley (1968), Gaines (1969), Frost (1972), Sheridan and Ferrell (1974), and Stassen (1975), and there is currently a renewed interest in the application of control theory to human factors problems (Rouse, 1977, Human Factors Special Issue). Some of the more important models will be covered briefly.

Based on linear systems theory came the transfer or describing function model, e.g. McRuer's 'crossover' model (McRuer and Jex, 1967), in which the output of the human operator could be divided into a linear system equivalent response (the describing function) and a remnant equal to the difference between the actual system output and the linear system equivalent response. This model is based on closed-loop stability considerations and though it has been applied to describe the human operator's behaviour when controlling fast response systems it is not really applicable to systems having a slow response. Also from linear systems theory came the optimal control model (Kleinman, Baron and Levison, 1971; Phatak, 1976). This model stated that the operator behaved optimally within his inherent limitations, and could for modelling purposes be represented by a Kalman filter, a predictor to compensate for the human data processing and response lags, an optimal controller, and model terms representing observation and motor noise. The optimal control model has been used to describe human control of fast response systems, and again has not been applied to slow response systems.

Besides these two important models, many non-linear models have been developed, being for the most part extensions of linear transfer function models with non-linear elements chosen intuitively to meet the non-linear outputs of specific experimental situations. Their applicability is therefore limited. Several workers (Elkind and Miller, 1967; Young, 1969) have proposed adaptive models to describe operator behaviour in systems with changing dynamics. The human behaves as an adaptive controller, modifying his transfer function parameters in order to achieve a stable and well-damped closed-loop performance. After the adaptation phase, his behaviour can often be approximated by a describing function model. A final approach, that of the finite-state machine (FSM) model has been suggested by Angel and Bekey (1968). In this model a response is obtained not by plugging a value into an equation but by entering at the appropriate point in a matrix or look-up table. Cooke (1965) and Beishon (1966) have suggested this approach for control of slow response systems, but call it 'system state/action state' control.

It can be seen that with fresh developments in modern control theory, engineering models of the human controller have become increasingly sophisticated and capable of describing human behaviour in a variety of tasks. This trend is likely to continue in the foreseeable future. However, Kelley (1968) has pointed out a fundamental objection. It is inappropriate to describe human operators in the general terms of control theory because there are fundamental differences in principle between the way human and automatic control systems operate. The differences originate in human consciousness, in the deliberate choice of methods (i.e. planning) to achieve goals which cannot be represented by transfer function models alone. Not only does the transfer function approach fail to take any account of individual differences and inconsistencies, to represent the operator's data reduction processes, or to incorporate any explicit representation of the task in question; most importantly it implies that an operator's response is a direct function of his immediate input - human memory, planning and prediction processes are ignored. Beishon (1966) has also criticised control theory 'black box' approaches, since they pay little attention to the processes inside the box. We have already seen that human control is essentially organised towards the future. Sheridan et al. (1964) and Sheridan (1966) have also commented on the inadequacy of quasi-linear describing function models to represent tasks where the operator can look ahead. Although preview models of the human operator have been developed (Sheridan, op cit; Rouse, 1973 a,b) the remaining fundamental objections to engineering models still apply. In short, people do not function in this way.

5.1.2 Statistical decision theory

An alternative approach to modelling the human operator has come from the area of statistical decision theory. A general introduction to decision theory and human behaviour can be found in Lee (1971), or Kaplan and Schwartz (1975). The literature in this area tends to be both vast and loosely structured - useful overviews have been published periodically in the Annual Review of Psychology series (Slovic et al., 1977; Rapoport and Wallsten, 1972; Becker and

McClintock, 1967; and Edwards, 1961). Most studies of human decision-making have been carried out in specific settings, so it is difficult to make generalisations. However, the main approaches will be outlined briefly.

The early distinction made between normative (the choices that a rational man <u>should</u> make according to probability theory) and descriptive or intuitive (the choices real people <u>actually</u> make) decision theories has now been largely abandoned, since in practice the distinction often became blurred. Laios (1975) notes that the advocates of decision theory now distinguish between static and dynamic decision theory. Static decision-making (Luce and Suppes, 1965) is concerned with a fixed problem environment or set of environmental states, a fixed set of alternative possible decisions, and a pre-determined payoff matrix. Once a choice has been made, that is the end of the decision-making. Static models have found favour because they and their associated methodology are simpler than dynamic decision-making models, which represent the characteristic fluid aspects of real world decisions. Dynamic models (Lee, 1971) involve a series of individual decisions which may be independent or dependent on each other. Sequential and multistage decision tasks can be further distinguished, the difference being that in the latter early decisions affect later ones, whereas in the former they do not. Real-life decision-making typically involves a series of inter-related decisions, each dependent on previous decisions. Often the consequences of an early decision reach far into the future. Much effort has been spent to investigate multistage decision-making problems within an operational research framework, e.g. dynamic programming (Bellman, 1957), but little effort has been spent by behavioural scientists due to mathematical and experimental difficulties of the approach. Laios (1975) has presented an extensive review of the whole area, and concludes that problems of human multistage decision-making in the real world stem from an incomplete knowledge of the objectives, multiple criteria, environmental states, consequences of decisions etc., together with problems inherent in the human decision-maker's inability to perceive and process the relevant information.

Two schools of thought have developed to predict a decision maker's policy: linear (regression) models and Bayesian models. Linear models have been used by a large number of workers to represent human decision processes, and can be formulated for static or dynamic decision-making tasks. Basically, the model assumes that a decision can be adequately predicted by a linear combination of weighted information inputs, plus a random error term and a scaling factor (Anderson, 1974). In fact, the evidence (Goldberg, 1968; Slovic and Lichtenstein, 1971) suggests that the linear model has a remarkably good predictive ability. However, it must be stressed that although linear models can accurately describe the outcome from a practical situation, they do not necessarily explain the human's decision processes. Many decision makers maintain that their decision-making activity is non-linear, and it may well be that linearity is imputed by an insensitivity of the analysis employed to non-linear effects (Green, 1968). Yntema and Torgenson (1961) suggest that the empirical processes used by decision makers in real life may be surprisingly simple, but that man's decision-making flexibility lies in his ability to simplify complex problems in ways that enable him to cope. According to Vaughan and Mavor (1972) this leads to good solutions, but not optimal ones.

Bayesian models have been applied largely in the context of dynamic decision-making, along with information purchase (also termed optional stopping) and operational research, dynamic programming models (Edwards, 1972). Bayesian information processing, derived from the work of the Rev. Thomas Bayes during the 18th Century, provides a mathematically amenable way of combining prior probabilities with the probability of an event occurring after a set number of observations, to yield posterior probabilities. The Bayesian approach is particularly useful for judgements made under uncertainty (Beach, 1975). Bayesian modelling is often identified with diagnostic tasks, though again the evidence suggests that the human decision maker may not behave in this way (Green, 1968). Most of the psychological experiments on optional stopping have been developed within the Bayesian framework, and the models have not been adequate to explain the observed results (Rapoport and Wallsten, 1972).

In summary, even though the weight of opinion is that the best descriptive model of the human decision process is a linear combination of the elements which influence the decision maker, this does not mean that the human functions as a linear decision system. He may use non-linear relations or look-up tables to reduce the complexity of the task to manageable proportions. As with linear control models, the evidence (e.g. Kahneman and Tversky, 1973) is that people just do not behave as statistical decision makers. In uncertain environments humans are basically conservative, that is they do not revise their opinions sufficiently in the light of fresh information. Edwards (1968) lists the main reasons for conservatism as misperception, misaggregation, and artefacts. Humans also do not extract all the information available in samples of data, and tend to place too much weight on early information (Dale, 1968). Peterson and Beach (1967) confirm that discrepancies exist between human inferences and those of an ideal 'statistical man'. A final criticism of the majority of decision theory literature is that it is far too abstract and theoretical to be of much practical use. Notable exceptions include the PIP (Probabilistic Information Processing) system of Edwards et al. (1968) and Laios' (1975) Predictive Computer Display (PCD), both explicit attempts at decision aiding.

5.2 Internal model concept

All the models discussed previously share one common aspect: they all assume that in order to control or decide effectively the human operator needs some knowledge of the system to be controlled and the properties of the system inputs, including knowledge of possible system disturbances and their consequences, together with an idea of the task objectives to be achieved. In other words, he needs an internal model of the process. Veldhuyzen and Stassen (1977) note that the existence of an internal model is implicitly true for quasi-linear transfer function models - for instance, McRuer's crossover model indicates to what extent the operator is able to adapt his control strategy to the dynamics of the controlled element. The optimal control model also shows very clearly the use of the internal model concept - the construction of the Kalman filter, predictor, and optimal controller require that the system dynamics should be known, as should the statistical properties of the disturbances. In addition, the various non-linear and decision theory models often imply that the human possesses a knowledge of the process he interacts with. The concept of an operator's 'mental model' or 'internal model' has been a recurring theme in previous sections of this chapter. It would now seem to be a fundamental tenet of control and decision-making behaviour. The development of the internal model concept will be considered further.

5.2.1 Early considerations

Craik (1943) was amongst the first to propose that mental processes represent within the brain the nature of the outside world. The brain, in other words, acts as a model in which neural processes symbolise the workings of the external world and thus allow us to predict the outcome of events and forecast the consequences of our own actions (Oatley, 1972). When one says: "I must leave now if I am to catch the train", the form of words presupposes mental structures (with neural processes underlying them) which represent time, the speed at which one can travel towards the station, its direction, distance and so forth. These mental structures mirror, in a symbolic form, objects and their inter-relationships. If mental processes did not accurately represent important features of the real world, the commuter would never catch his train; nor indeed would trains ever have been built. Only occasionally would random behaviour ever bring about a favourable outcome through pure chance. Since the brain model represents the kind of events that can occur in the outside world, "What would happen if?" questions can be tried out without the usually useless and occasionally dire consequences of actually doing it. Oatley (op. cit.) poses the rhetorical question "Why is all this talk of representation necessary? Is there not a perfectly good external world already there towards which we can direct our behaviour, without making models of it?" The answer is that without a means for creating a model world with similar properties to those of the real one, but nevertheless independent of it, such human attributes as thought, language, perception and purposive action would be impossible. The latter is particularly important for present considerations; without an internal model we could only react to stimuli in the immediate range of our senses, never predict the outcome of any action, never behave purposefully since that implies working towards a state of the world that does not yet exist.

Tolman (1948) indirectly reiterated Craik's view in his 'field' theory. Rats (and people) were held to construct a broad, cognitive map of the outside world, which they could then use to negotiate their way through their environment. The mental processes involved in constructing such a map were presumed to be much more sophisticated than any simple stimulus-response relationship. These intermediate processes between sensory input and behaviour have been variously referred to as 'thought' (Craik, 1943; Hebb, 1949), 'schema' (Bartlett, 1932), 'hypothesis' (Gregory, 1970), 'readiness' (Bruner, 1957) and 'internal model' (Kelley, 1968).

Sheridan (1966) and Smallwood (1967) have presented control models which include, in equation form, a model of the controlled element which can be used in fast time to make predictions of future behaviour on which control can be based. Prior to Smallwood's application, the internal model notion had also been implied by studies of the human operator's monitoring behaviour, especially in the control of slow response systems (Senders, 1964; Cooke, 1965).

In such systems the quantities to be monitored change so slowly that continuous viewing is neither desirable nor practicable - an internal model of process behaviour is thus necessary to determine an efficient sampling policy. Fogel et al. (1966) have discussed finite-state machine controllers which include internal models of the external environment. And Bainbridge's (1967) predicting controller model contained templates for system behaviour which were used for prediction.

5.2.2 Kelley's Model

Perhaps the most consistent and influential proponent of the internal model hypothesis has been Kelley (1968). He sees an operator's internal model as a fundamental component of goal conception and selection (the first two stages in Kelley's general schema of control activity, described in an earlier section), and notes that the process by which man conceives of and selects amongst possible future states (or goals) is the most important yet least understood part of the control process. The following quote is from Kelley (op cit): "Man receives information through his senses and applies information stored in memory to create internally, from the little understood materials of consciousness, a dynamic model of the world about him. This model not only represents the spatial structure of the environment, but also incorporates its rules of operation, e.g. temporal order, cause and effect relations. The model represents the individual's perception and understanding of his environment. The nature of the modelling process is such that it is not limited to past and present but can be used to create representations of possible (and impossible) future states as well".

Kelley believes that the mental modelling process operates on a fast-time rather than a real-time basis, in that events may be thought about much faster than they can occur. Human consciousness schematises and compresses events, with significant points and end results included, but much of the remainder omitted. For this reason, the human operator in a control system may consider several possible courses of action in less time than it takes to carry out one of them. Kelley further distinguishes between a full internal model incorporating all aspects of a situation, and simpler derivative models. The latter are built around the display itself, permitting a rate or acceleration to be perceived and represented conceptually as a position, so effectively reducing the control order by a factor of one or two. This distinction seems to match the difference between 'course' and 'speed' anticipation mentioned earlier. Since many physical laws of movement can be expressed in the form of differential equations, the derivative model may behave in a way that is simpler and easier to understand than the process itself. As the model usually bears a clear and straightforward relationship to the controlled variable, it can be employed directly for control. Such derivative models are particularly appropriate for vehicle and other high-order manual control systems.

A man operating a control system for the first time has some understanding of the system, and some expectation (however crude or misplaced) of how it will behave in response to his control actions. In Kelley's terms, he starts out with some kind of internal model of the system. His initial predictions are often in error and force him to change his model, so control based entirely on his internal model would be unwise at this early stage. Hence the naive operator usually behaves as a feedback controller. As experience is gained, however, the operator's internal model is adjusted repeatedly to further reduce errors between predicted and actual output values. A point is reached when the operator's performance no longer improves with practice, errors in prediction have levelled out at an acceptably low level, and his internal model has stabilised. The process of building up an accurate internal model is the primary ingredient in training for skilled performance. An accurate model leads to accurate predictions, which in turn are the basis for skilled control. The same process by means of which the model is developed and refined is employed by the operator to make adaptive changes. The operator's model is changed so that predictions based on the model and control activities based on the prediction reflect the adaptive changes.

Gregory (1970, 1973) endorses Kelley's point of view when he states that perception is a process of selecting 'internal models' in terms of which incoming data are used to shape behaviour. Only through an internal model based on previous experience can so little stimulus information control so much behaviour. Control is rarely direct, except in the special case of reflexes, but is via internal neural models of reality. Gregory cites the occurrence of perceptual illusions as a case where an inappropriate internal model has been chosen to explain the input data.

5.2.3 Verbal protocol approaches

Though the internal model concept has been widely quoted, its nature remains somewhat elusive. Stříženec (1976) has noted that until recently no precise definition has been forthcoming concerning problems of inner, mental models. Verbal protocol techniques have made some headway in their study, however. Cooke (1965) in his water bath task, and Beishon (1969) in his study of cake ovenmen's behaviour, both rejected traditional control and decision theory models and adopted information processing models which incorporated an internal model to explain subjects' behaviour. Both authors based their models largely on protocol data. Protocols from Cooke's water bath task provided several instances of prediction of temperature or control changes and their effects, which could only be satisfactorily explained by postulating the existence of an internal model. Beishon's protocols suggested that his ovenmen had internalised look-up tables, e.g. of baking times and oven temperatures for different types of cake. In addition, Beishon proposed an 'advanced planning' or executive routine which functioned to organise lower routines, handle interruptions, anticipate future events and activities, and maintain a current list of what was to be done next. An executive routine to organise the overall sequence of activity is a common element of program models which aim to produce the same sequence of complex cognitive activities as revealed by human protocols (Reitman, 1965; Baker, 1967).

5.2.4 Bainbridge's Model

The protocol/mental model approach has been exemplified by the work of Bainbridge et al. (1968), Bainbridge (1972, 1974, 1975a). She investigated subjects' behaviour in a simulated power demand task, again using verbal protocols as her main source of data in an attempt to shed some light on subjects' underlying decision processes. An action/information tree model was first tried, the protocol analysis showing that subjects did not follow a continuous path down the tree structure. A program model approach was then attempted. However, in accordance with other workers it was found that an executive routine was required to select specific routines. Bainbridge (1975b) found that the inherent flexibility of her protocol data could be successfully modelled by making explicit the working storage implied in the routines. In her so-called 'head box' approach, the data item required by the operator is represented by the item's name, its present value location (or 'box'), and an address of the routine or routines for obtaining that value. The philosophy of this approach will be familiar to computer engineers.

Bainbridge stresses that by separating routine from purpose, a routine can be accessed from different head boxes, so that any given routine may be used for several different purposes. In addition there might be several different ways of obtaining a data value, e.g. by judgement or calculation, the latter by mental arithmetic or slide rule, etc. Two types of working storage were specified: temporary storage for values generated within a routine, and longer-term storage for more general information. Bainbridge then makes use of standard flow diagram methodology, with conditional statements allowing jumps to different routines in different circumstances, to explain the sequencing of routines for situations where the control error is acceptable, unacceptable, and so on. Her sequencing diagrams are similar to Beishon's (op. cit.) advanced planning or executive routine. Bainbridge notes, however, that mechanisms for dealing with external interruptions, together with a means for deciding between alternative routines accessible from within a single 'head box', must be considered in addition to the flow diagram sequencing mechanism. The latter is especially difficult to account for, since as was noted earlier controllers rarely give the reason for their choice of action. Controllers' behaviour suggests that the routine used at a particular time is chosen according to the time and working storage available, the difficulty of the operations involved, and the accuracy required.

The sum total of the flow diagrams (describing what to do when), plus knowledge of static (long-term) and dynamic (temporary) process characteristics, were held to constitute the operator's 'internal model' of his process. Bainbridge distinguishes between an operator's 'mental model' of the process and his 'mental picture', the latter being the contents of his working memory comprising the actual data values stored at a particular time. These values provide the context in which decisions are made. They determine the sequence of behaviour, and are themselves the items found by the main routines.

5.2.5 Rasmussen's Model

Remaining in the area of industrial process control, Rasmussen (1974, 1976) and Paternotte (1976) have both shown an interest in process operator's mental models. Rasmussen in particular has used protocols to develop the idea of a hybrid hierarchical model of an operator's monitoring and control behaviour, having conscious and subconscious components. The notion of hierarchical control is of course not new, having been previously taken up by Miller, Galanter and Pribram (1960) in their 'Plans and the Structure of Behaviour', and also by Kelley (1968). Broadbent (1977) reviews the question of control hierarchies in detail. Rasmussen notes that the process operator's data processes must be controlled by a representation of some kind of the functional properties of the plant. This representation obviously can be derived from different sources, for example, from previous experience on the plant; from knowledge of its internal construction and functioning; or from prescribed rules and instructions. He distinguishes between different mechanisms for data processing: the operator may respond 'automatically' to a situation, or he may identify a problem and 'think' out its solution. These mechanisms correspond broadly to subconscious and conscious model components respectively. Only the latter component is accessible by means of protocol analysis. During a long period of interaction with a system a trained operator will develop a large repertoire of complex and partly subconscious routines, which are controlled by a conscious (and hence verbalisable) sequence at a high level of abstraction.

According to Rasmussen, the subconscious component of the model comprises a high-capacity, parallel processing system serving functions related to perception, sensory motor responses, etc. It resembles in its operation a goal-oriented, self-organising, associative network operating by dynamic matching of input information patterns to stored patterns. A dynamic model of the external environment is included, constituting the operator's 'process feel'. The internal model also directs and controls attention, and accounts for prediction so that efficient feed-forward control is possible in sequences too rapid to allow for sensory feedback. If a mismatch occurs between the behaviour of the real world and the predictions of the internal model, or if an appropriate model is not available, this is detected and the conscious processor usually alerted. Failing this, an error of judgement or perceptual illusion may result.

The conscious component of the model, however, is an extremely versatile sequential processor, but of limited speed and capacity. It acts as a high-level co-ordinator of the subconscious processes, and functions in unfamiliar situations which demand that unique responses be thought out. The conscious processor can function at various levels of abstraction; it can call up information from the lower perceptual system, and can calculate the consequences of possible actions via different types of representations (mental models) of the physical system considered, e.g. causal models for rational deductions, representation of typical system behaviour as in 'visual thinking', or prescribed algorithms for control rather than a structured model. The model chosen would depend on the task in question. Edwards and Lees (1973) note that the model may be in the form of an equation, a graph, a table, a linguistic expression, and so on.

Perhaps most importantly, the two processing mechanisms co-operate, the subconscious processor by virtue of its large repertoire of automated subroutines relieving the limited capacity of the conscious processor. The latter controls a sequence of such subroutines via an executive or sequencing programme, in which the level of abstraction rises as training increases the efficiency and complexity of the subroutines. Conscious control is, however, significantly influenced by the subconscious processes directing attention, supplying intuitive hypotheses, and so forth. The foregoing is a summary of Rasmussen's hybrid model. The systems engineering implications of the suggested model and its psychological references are discussed in detail elsewhere (Rasmussen, 1974).

Rasmussen's views are somewhat close to those of Bainbridge (op cit). Routines are again selected as a function of their output rather than what they do. Behaviour is organised at several levels of complexity, due to the hierarchical nature of the process operator's total task. However, the organisation is not in the form of a straightforward hierarchy commonly found in problem solving experiments. Rather the efficiency of skilled performance is due to the ability to compose the behaviour needed for a specific task as a flexibly-linked sequence of standard subroutines which are useful in different contexts. Rasmussen likens the data processing steps making up a sequence to a 'ladder of abstraction', with one leg upwards for analysis of a situation, another downwards for planning of the proper action. The number of steps actually taken, however, depends on the skill of the operator. Habits and rules-of-thumb act as short cuts to connect the two legs of the ladder for experienced operators; only a novice would follow the full sequence as set out in the ladder. Actual skilled performance is described in terms of 'shunting leaps' within the basic sequence. The short-cuts are equivalent to a shunting-out of activities, from the higher levels of abstraction which call for complex conscious reasoning, to automatic lower levels, and a considerable increase in data handling capacity results. Rasmussen also notes that the skilled operator need not enter the sequence at its entry point - he can use his 'process feel' to start at a later stage in the sequence. His internal model may even enable him to perceive directly in terms of system state rather than observing separate items of information.

5.2.6 Advantages and disadvantages of internal model approach

To recap, the advantage of control and decision-making through an internal model can be expressed as follows (Gregory, 1970):

1) it makes use of the redundancy of the real world by extracting key features, thus reducing input requirements and relieving higher cognitive activities;
2) it confers predictability, so circumventing the effect of human response lags; in addition the consequences of alternative courses of action can be calculated;
3) it gives continuity in the absence of a continuous input;
4) it allows generalisations to be made to similar but novel situations, or to aspects of the same situation not previously considered.

Weighed against these, two potential disadvantages of the internal model concept must be considered.

1) The operator can be systematically misled by an inappropriate internal model, as in the case of illusions. (Since it would not be possible to store an internal model for every eventuality, flexible models must be assumed which can be 'scaled' to fit reality. In the absence of reliable scaling information, the 'average' past value is often assumed leading to perceptual distortions).
2) Models are essentially conservative, reflecting as they do the past rather than the present. In other words they are resistant to change.

Střížeňec (1976), and Baum and Drury (1976) have both reviewed the status of process operator mental models. Baum and Drury comment that the available industrial and laboratory evidence tends to support a loose, hierarchical goal-directed model as proposed by Bainbridge (1974) and others. Střížeňec concurs, believing that a combination of cybernetic and internal model view points may yield the best solution.

Having reviewed the literature, it seems clear to the present author that in considering process operator models it is not a question of whether the operator has an internal or mental model of his process, but rather what form this model takes and how it is affected by perturbations in the environment. Further work is needed in this area. Given that man does form an internal representation of the outside world, it seems that the value of predictive displays may well be in helping him to develop and maintain an accurate internal model by providing him with immediate feedback of the results of his decisions and control actions, either on-line or in a trial-and-error test facility mode.

CHAPTER 2

PREDICTIVE DISPLAYS

1. INTRODUCTION

The previous chapter has reviewed process operator skills, models of the human operator, and the role played by prediction. It has become clear that the ability to anticipate future events is a main feature of skilled control, but that this ability in humans can be far from perfect, particularly with high-order, lagged systems. Engineers and psychologists have thus been encouraged to develop aids for control and decision-making, and the predictive display concept has been to date one of the most promising ventures in this area. The next section will outline the historical development of predictive displays and describe briefly how they work, the second section will review the more notable applications, and the third section will present the experimental evidence on factors affecting predictive display performance. The final section gives an overview of the introductory chapters.

1.1 Background and early attempts to assist the operator

The predictive display concept was invented by Kelley (1958) in a logical yet innovative development from Ziebolz and Paynter's (1954) theory of two-time scale computing. The latter was an entirely automated theory which did not permit intelligent overriding of the control system by the human operator, whereas the predictive display concept recognises that the man has a valuable role to play in the control loop. In a classic example of 'designing the machine to fit the man', the predictor instrument compensates for man's inherent response lags and lack of predictive capacity by displaying future as well as present system status information to him, whilst at the same time making full use of his outstanding perceptual and intellectual abilities and his flexibility. Thus the predictive display concept uses a form of 'fast-time' model as suggested by Ziebolz and Paynter, but replaces their automatic controller by a human operator.

Kelley (1968) presents the authoritative history of predictive display development. It must be stressed at the outset that the predictive display is quite distinct from other forms of control assistance which became popular at about the same time, namely 'aiding' and 'quickening' (Birmingham and Taylor, 1954). Aiding is a method of compensating for the operator's relative inability to obtain derivative (rate of change) or integral (summation over time) information from a display, by moving any derivative, integral or algebraic summation functions to within the controlled mechanism itself (Murrell, 1976). The operator thus acts as a simple amplifier through the control stick, a single adjustment of which effects a change in the position and rate of movement of the controlled element. The practical utility of aiding is generally limited to pursuit tracking applications, e.g. gunnery, and those systems having negligible inertial lags.

Quickening is a second method of compensating for the operator's inability to extract derivative information. It differs from aiding in that it does not act directly on the controlled system, but rather functions indirectly to provide a simplified display to the operator who then responds through a conventional controller. Hence this general class of assistance is sometimes termed 'display augmentation'. Feedforward loops compute information on output position, rate, acceleration etc. which is then subtracted from the desired input value to form an error signal which the operator must minimise. Quickened displays can be used in the control of high order systems and those which include response lags. Control performance is found to be substantially improved and learning times reduced in such applications as helicopter hovering and chemical process control, both of which typically involve high-order dynamics and response lags. Quickening can never be achieved with perfect success, however, since it would be self-defeating: in theory there would be no need to retain the operator in the system at all if his role could really be reduced to that of a simple amplifier. Further disadvantages of quickened displays are that they are of no help when dealing with procedures which cannot be pre-programmed into the quickening circuits, and there is a real danger that the operator will mistake quickened output for status information. Since it is not usually possible to pre-program every contingency, the application of quickened displays is quite limited.

Many observers comment on the apparent similarity between quickened displays and predictor displays, particularly those based on simple prediction models. Both calculate derivative terms in order to display immediate knowledge of results to the operator. Both (it will subsequently be shown) result in considerable improvements in performance and reductions in learning time. Neither have been exploited commercially. A fundamental difference, however, is that whereas quickened or 'command' displays tell the operator what to do through minimising an error signal, the true predictor display tells him what is happening. Command displays do not usually incorporate system status information (or do so clumsily with a separate display), and without such information the operator is entirely dependent on an error signal. In addition he cannot 'see' the trajectory or associated characteristics of the controlled system with a command display. It is now clear from a fair body of research that as a general principle the more closely the information given to the operator relates to that which he actually needs in order to control effectively, the quicker will be the learning process and the lower will be the likelihood of control errors (Murrell, 1976). It transpires that what the operator actually needs, besides a general knowledge of his objectives, is information on the current state of the system and its future behaviour (Singleton, 1972). As Bainbridge (1975a) has pointed out, the ease with which process behaviour can be learnt is particularly affected by the way process information is displayed. The progressive inability of humans to control systems having an increasing number of integrations in the forward path is not a response time limitation, but is a conceptual difficulty arising because the required manual control action must be quite unlike the immediate observed response of the system (Fargel and Ulbrich, 1963). If the operator is to learn the effect of an action or input on the process output then he needs clear information about these effects, and so it would be a mistake to display an error signal alone without the underlying output and target value. This matches Poulton's (1974) recommendation that pursuit displays are in general preferred to compensatory displays.

In fact, experimental comparisons of conventional, quickened and true predictor displays have shown the latter to be reliably superior (Kelley, Mitchell and Strudwick, 1964; McLane and Wolf, 1966), though a predictor display has yet to be compared with a rudimentary velocity vector display (Poulton, 1974). In addition, predictive displays are more flexible than separate quickened and status displays, as all the information the operator requires can be integrated on a single display. However, motion predictors have not been widely applied for two reasons (Smith and Kennedy, 1975). The first is that, like most sound but untried ideas, they are still regarded as a kind of sophisticated toy. A second, more fundamental, reason is that their use assumes that the operator has a meaningful role to play in the control system, and as such directly contradicts Birmingham and Taylor's (1954) design philosophy, a philosophy which has for too long dominated human factors thinking.

1.2 Predictive display operation

In essence, predictive displays make use of a mathematical model of the controlled system which can be run repetitively ahead in 'fast-time' to predict the future response of the system to changes in the controls, the resulting excursion being displayed to the operator. Its ability to indicate what the operator must do to achieve a desired future state is the predictor instrument's main attribute. The mathematical model on which prediction is based need not necessarily be sophisticated: a previous section (Chapter 1, section 3.2) has already outlined the wide range of mathematical prediction techniques available to the engineer. Simple prediction models were found to suffice for some forms of fully automatic control.

Bernotat and Widlok (1966) have produced a useful classification of predictive display types, based on the fidelity of the prediction model employed. The experimental findings relating to each class will be reviewed in a later section. Basically, Class I is the simplest form of prediction model and employs a power series (typically the Taylor series) to extrapolate future values from past data points. The purely mathematical nature of the extrapolation does not account for the unique response characteristics of the controlled system, so predicted values become progressively less accurate as the time over which prediction is required increases. Bernotat (1972) and his West German co-workers have made extensive use of this technique for short-term aircraft stabilisation and guidance problems, often using a single predicted endpoint rather than a two-dimensional curve. Class II prediction extends the simple model of Class I to include the individual controlled system's response characteristics. Kelley's work (1968) has been largely concerned with Class II predictive displays, in applications requiring a more accurate, medium-term prediction, e.g. submarine guidance. The final type of model described by Bernotat and Widlok, Class III prediction, again extends the fidelity of the prediction model by including the effects of such external influences on the environment as can be pre-programmed. Class III predictive displays can therefore be used for those applications requiring a highly accurate, long-term, navigational feature, e.g. spaceship trajectory prediction. For most practical situations Class II prediction will suffice, since rapid updating of the model will minimise the deleterious effects of external disturbances.

It is of interest to consider Kelley's (1968) original description of his Class II device, shown in Figure 1a. Kelley writes: "The heart of the predictor instrument is the fast-time model of the controlled element. This model can be mechanical, electromechanical, or electric, using either analog, digital or hybrid simulation methods. Frequently the fast-time model is a simulation by means of a repetitive electronic analog computer. Sensing instruments in the real system provide signals which are transduced into DC voltages and scaled to equal the voltages representing corresponding quantities in the analog model. In this way, the sensing instruments provide initial conditions for the analog system, conditions that begin each cycle of its operation. If the cyclic resetting device resets 50 times/ second with a 1 ms reset time, and the analog operates on a time scale 500 times that of real time, the analog system will represent the period from present time to 9.5 seconds into the future. The predictor instrument is completed by using the predictive signal from the output of the analog system and a sweep or other timing signal to generate and display the future state of the controlled element, future path, or other form of predictive information. The display symbol may correspond to all or part of the prediction period".

Kelley goes on to discuss the various control actions which the operator must be assumed to make during the prediction period. Typical assumptions might be that the operator returns his control to a null position after an appropriate lag, or he might move his control between either extreme. The simplest assumption and the one made in the predictive display systems of the present thesis, is that the programmer of Figure 1a feeds the output from the operator's control device directly to the fast-time system model.

By way of illustration, Figure 1b gives a simple example of what a predictor display used for guiding a piloted aircraft onto a runway glidepath might look like. In the example shown, the predicted trace indicates that the aircraft will undershoot the runway if the pilot makes no further control adjustments. He will therefore insert one or more control inputs until his aircraft rises onto the glidepath and the

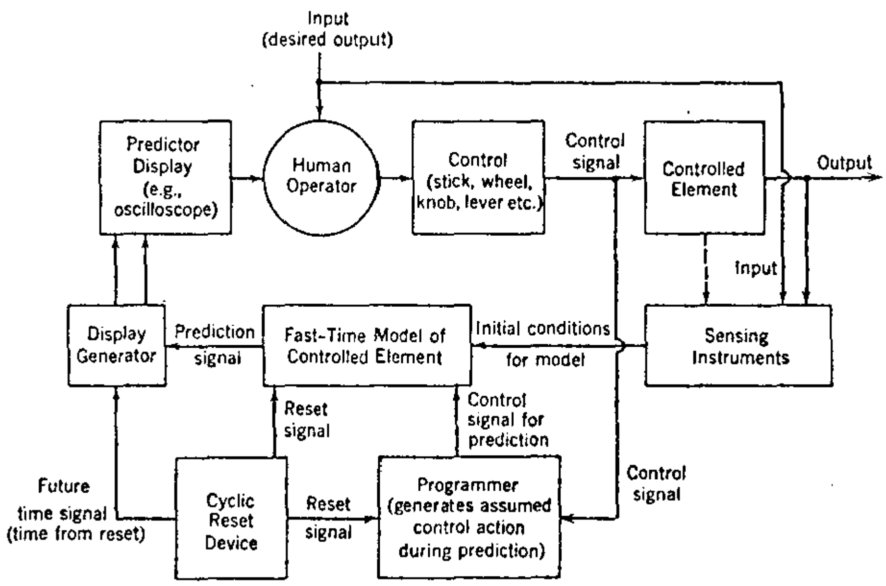


Figure 1(a) Block diagram of a manual control system employing a standard predictor instrument (from Kelley, 1968, page 140).

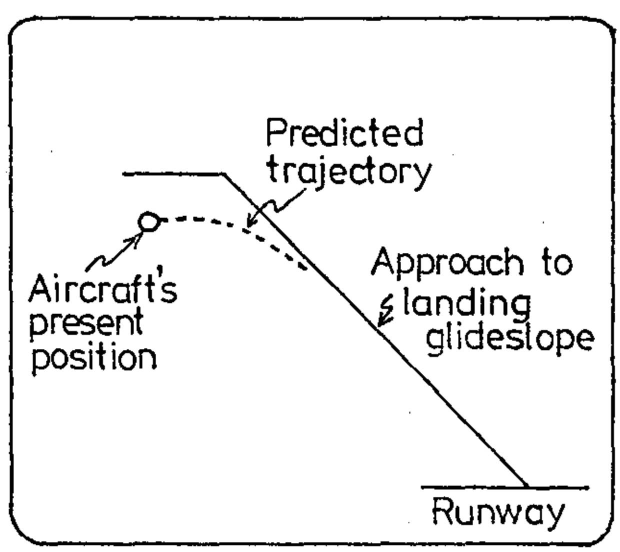


Figure 1(b) Example of a simple two-dimensional predictor display, showing an aircraft approaching the target glidepath to a runway (from Smith and Kennedy, 1975).

predicted trace is linear and superimposed on the glidepath. This form of display is similar to that employed in Chapters 5, 6 and 7 of this thesis.

There is, of course, no reason why the predicted trace should be continuous. Several workers have successfully displayed only the endpoint of or selected points along, the predicted trace (e.g. Bernotat and Widlok, 1966; Wierwille, 1964). With discrete systems such as occur in scheduling applications, the prediction model will still incorporate the characteristics of the controlled process, but a more appropriate display format would show the predicted states of the system at set times into the future, on the basis of actual or hypothetical system inputs. This form of predictive scheduling aid has been used in Chapters 3 and 4 of this thesis. Various examples of predictive display applications in discrete and continuous systems are reviewed in the next section.

A fairly recent development in predictive display technology will not be considered in detail in this thesis, but is mentioned here for the sake of completeness. These are the adaptive techniques ('adaptive displays') devised by Kelley and his associates (Kelley and Prosin, 1972; Prosin and Kelley, 1973). In these applications, previous predictions from the system model are compared with current performance. If a discrepancy appears, this indicates that the real-time system is no longer following the same laws as expressed in the model and so adaptation by the operator is necessary. Simulation exercises (Kelley and Prosin, op cit) have shown that the technique helps the operator to update and maintain his internal model of the system in the case of significant system changes. The extent of such changes could be diagnosed more accurately and more quickly, with the adaptive display. There have, however, been no practical applications of the technique to date, and it is not considered further.

2. APPLICATIONS OF PREDICTIVE DISPLAYS

Computer aiding for the human operator usually falls into one of four areas (Whitfield, 1977): perceptual aids to assist with pattern recognition, alarm analysis, and the detection of signals from noise; decision-making and control aids for problem diagnosis and action selection; memory aids to provide information storage and retrieval facilities appropriate to human search; and output aids to extend man's capability for action in examples such as teleoperators. Predictive displays come under the second heading, since they are concerned with employing a computer-based predictive model to help with the choice of an appropriate decision or control action. Reference may be made to Whitfield (1975) for a broader coverage of man-computer symbiosis issues.

In the first chapter, the distinction was made between discrete and continuous processes, though at the time it was noted that the human's processing of any information is via a discrete mechanism, so the discrete-continuous dichotomy may not be that important in human terms. As with processes themselves, a distinction may conveniently be drawn between scheduling aids for discrete systems, and tracking aids for continuous control applications (Bird and Whitfield, 1975). The distinction is upheld because practical considerations usually mean that discrete aids are incompatible with continuous applications, and vice versa. As an example, the prediction span* of scheduling aids is often quite long (hours); whereas the corresponding span of tracking aids is typically fairly limited (in the order of seconds). In subsequent sections, the evidence relating to predictive aids for discrete and continuous applications will be considered separately.

2.1 Discrete applications

Discrete decision aids have been devised for a wide variety of scheduling situations, where a number of items or events must be arranged to match limited resources by forming an array which satisfies a number of performance criteria. Applications can be grouped under the headings of static decision aids (stateboards), production or job-shop scheduling, air-traffic control, and the scheduling of steel plant soaking pits. The following subsections present a brief review of these areas.

2.1.1 Stateboards

The stateboard can be a simple and inexpensive aid to manual scheduling. It gives immediate information regarding the current status of the process, and can also be used for advanced planning. Shackel and Klein (1976) report an early application of the principle from 1968 in scheduling the refuelling operations of aircraft at the Esso London Airport Refuelling Centre. Prior to the introduction of the magnetic stateboard the operators found the task of allocating the available manpower and equipment whilst satisfying the conflicting objectives somewhat difficult. The stateboard served a dual role as a memory and a planning aid, and proved effective in helping the scheduler to make best use of his limited resources. Brigham (1974)

---

* defined as the time period over which predicted plant response is displayed.

and Sarkar (1972) cite further applications of the technique in a container terminal and x-ray department respectively. The 'Planalog' of O'Brien (1969) is an extension of the concept to production scheduling. Oliverson (1971) has noted that no scheduling board, whatever its cost, can be successful unless the input data are accurate and are kept current.

More recently, Gibson and Laios (1978) have developed a 'scheduler's abacus' on which to evaluate three different graphic methods of presenting scheduling information from a job-shop scheduling environment. The abacus (based on a Gantt chart representation) forms the basis of Chapter 4 in the present thesis, and a full description can be found there. Gibson and Laios report that each of three graphic methods proved more effective in helping subjects to produce efficient schedules than a conventional, card-based numerical presentation. In particular one method which used a machines-by-time organisation and identified machines by colour code proved superior to the others tested, as it facilitated solution of the scheduling problem by perceptual rather than computational means. A follow-up to this study is now in progress at Loughborough (Gibson, 1978) aiming to test an interactive computer-based version of the schedulers' abacus in an operational job-shop.

It would seem that the stateboard concept is only really suitable for relatively simple systems where a limited number of alternative courses of action need to be compared. In those scheduling situations where the number of alternative courses of action is excessively large, the combinatorial explosion dictates that a computer-based aid is to be preferred. It is widely accepted that fully computerised scheduling cannot provide the degree of flexibility needed to cope with the unexpected, and consequently the use of man-computer interactive scheduling systems has been favoured.

2.1.2 Production scheduling

Ferguson and Jones (1969) give the earliest account of a scheduling aid in a computer simulation of a job-shop scheduling system. A 'simulation subsystem' (i.e. fast-time predictive mode) was included, enabling the user to ask "What if?" questions and thereby evaluate the use of various combinations of rules. The predictive facility was found to be useful during informal trials by managers and academics, especially in helping the scheduler to obtain the best compromise between short- and long-term decision criteria. Participants were universally impressed by the flexibility of the system. It is interesting to note that given the task of devising a manual schedule, most subjects first laid out some feasible schedule in a Gantt chart representation and then pushed, pulled, squeezed and otherwise manipulated their schedule. Those few subjects able to devise a manual schedule to equal their computer-aided equivalent could nevertheless perform the task much faster with the computer aid. Operators appreciated that with the computer aid they could generate alternative schedules instead of feeling constrained by time to continue working on their first schedule. This study demonstrated the advantages of interactive planning by a well-integrated combination of man and computer aid.

A further study by Jones et al. (1970) replaced the typewriter terminal used previously with a display terminal, which was found to facilitate communication between scheduler and computer. A shift in problem-solving technique also occurred: schedulers using the display tried out alternatives to find the best combination of rules, whereas those using the typewriter saw the computer as a means of evaluating rules and attempted to reason out the logical value of each combination. Tobey (cited by Hall, 1970) conducted a project at about the same time in which techniques of man-computer communication applicable to job-shop scheduling in a sizeable factory were explored. An on-line system was developed in which the scheduler selected one of a number of options available at any branch point in an extensive 'tree' of alternative displays and computer routines. Results stimulated local factory management to continue development of the system for their own computers.

Practical examples of computer-aided scheduling have included studies by Rice (1969), Godin and Jones (1969), Bollenbacher (1970), and Haider et al. (1977). Rice found man plus computer to be the best solution in scheduling the output from a corrugator - a machine which fabricates corrugated cardboard. The computer was used to research and review all possible combinations of orders, allowing the scheduler to choose between those combinations providing the most efficient corrugator runs. This approach proved to be superior on all counts to other fully computerised methods where selection amongst alternatives was performed by the computer. Rice's finding was later echoed by Haider and his colleagues in a study of a simulated job-shop scheduling problem. In Godin and Jones' study, up to 10^{62} possible permutations of operators, machines and coils had to be considered in the scheduling of a coil winding shop. Left to his own devices the scheduler could consider only a few alternatives, but with the aid of an interactive computer system having a simulation facility he could test alternative assignments and anticipate future problem areas. Bollenbacher also found that a computer scheduling system could ease small parts scheduling in a manufacturing company. The number of different parts and machines in use was so great that the human scheduler was incapable of analysing the situation over sufficiently long periods of time, resulting in poor schedules and excessively high levels of safety stocks. By using the computer not to replace the human scheduler, but merely to relieve him of the tedious calculations he had previously been obliged to perform, smoother schedules were produced. Forward production planning was also possible - different levels of production needs could be input to the program to find when operators, machines and shift numbers would change.

There have been a number of other applications of the man-computer interactive approach in recent years (for example Mason, 1976; Petit and Favrel, 1976). Brewer (1971) has described the NASA computer-aided system for world-wide resource allocation and scheduling. When the user had specified his desired schedules, computer algorithms searched for and reported any 'schedule conflicts' and the user could modify his input accordingly. Wilkinson (1972) has presented a theoretical discussion of three iterative man-computer dynamic scheduling systems. A set of computer algorithms again calculated a set of optimal alternative routes for the human scheduler to decide between.

More recently, Smith and Crabtree (1975) used a scaled down version of a job-shop scheduling problem modelled on the computer. Subjects acted as supervisors of the job-shop and their objective was to manufacture a number of items specified in a hypothetical order book by set target times, using information on machine states and material distributions displayed by the computer. A predictive facility was again provided, enabling the scheduler to ask "What if?" questions by running the simulation forwards or backwards through time thus rapidly evaluating the consequences of his decision strategies. Results showed that extensive use was made of the predictive mode when getting the feel of the system, for investigating different control strategies and lastly as a short-term 'error-correction' or steering procedure. Schedulers with the predictive aid performed only marginally better than those without (not statistically significant, perhaps due to the small number of subjects used). Furthermore the search strategies of both groups were similar. Smith and Crabtree suggest that the high task complexity and system lags together with short-term memory limitations prevented the deep level of search necessary, and may account for the small performance difference between groups.

Smith (1976) has also looked at how solution search is affected by different representations of a problem environment. One group of subjects undertook a resource allocation problem using a PERT-type network graph, another group were given a mathematically identical problem formulated as a warehouse packing task using a pattern representation. Both groups had access to a simulation facility which enabled them to investigate various alternative actions in their search for an optimal solution. The pattern group performed better than the network group, largely as a result of concentrating their efforts on those areas likely to yield the most profitable outcome. The difference was essentially one of a 'wide' versus a 'deep' search strategy.

2.1.3 Air Traffic Control

Rouse (1970) in the second of a series of experiments devised an application of predictive scheduling in a simulated ATC task. Subjects were required to marshall and guide three aircraft through a runway 'gate' by specifying correct headings and speed commands. A computer generated predictor display showed calculated aircraft paths 20 seconds into the future, and this was compared with a no predictor condition. Though the predictor display had yielded better performance in a prior experiment involving guidance of a single aircraft, no significant differences were observed when the task consisted of the more complex simultaneous guidance of three aircraft. What performance differences there were between aided and unaided conditions narrowed with practice. Eventually subjects used the predictor only as a checking device and ignored it completely when overloaded. It should be noted that Rouse's study was unique in that it involved a double-interaction situation. While interacting directly with the computer, subjects did not interact with or control the aircraft directly but gave verbal commands to a 'pilot' (i.e. experimenter) who operated the controls. This additional 'link' could well have contributed to the negative findings of the complex task.

Rouse's findings were echoed by Kreifeldt and Wempe (1974). Short-term tactical path predictors were evaluated as part of a complex study of distributed versus centralised ATC procedures. The pilots clearly preferred a display that included a flight path predictor for their own aircraft, but this did not significantly improve their performance. A more advanced Interactive Conflict Resolution (ICR) system has since been developed by Ball et al. (1975) working at the Royal Radar Establishment, Malvern, and makes use of a fast-time computer model of aircraft trajectories to predict airspace infringements (conflicts) between aircraft in an airways sector. The RRE method calculates a long-term prediction (up to 20 minutes ahead) of aircraft trajectories based on a probabilistic method, so that confidence limits can be placed on the predictions. The air traffic controller therefore has ample opportunity to evaluate alternative strategies by 'game playing' with the computer before implementing one of them. Ord and Whitfield (1977) report that results from a small-scale simulation experiment to evaluate the system from both objective and subjective viewpoints have been encouraging.

2.1.4 Soaking pit scheduling

Inefficient scheduling of the soaking pit complex accounts for perhaps as much as 10% of the total running cost of a steel plant, so any improvements are likely to be of significant cost-benefit. One would perhaps expect predictive techniques to make a valuable contribution in this area. Ketteringham and O'Brien (1970, 1974) at BISRA undertook a simulation study of computer-aided soaking pit scheduling in an attempt to evaluate an integrated manual and automatic scheduling system under comparatively realistic conditions. The main objective of the soaking pit controllers in their task was to keep a sufficient flow of correctly heated steel ingots to a costly rolling mill. So that no delays in the mill operation were caused, decisions were concerned with allocating a complete cast or part of such a cast to one of a number of soaking pits between the melting shop and en route to the rolling mill. Each pit had different characteristics and each ingot a different duration of soaking so there were problems in predicting the long-term effects of a decision. Under the old system the schedulers used their experience and a few crude rules-of-thumb to satisfy short-term, cost-based objectives, but were incapable of calculating the longer term effects of their decisions. They tended to adopt 'safe' (conservative) strategies which failed to maximise throughput, but by using a computer-based system model it was possible to investigate the outcome of different decisions and thus satisfy long- as well as short-term criteria by 'game playing' with the system and planning pit changes to match rolling mill requirements. Information on pit states was displayed to the scheduler on a c.r.t. screen. He could communicate with the computer using a touch-wire display.

Results of limited simulation trials showed a substantial increase (10.5%) in the average number of finished ingots rolled per shift due to a significant reduction in mill delays and pit over-soaks, indicating the potential of the system over existing scheduling methods. Operator's comments also indicated that the system relieved the tedium of previous pencil and paper approaches, and left them free to concentrate on the scheduling. The findings from field validation trials, however, were more equivocal (Bibby, 1974). Although the technical performance of the system was satisfactory (the scheduler's display being particularly successful) only a very marginal increase in steel throughput was achieved during the trial period. In general, the level of improvement that the simulation trials had led management to expect was not achieved. McEwing (1977) notes that the major cause lay in the scheduler's inability to make full use of the system's planning facility. Schedules developed using the predictive facility frequently had to be abandoned because the expected patterns of soaking pit availability and steel arrival at the pits were not achieved. This was due not only to unpredictable plant breakdowns (representative patterns of which had been simulated), but to organisational problems such as poor co-ordination between adjacent areas, slow communication of decisions (there was no direct link between the rollerman and the soaking pit controller) and poor working relationships between key job holders. The simulation had not encompassed organisational as opposed to technical realism - the role of operators in adjacent areas had been played by members of the experimental team. Had a full task analysis been carried out initially before the simulation study, and the scheduling aid modified (e.g. by using probabilistic arrivals information) to take into account this additional uncertainty, then results from the field trials may have been less disappointing. Bibby comments that a computer system with information displays linking more parts of the system, plus a more efficient communications network with up-to-date data entry, could help realise the potential of the scheduling system, as well as improving the working climate in the plant. A major warning from this series of studies is the way in which the ill-defined, uncertain nature of real-world processes can negate findings from well-defined simulation studies. If at all possible, steps should be taken to incorporate the possible effects of such factors at the simulation stage.

Following on from the above studies, Laios (1975)* developed a laboratory scheduling task similar to Ketteringham et al's and used it as a basis for his 'Predictive Computer Display' (P.C.D.). This combined a dynamic version of Shackel's (1976) stateboard concept with various ways of conveying information on the probability distribution of ingot arrivals (see Herman et al. 1964). As in Ketteringham's task, the scheduler endeavoured to maintain a constant time interval, this time of 5 time units, between departures from the soaking pit complex. This he did by judicious allocation of ingots to soaking pits. Judgements were made on the basis of displayed estimates of ingot arrival times, updated regularly. At the same time system constraints had to be satisfied, such as allocating ingots to pits as soon as possible after their arrival, and leaving a gap of 3 time units after a pit emptied before it could be reloaded.

Laios' early work looked at the effect of uncertainty associated with ingot arrival time estimates on <u>unaided</u> decision performance, using an arrivals display (F3) and a pit state display (F2) only. A pilot study with 6 subjects showed that unaided scheduling performance was degraded by the introduction of a moderate level of uncertainty. Use of the basic displays provided was higher with uncertainty present. These preliminary findings were confirmed in his first main experiment using 12 subjects (see Laios, 1976). The effect of 0 bits (i.e. absolutely certain), 3.4 bits and 4.4 bits of uncertainty measured with respect to a common reference point was examined, again using F3 and F2 displays only. A significant decrement was found in the performance index (relative number of scheduling errors) upon

---

* since summarised in Laios (1978).

introduction of a moderate level of uncertainty. However, no further decrement was observed when uncertainty was increased still further - in fact a slight, though non-significant, improvement occurred. A similar significant increase in on-line activity was observed with the introduction of uncertainty, but on-line activity was not affected by changes in the amount of uncertainty. The type of strategy adopted by the operators was held to account for the lack of effect when uncertainty was varied.

In a second experiment a predictive facility (PCD 1) was introduced with which the operator could plan in conjunction with the F3 display allocations and utilizations up to 35 time units ahead of current time. The effectiveness of the predictive display and an optional heuristics display were tested under deterministic conditions (i.e. no uncertainty present) using 8 subjects. The predictive facility resulted in a significant improvement in the performance index and a concomitant increase in on-line activity. Subjects 'decision horizons' (how far ahead they were planning) showed them to be looking well ahead and attempting to satisfy long-term as well as short-term criteria. The optional heuristics display had no apparent effect.

The latter study was repeated under a moderate level of uncertainty in Laios' third experiment. Though no statistically significant improvement due to the predictive display was observed, a marked bias in the averaged performance scores was evident in favour of the predictive condition. Laios attributes this to the high variability of individual scores. A statistically significant increase in on-line activity was, however, observed with the predictive facility. The heuristic display again had no apparent effect. Laios concludes that deterministic decision aids, i.e. those designed for certain environments, may be of little use in uncertain real-world conditions.

Consequently, in his final experiment Laios set about examining different ways of presenting information on the uncertainty characteristics associated with ingot arrivals. Two new displays were developed. The first combined numerical estimates of ingot arrival times (F3) with the previous predictive display (PCD 1) to form a new combined display (PCD 2). The second display went one stage further by displaying the actual interval within which an ingot would arrive, as opposed to a numerical estimate (PCD 3). In both displays, the need to switch to a separate display to infer ingot arrival times was eliminated. The experiment was again conducted under a moderate level of uncertainty using 8 subjects in a repeated measures design. Results indicated that the introduction of arrival intervals showing the extent of the uncertainty present gave a significant improvement in performance. However, reducing the width of the intervals by approximately 20% using Bayesian mathematics made no statistical difference - in fact performance became slightly worse. This result casts doubt on the validity of the Bayesian approach in these situations. Laios attributes the improved performance scores directly to the provision of interval information about uncertainty characteristics, coupled with the subject's increased ability under these conditions to delay ingots and so achieve a smooth output flow. It is interesting to note that on-line activity was not affected by the different information presentations.

Laios' study represents the first systematic investigation of factors affecting scheduling performance. His results indicate that not only can uncertainty detract from the effectiveness of deterministic decision aids, but that the display of information in a suitable form about that uncertainty might be a useful approach in real-life scheduling applications where uncertainty is an inherent system component. Laios' work has been quoted in detail as it is of direct relevance to the experimental material of Chapter 3, and in addition was the original stimulus which led to the questions posed in this thesis. Laios' work does not, however, present a complete picture. As Whitfield (1975) has commented: "Further research is obviously necessary to investigate the effects of variations in the environment and within the predictive model, in aided scheduling tasks". This point will be taken up in Section 3 of this chapter.

2.1.5 Summary

Discrete decision aids have been shown to assist scheduling performance in a wide variety of industrial and non-industrial situations. Two basic approaches may be delineated. In the first, the human operator is free to select from a computer-derived set of optimal schedules, calculated according to criteria expressed as a formal algorithm. Examples include the studies by Rice and Bollenbacher (op. cit.). The second approach allows the scheduler to 'game play' with the computer and explore the consequences of individual decisions. This approach has been favoured by behavioural scientists, and examples include the studies by Smith and Crabtree, Ketteringham and O'Brien (op. cit.). A common finding in this area of research is the large variation between individual subject's strategies. Further research is necessary to examine the behavioural effects of variations both within and outside the predictive model on aided scheduling performance. The present author has attempted such an investigation, described in detail in Chapter 3 of the present thesis.

2.2 Continuous control applications

Since its arrival in 1958, the predictor instrument has been proposed for use in a wide variety of continuous control systems. Kelley (1968), Warner (1969), Smith and Kennedy (1975), and Bird and Whitfield (1975) have reviewed the available literature. It appears that most of the experimental work has concentrated on simulation studies of military and vehicular systems, to the apparent exclusion of any operational applications (though several have now been planned).

This is not to say that the value of predictors is unproven. The literature abounds with simulation studies demonstrating the efficacy of the predictor technique in situations as far removed as piloted aircraft landing, submarine depth control, lunar rovers, spacecraft guidance, docking ocean-going vessels, VTOL and helicopter hovering, and remote control systems: in short, any situation where system dynamics and a man's inherent limitations combine to render unaided control difficult or impossible. And yet in spite of the overwhelming evidence in their favour "no operational applications have yet been made, or no documentation of such applications exists" (Smith and Kennedy, 1975).*

---

* Unconfirmed reports (Pitrella, 1978) indicate that a form of predictor display has been included as part of the standard instrumentation in the Douglas DC-10 aircraft to assist pilots to monitor the automatic landing system. Other operational applications may well exist in the military, but for security reasons are not documented.

The following subsections present a brief review of some of the principal application areas which have been envisaged. Though the applications are numerous and cover a wide range, it should be noted that many of the processes studied have similar transfer functions and are thus equivalent in control engineering terms.

2.2.1 Piloted aircraft

Judgement of the point on the runway at which it is safe for a jet aircraft to take-off is by no means a simple matter for the pilot, being dependent upon several parameters. Hainsworth and Olinger (1958; cited by Warner, 1969) were amongst the first to envisage the use of predictive information in their proposed Safe Take-off Predictor (STOP), which would provide a pilot with current aircraft position, a predicted take-off point and a last safe stop point, all on a single scale. Price, Honsberger and Ereneta (1966a,b) also foresaw the need for predictive information in the various flight management and control functions associated with aircraft flying at supersonic speeds.

In a study by Sweeney, Todd and Heaton (1965) a predictor display was found to yield superior results over conventional displays for altitude control of high performance aircraft. The authors also found an improvement in a simulation of a low-level terrain following task, provided the dynamics of the automatic flight control system were included in the predictor model. A synthetic predictor display showing the aircraft velocity vector was eventually proposed, owing to the high computational requirements needed to generate the predicted path. Williams (1969) also found improvements in low-altitude, high-speed flight using a predicted pitch display.

Vertical take-off and landing (VTOL) aircraft are characterised by some very difficult control problems, notably in the transition phase from horizontal to vertical flight. In a VTOL simulation study, Kemp (1969) found that test engineers were unable to control a hovering aircraft without an altitude prediction display. This problem has also been extensively investigated by Bernotat, Dey and their West German co-workers, using Taylor series prediction models. Collectively they found that even naive subjects could stabilise pitch/roll axes using a predictor display, but were unable to do so properly without it. Dey and Johannsen (1969) report an 80 per cent reduction in training time using the predictor. A consistent enhancement of control performance is also found. Dey (1971) for example obtained a reduction from 7.92 to 2.48 in the mean square deviation of a simulated VTOL from a desired 'course' using a Taylor series extrapolation model. Gallaher et al. (1976) also report favourably on the use of a Bernotat-type regression approach on each axis of a 6 d.f. simulator to generate a composite display of future aircraft position.

Landing on the undulating deck of an aircraft carrier poses another complex control problem for the pilot. Wulfeck, Prosin and Burger (1973) used a sophisticated 6 degree of freedom F.4 carrier landing simulation to represent night landing on a carrier at sea. Pilot approach and landing performance on the simulator using existing operational landing aids was compared with performance using a predictor display. Though differences in mean altitude and lateral errors were negligible between the displays, the predictor display reliably produced substantially smaller altitude and lateral error variances than the conventional display. The experienced Naval pilots used as subjects could land their aircraft much more accurately using the predictor display. In addition the finding that variability amongst the predictor pilots was considerably smaller suggests that in its real world application use of the predictor instrument would lead to fewer accidents and abortive landing attempts. The somewhat simpler task of landing on a conventional runway was also found to be facilitated with a predictor display (Smith et al., 1974; cited in Smith and Kennedy, 1975). It is interesting to note that learning from the predictor display transferred to a conventional non-predictive display, so that those subjects trained with the predictor performed better subsequently than those subjects trained on the conventional display. Kennedy et al. (1975) in a carrier landing simulation study found that a predictor display also facilitated landing performance by pilots both experienced and inexperienced on F.4 carrier night landings. Although the experienced pilots had achieved better performance than the inexperienced pilots in a control condition, performance levels with predictive aiding were equivalent for the two groups, again demonstrating the potential of the predictive display as a training device.

Kreifeldt and Wempe (1973) found that professional pilots could more accurately perform a standard instrument procedure turn on a simulated plan view display of the ground beneath using a predicted ground path trace than without such a predictor. Wempe (1974) further notes that using such a 'Horizontal Situation Display' (HSD) together with a flight path predictor which calculated the effect of airspeed, bank angle and wind, not only was the lateral error when turning in the presence of wind gusts reduced by one-half, but that individual differences amongst pilots also decreased. Pilot acceptance of the flight path predictor was good. The map usually depicted on an HSD can have several orientations: north uppermost, fixed map with moving aircraft symbol or translating map with fixed aircraft symbol. Baty (1976) found that a flight path predictor helped improve the performance equally well in either a 'north-up' or a 'heading-up' map orientation.

Rouse (1970) in a further study of map-based predictors asked subjects to guide a simulated piloted aircraft through a runway approach 'gate' with correct speed and heading for the aircraft runway. The predictor yielded better performance than conventional display and faster learning of the task, but the differences between performance with and without predictive aiding narrowed as learning proceeded. It should be noted that Rouse's experimental set-up was a-typical in that a double interaction situation was present - whilst interacting directly with the computer subjects did not control the aircraft directly but gave verbal speed and heading commands to a 'pilot' (i.e. experimenter) who 'flew' the plane.

Warner (1969) simulated a minimum-time terminal control task for a pure inertia system and reports that an off-line exploratory predictor gave slightly better results than its on-line equivalent (though not statistically different), assuming that sufficient time was available to use it. Performance without any form of predictor was more variable, and in general, worse.

2.2.2 Spacecraft

Guidance during various stages of a spacecraft mission presents a unique series of control problems, many of which the predictive display concept is particularly well-suited to deal with. In a study of predictive aiding during the launch phase, Gilchrist and Soland (1967) used a technique called a Predictive Model Guidance Scheme employing a fast-time model to generate a predicted trajectory in an altitude vs.

velocity display. The pilot's role was one of indirect participation through adjustment of the initial conditions of an optimal steering program. The authors felt that the predictive display helped the pilot to 'shape' the vehicle's trajectory. The use of a predictive display system by a range safety officer in observing the flight of a missile has also been proposed (Fogarty; cited in Warner, 1969).

Rendezvous between the orbiting spacecraft is a critical phase of many space missions, as the manoeuvre must be completed in a given time and with minimum fuel expenditure. Though actual docking can be accomplished using visual and radar-supplied cues, the approach to docking is more complex and requires special aids. McCoy and Frost (1965) in a series of experiments have shown that use of a predictor display facilitates performance in such a coplanar rendezvous situation. A first experiment compared display of current position only with a time-history display, and demonstrated that whilst rendezvous could be achieved in either case, performance was better with time history included. A second experiment compared a predictor display with a display of current position only. Again rendezvous was always possible, but with the predicted flight path fuel consumption was significantly reduced and smoother rendezvous trajectories were flown. Similar results were obtained by Mano and Ulbrich (1965), fuel savings being achieved with an exploratory predictor display. In addition these workers suggested that predictive techniques would be useful as a training device. Later work by McCoy and Frost (1966) compared on-line with off-line (exploratory) prediction. Off-line was found to yield better performance than on-line. The authors also reported that naive subjects were able to perform the rendezvous with essentially no training whatsoever, using the predictor display. A study by Shannon (1976) confirms this finding.

With the advent of manned lunar landings, manual control of the lunar module's descent phase is necessitated by the need for flexibility in the selection of a suitable landing site. Conservation of fuel is again a primary consideration. Fargel and Ulbrich (1963) simulated a lunar landing task and found that tracking accuracy and fuel economy were both enhanced using a predictor display in a height vs. rate-of-change-of-height configuration compared with a display of current values alone. Results were similar for both one dimension (altitude) and two-dimensional (altitude and translation) tasks, the latter proving harder to interpret as the two axes were presented orthogonally. Training times were such that a stenographer (typist) assisted by the predictor display was able to do better than unaided test pilots and astronauts after a few tries. Besco (1964) simulated a 3-axis spacecraft attitude control task and investigated control manoeuvres such as attitude hold, stabilisation and attitude change. Lower fuel consumption and r.m.s. errors were obtained with a predictor display incorporating pitch, roll and yaw information at 0, 10 and 20 seconds into the future than were achieved using more conventional displays. In addition the experienced pilots used as subjects preferred the predictor display.

The final re-entry phase of a space mission again comprises a complex series of decisions requiring the flexibility of human control. Several authors have investigated predictive displays for re-entry employing a Ground Area Attainable or 'footprint' display indicating the area on the earth's surface that can be reached by the vehicle, with the target destination marked (Wingrove and Coate, 1961; Austin and Ryken, 1963). Warner (1969) cites a study at Lear Siegler Inc., Grand Rapids, Michigan, which showed that a digital display of such predictive information allowed more precise control when the pilot had several tasks to perform concurrently.

2.2.3 Remote control systems

The transmission delays encountered, for example, when controlling a lunar rover from earth are in the order of seconds. Such pure time delays effectively reduce conventional displays to those presenting historical rather than current information. Cohen (1962) and Kelley (1963) were amongst the first to envisage the application of predictor displays in this context. Arnold and Braisted (1963) in fact demonstrated that predictive information could be used to compensate for these delays. A simulated lunar rover transmitted a delayed picture of the terrain ahead to a remote operator. A predictor display was available to the operator, the prediction span being set to equal the transmission delay. Performance using the predictor was approximately equivalent to direct, no-delay control with the result that higher lunar vehicle speeds than were previously possible could be attained.

Smith and Kennedy (1975) comment that earthbound remote piloted vehicles are receiving increased attention for military reconaissance and weapons system use. The authors report that in their work (Smith and Queen, 1975) a predictor display was demonstrated to assist flight profile control in a simulated approach-to-landing task.

2.2.4 Ocean going vessels

Many of today's ocean going vessels are similar to the aerospace applications previously discussed in that their response characteristics are complex and frequently involve large time lags. The supertankers of the seventies are a prime example. Kelley (1960b; cited in Kelley, 1968) carried out the first laboratory test of the predictor technique in an application to submarine control. Using a simulated high-speed submarine he found that levelling out after a depth change was a simple matter using the predictor, but was virtually impossible without it. Reductions in training time due to the predictor were also found. One display format presented the effect on predicted depth error of full rise, full dive and centering the controls. Alternative configurations are discussed by Berbert and Kelley (1962).

McLane and Wolf (1967) evaluated a symbolic predictor display compared with a symbolic quickened display and two forms of contact-analog display in a simulated submarine guidance and homing torpedo avoidance task. They found no statistically significant difference between the displays in terms of tracking error, but the predictor display was significantly superior to the others in the torpedo avoidance subtask as it resulted in fewer collisions.

Brigham (1972) has put forward two applications of predictive techniques in ship control. Collision avoidance involves extensive decision making and considerable skill in the interpretation of the ship's radar information. Several new radar systems offer automatic facilities to aid in collision avoidance. As well as displaying the present position of all targets, the Marconi 'Predictor' radar can show the target positions two, four and six minutes previously. In the 'predicted relative tracks' mode a proposed course and speed change can be entered. The tracks of all the targets then move accordingly so enabling the effectiveness of a proposed manoeuvre to be evaluated. A similar system, the AEI 'Compact' radar, projects a vector from each selected target to indicate course and speed. A collision warning alarm is given if any of the predicted tracks will come within the home ship's 'closest point of approach' circle during the next 30 minutes. A 'trial track' facility also enables course and speed changes to be evaluated with respect to ships and other possible hazards. Brigham noted that these radars are expensive and had not been extensively tested in sea trials at the time of writing.

In berthing a supertanker and during manoeuvres in restricted waters the pilot and master have to judge distances, velocities, accelerations and rates of swing with an accuracy below the minimum psychophysical threshold which can be perceived directly. Their judgements are therefore of necessity based on a shortage of information. Brigham (op. cit.) propose the use of a predictor display as an aid to berthing. Such a display would show current positions of jetty and ship with predicted ship's positions at chosen time intervals ahead. Corrective action could then be taken in advance of the ship's impending collision with the jetty. However, Brigham felt that the use of such a display was not feasible at that time due to problems in engineering a prediction model of sufficient complexity to incorporate all the possible system inputs (ship's position, equations of motion, effects of engines, helm, tugs, wind, tide and undercurrents) many of which cannot be precisely determined. A simple extrapolation model may be applicable here.

2.2.5 Process control applications

Control of many continuous industrial processes, notably chemical process control, is typically associated with long time delays and high order plant dynamics which place severe limitations on the predictive abilities of the human controller. There would appear to be considerable scope for predictive displays here. It is interesting to note that Ziebolz's original work was in the area of process control, though most subsequent envisaged applications were in the field of aerospace. The situation has now come full circle and Crawley (1968) forecasts that predictive techniques will have an important part to play in future integrated man-computer systems for the steel-making industry. As yet, however, application of such techniques (other than as parameter estimation within a totally automated control program - see Chapter 1, section 3.2) has been limited.

Blake (1968) in a discussion of display techniques for batch, continuous, semi-continuous and start-up processes has stated that for him "predictive displays are a most important part of any start-up system". Predicted values could be displayed on numeric readouts or on a line-drawing c.r.t. Typical messages that a predictive display could produce might be the speed a turbine would reach after a given time, or the fact that the safe operating temperature of a boiler would be exceeded if firing continued at the same level. Rasmussen (1968) has also suggested the introduction of a trend display or predictive facility into a c.r.t. display for power plant controllers. A visit late in 1975 to the British Gas grid control centre at Hinckley, responsible for monitoring and controlling the country's gas supply, revealed that a sophisticated, off-line 'Control Advisory Program'

(C.A.P.) was at the development stage. When operational the CAP would analyse and forecast the state of the grid system 24 hours ahead, and could also display the results of simulation runs. It was estimated then that the complete system would take 5 years to implement and debug.

McEwing (1977) reports that British Steel are currently developing "displays of continuous estimates of process states derived from a computer-based parameter estimation procedure and dynamic process model" for control of basic oxygen steelmaking. And Verhagen (1976) at a recent NATO Advanced Study Institute on Man-Computer Interaction reported a predictive control system for a simulated distillation column: at a later stage it was planned to display computer predictions direct to the operator.

Ratcliffe (1977) has also reported that an off-line conversational predictive model is under development for stock control of a complex continuous ammonia soda plant at ICI Mond Division. Lags made the process difficult to control, and limited stocking capacity resulted in stock control problems. When operational as part of a suite of supervisory programs, the predictive facility would enable the operator to investigate the effect of changes in plant throughput, on the stock levels in 5 storage tanks up to 12 hours hence. The program could be used to estimate stock levels if present conditions were maintained, or to establish the best way of shutting down the plant without emptying or overfilling the stock tanks, or to build up levels prior to a maintenance shutdown. The predictive program could be accessed from a VDU terminal, changes in throughput entered numerically, and the predicted alterations in stock levels displayed graphically or in tabular form. Ratcliffe sets out a set of guidelines for implementing such a system, and it will be interesting to see what operational benefits result.

2.2.6 Summary

Surveying the reported literature on predictive display research, some common findings emerge. First and foremost, human operator effectiveness is improved through the use of predictive displays in a variety of control tasks, such as controlling non-linear systems or linear systems with pure time delays. Learning times are reduced, often to the point where novice operators are able to control relatively complex systems with essentially no training whatsoever. Several workers (Rouse, 1970; Bernotat, 1972) have commented that the eventual stabilised performance levels of aided and unaided groups are similar. The aided group, however, achieves this performance level long before the unaided group. Predictor-aided control can approach optimal control with respect to a particular performance criterion as the operator is able to plan the best courses of action. Lastly, the information processing requirements on the human operator can be reduced, notably in multi-dimensional control tasks.

It must be stressed that the experiments reviewed have been largely restricted to specific military or vehicle simulations. No general, cross-system studies appear to have been conducted. From the literature, it seems that those applications most likely to benefit from predictor displays include at least one of the following characteristics (Warner, 1969):

1) the dynamics of the system to be controlled are both complex and slowly responding;
2) the number of separate tasks to be performed is relatively large;
3) the nature of the task requires considerable anticipation by the operator;
4) the task is system-paced, i.e. time constrained;
5) flexibility of control actions is required.

Several potential applications having one or more of these characteristics can be envisaged in addition to those already discussed. Control of automobiles might be facilitated with a predictive display, especially for learner drivers. In railway sidings, a predictor display could assist smooth and precise shunting of wagons so preventing damage to rolling stock and contents caused by the intermittent application and release of brakes.

It is evident from section 2.2.5 that predictive displays are a major growth area for the control of chemical plant. At the time of writing, however, the present author has carried out what is believed to be the only known comprehensive experimental study and field validation of the predictive display in this area. This work is described fully in Chapters 5, 6 and 7 of the present thesis.

3. FACTORS AFFECTING PREDICTIVE DISPLAY PERFORMANCE

Whatever the application, there are certain parameters internal and external to the predictive display system which will affect the operation of the entire control or decision-making process. These parameters can be defined as follows:

1) Input uncertainty - a measure of the accuracy of the updating information fed to the prediction model. This ih turn is dependent on the degree of signal contamination due to noise in continuous systems, the unreliability of input information in scheduling applications, or simply the normal variability inherent in the system operation being controlled.
2) Prediction span (extrapolation interval) - the real time period over which predicted plant response is displayed. This is often the same as the prediction time - the real time interval over which predicted plant response is computed by the prediction model - particularly where a single predicted end-point is displayed.
3) Prediction model fidelity - the accuracy with which the prediction model represents the controlled system's behaviour. Usually expressed in terms of Bernotat and Widlok's (1966) three stage classification. Prediction model fidelity can be thought of as internal inaccuracies within the predictive display.
4) Process dynamics/response characteristics - for continuous systems, this is usually expressed in terms of the plant gain (K) and the control order or effective number of integrations in the control system. In scheduling applications, response characteristics are a function of the complexity and speed of the process.

5) Repetition (refresh) rate - the number of successive predictions displayed to the operator per unit of time. It is often the frequency at which the prediction model is updated with the present state of the plant.
6) Mode of control - refers to whether the predictive display is arranged in on-line or off-line configuration.

The above characteristics also apply in principle to fully automate predictive control systems, where man is replaced by a logic decision element, as well as to manual control systems where the predictive display is employed to extend man's capabilities. Factors peculiar to manual systems not listed above mainly involve display format issues, and the question of which system variables need to be displayed to the operator. The answers to the purely 'knobs and dials' ergonomic questions can frequently be gleaned from the body of existing knowledge available in such human factors texts as van Cott and Kinkade (1972) or McCormick (1976). Otherwise a system specific experimental study may be necessary. It is also important to determine which system variables need to be displayed to the operator, as displaying data which has only potential relevance is not only ineffective but can actually degrade performance (Baker and Goldstein, 1966).

Warner (1969) and Smith and Kennedy (1975) have pointed out that no general quantitative studies of predictor characteristics have been conducted. The need for such studies was noted by Sheridan as long ago as 1962: however, with the exception of a few 'mini studies' little seems to have been done to meet this need. As the following discussion of predictive display characteristics implies, there is clearly much to be learned in this area.

3.1 Input uncertainty

This refers to the accuracy of the information fed to the prediction model, and is not the same as prediction model fidelity (discussed under 3.3). Input uncertainty is caused by normal variability in system operation or by external disturbances to the controlled system (e.g. cross winds affecting aircraft flight, transmission noise on signal lines, unreliable information fed to a production scheduler). The net effect of these variations is, however, similar to prediction model inaccuracies in that they both serve to reduce the credibility of the predicted information displayed to the operator. If the uncertainty cannot be incorporated into the predictive trace, for example because its form is entirely unknown or because of limitations within the prediction model, then there will be a discrepancy between actual and predicted paths which can only serve to mislead the operator. In this case the useful prediction span may have to be reduced. If however the nature of the uncertainty can be forecast and incorporated into the predicted trace then the display has the addition of a diagnostic feature. One suggested approach is to display multiple predicted paths corresponding to the mean predicted path with extreme ranges to either side. There has been little quantitative research in this area (see Tainsh, 1977). Another approach, adopted by Herman et al. (1964) and Laios (1975), has been to display probabilistic information showing the range of values within which the actual input may lie.

It is well known that in non-predictive control systems the main effect of uncertainty is to degrade performance. Howell and Briggs (1959) for example looked at the effect of visual uncertainty in the form of perturbations on the display signals in a pursuit tracking task where input, response, and input and response together could be perturbed, and in a compensatory tracking task where the error signal was perturbed. In general, tracking performance became worse as the level of uncertainty increased. It is interesting to note, however, that tracking performance was not noticeably affected when noise was present on the response display only: apparently reliable visually coded feedback information is less critical to the operator than is accurate input information.

At a psychological level, Garner (1962) has proposed the existence of uncertainty as a fundamental psychological concept. Kahneman and Tversky (1973) note that in making predictions and judgements under uncertainty, people do not appear to follow the calculus of chance or the statistical theory of prediction. Instead they rely on a limited number of heuristics which sometimes yield reasonable judgements, but which sometimes lead to severe and systematic errors.

In a study of unaided dynamic decision-making, Rapoport (1966) found that the introduction of a stochastic (uncertainty) element into his version of the Reader's Control Problem resulted in a 50% deterioration in performance. Ebert (1972) however failed to replicate this finding in a complex production scheduling problem, possibly because his subjects had detailed information about the stochastic element throughout the task, or because the task itself was of greater complexity. Levine and Samet (1973) and Laios (1976) however both found that the introduction of uncertainty brought about a worsening of decision performance. Laios also found that increasing the level of uncertainty beyond a certain point resulted in a slight improvement in performance.

Given that uncertainty degrades unaided control and decision performance, one would perhaps expect it to have a similar effect on performance when predictive displays are employed. Laios (1975) has presented one of the few experimental studies on the effect of uncertainty in a predictive display system, and because of their importance to this thesis his experiments were reviewed in detail under section 2.1.4 of the present Chapter. Bernotat and Widlok (1966) also describe a noise-contaminated continuous predictor, but do not report empirical findings. In fact, Laios found that the level of his subjects' performance with uncertainty was considerably worse compared with performance under deterministic conditions. Bibby (1974) reports a similar finding when a predictive scheduling aid designed and tested under near-deterministic laboratory conditions was implemented in the uncertain real-world environment of a steelworks soaking-pit scheduling complex. However, it should be noted that Laios evaluated directly only a single value of uncertainty, and there is clearly a need to establish the effect on predictive display usage of a wide range of uncertainty values. It would also be highly interesting to investigate the interaction of uncertainty with such predictive display parameters as prediction span and prediction model fidelity.

3.2 Prediction Span

The evidence relating to how far ahead prediction should extend is conflicting. One would expect the length of useful prediction span to be affected by other system characteristics. Kelley (1960a) in an early predictive display study found that whilst approximate prediction models could be of some assistance, useful prediction spans decreased with decreasing model fidelity and learning times for effective control were increased. Bernotat and Widlok (1966) however report the opposite.

As the order of their extrapolation model was reduced, useful prediction times increased by a few seconds.

Subjects in a submarine control task (Kelley, 1960b) when permitted to adjust prediction span elected to reduce it as vehicle speed was increased. Kelley (1962) thus recommends that slow, sluggish systems such as submarines are best served by a long prediction span (25 to 30 seconds), whereas quick-changing, high-frequency systems such as helicopters require a shorter span (5 seconds). Dey and Johannsen (1969) on the other hand suggest that the faster the control task, the longer the prediction time span should be. Dey (1971) also found optimum prediction time to increase as the controlled process increased from a second to a third order system. The latter authors were concerned with extrapolative predictive displays for VTOL aircraft hovering.

In the limiting case, extremely short spans provide insufficient information and control instability ensues. An unnecessarily long span is, by definition, unnecessary and may even act as system noise. Whether such noise distracts the operator and degrades his performance or merely acts as superfluous information is unknown, the research findings being inconclusive. Rouse (1970) for example found that a 40 second span yielded worse overall performance than a 20 second span in an aircraft guidance task, as subjects wasted time correcting distant errors that would never arrive. This illustrates the concept of an 'optimum prediction span'. Rouse's results may be a-typical, however, in that his task was a 'double-interaction' situation: subjects acting as air traffic controllers gave heading and speed instructions to a 'pilot' flying an aircraft. Williams (1969) using an aircraft predicted pitch display reported that performance remained the same with spans of 3.5 to 6.75 seconds, but deteriorated for spans of less than 3.5 seconds.

Besco (1964) in a simulated spacecraft attitude control task tried prediction spans from 10 to 30 seconds but found no significant difference on performance, perhaps due to inaccuracies in the prediction model used. McLane and Wolf (1967) investigating predictor displays for submarine course and depth control also reported no significant difference between prediction spans of 20, 30 and 40 seconds, though the 40 second span did result in larger overshoots. There was also some evidence that had a more stringent tip-of-predicted-path-in-circle tracking task been employed, r.m.s. error would have risen with lengthening prediction spans.

In a study of a simulated jet aircraft landing by Kennedy et al. (1975) control performance not only increased sharply as spans increased from 5 to 20 seconds but there was an indication that much higher performance would occur with even longer spans. Yet in a follow-up study using 'experienced' subjects from the first experiment, the authors found no difference between spans of 10, 20 and 30 seconds (Smith and Kennedy, 1975). Perhaps a wide range of spans may be equally effective for experienced operators. Smith and Kennedy note that their experience in the Dunlap Labs, where Kelley also carried out much of his work, indicates that operators make use of the first or central segment of a predictor trace rather than its end-point. This procedure effectively minimises the time to reach the desired trajectory. In cases where time is not critical, however, there is probably no advantage in using any particular segment of the trace, again suggesting that a broad range of prediction spans may be equally facilitating.

Notwithstanding the above, practical considerations usually require the selection of one (or several alternative) prediction spans, unless the operator is given the freedom to adjust the prediction span for himself. Different systems will undoubtedly need different prediction spans, probably related to the 'responsiveness' of the system and to the magnitude and frequency of unpredictable disturbances. Bernotat (1972) in this context comments that the proper choice of prediction time can improve performance by as much as 70%. Kelley (1960b) has noted that for some tasks span should be in terms of distance rather than time. Though Rouse (1970), Dey (1971) and Bernotat (1972) have all found optimum prediction spans/times in laboratory simulation studies, currently there is no research which points to optimum spans for any operational system.

3.3 Prediction Model Fidelity

A 'perfect' predictor instrument is one which predicts the future state of a controlled process, by displaying to the operator one or more future states in addition to the present system state. As was previously noted, three classes of prediction fidelity have been put forward by Bernotat and Widlok (1966);

CLASS I PREDICTION uses a mathematical power series to extrapolate repetitively from the current value of the controlled system and its derivatives. System time-history is ignored in favour of present movement. As the prediction process is independent of system characteristics, Class I prediction cannot be used to predict accurately far into the future. Its necessarily short prediction spans are applicable to stabilisation and guidance problems, the absolute value of the prediction span being dependent on the dynamics of the system and its external disturbances. This technique has been widely used by the West German school (Bernotat, 1972; Dey, 1972). Bernotat and Widlok (1966) note that although the extrapolations are not as accurate as (for example) Class II prediction, because man extrapolates very coarsely the method is more accurate than anything the human can manage and in most cases will suffice to provide some lightening of the load. Some degree of model inaccuracy can also be tolerated due to the human operator's adaptability.

Extrapolation according to this method is a problem of approximatior the function being approximated by a power series. A Taylor series expansion is typically used, of the general form:

$$y(t + \Theta) = \sum_{n=0}^{N} \frac{\Theta^n}{n!} \cdot \frac{d^n y(t)}{dt^n}$$

where N is the number of derivative terms in the extrapolation. Dey (1972) notes that the order of extrapolation should be one less than the order of the controlled system. For example, a controlled process having three integrations in the forward path would require a second order extrapolation. It is a property of the Taylor series expansion that deviations from the true path will increase with increasing prediction time. Accuracy on the other hand depends on how many terms of the series are used, and this is limited by the technical possibilities

It is also influenced by the level of noise contamination, which will be amplified as N, the number of derivative terms, is increased.

CLASS II PREDICTION differs in that it assumes the controlled system's transfer function or response characteristics are known and can be included in the prediction model. Usually an analog model of the process (though there is no reason why this should not be accomplished digitally) is run in 'fast-time' alongside the real-time system to be controlled. The fast-time model is fed with exactly the same control inputs as the real-time process, and so extrapolates the predicted path of the system from its present state. Because of its greater accuracy Class II predictive displays permit prediction further into the future than Class I. However, the two-time scale modelling technique does not achieve perfect extrapolation since it does not include factors external to the system; hence unlimited length prediction spans are not possible. Its applications include longer term stabilisation and guidance problems, and this approach has been developed and widely used by the American school, notably Kelley (1958, 1960a,b, 1962, 1968, 1972) and his colleagues. Kelley's original design of predictor instrument has been fully discussed in section 1.2, and will not be repeated here. Most of the applications cited in section 3.2 centre around Class II instruments. It is worth pointing out that the Class II approach can be thought of as providing the best estimate of all the terms in a Class I Taylor series expansion.

CLASS III PREDICTION approaches the hypothetical perfect predictor in that important external disturbances which are to some degree predictable are included in the fast-time model. Obviously the incorporation of all possible system disturbances (in a space mission for instance) tends to stretch computing facilities to their limit, and therein lies the principal drawback of this method. Class III predictive displays can, however, be used to extrapolate far into the future, and their long and accurate prediction spans provide a useful navigational feature.

In practical terms, Class II prediction will usually suffice for most operational accuracy requirements since rapid updating of predictions will tend to offset external disturbances. However, if Class I can be successfully used a substantial saving in computational power will result. There has apparently been little empirical research on what level of prediction model fidelity is required in relation to other system characteristics.

3.4 Process dynamics/response characteristics

The human operator demonstrates considerable talent in predicting the response of quite complex systems - but only when he has had a good deal of training. As the complexity of the plant increases, however, it is not always possible for the operator to form an accurate mental model of the system and his predictive abilities are impaired. It is at this level that some form of control or display aid, such as a predictive display, is necessary. But how is performance with such an external aid affected by changes in system gain and control order? One would perhaps expect performance to be adversely affected for very fast or very slow systems, as fast systems are uncontrollable and slow systems move too slowly for changes to be noticed.

Several workers have addressed experimentally the question of plant gain and control order. Warner (1969) reports that in his terminal control task operator performance was independent of system parameters over the ranges investigated, i.e. from high gain (short response times) to low gain (long response times). However, the more important system orientated performance measures did show a dependence upon system parameters. It appears that the sensitivity of the performance measure(s) to control action timing is an important issue, and that performance evaluation is more representative if based on overall system measures rather than just the operator's behaviour. Bernotat and Widlok (1966) found that although their extrapolative predictor display improved the quality of control over a no predictor condition for all values of plant gain investigated, the improvement was most perceptible in the area of high gain. In absolute terms, error scores using the predictor reached a minimum level for medium gain values and rose at low and high gains.

As far as control orders are concerned, Bernotat and Widlok (1966) report that in their stabilisation problem the greater the number of process integrations, the larger was the benefit obtained from the predictor. Bernotat (1972) notes that, as control order increased from two to three integrations, errors rose considerably as did the amount of control effort required of the operator to achieve that level of performance. Bernotat comments that depending on the task an error/control effort trade-off may be expected. In some tasks, minimum error scores must be achieved whatever the cost, whereas in others a higher degree of error can be tolerated but operator effort (which is usually synonymous with fuel consumption) must be kept to a minimum.

Rouse (1970) summarises the situation when he suggests that predictive aids may be beneficial only in tasks of medium difficulty: they are strictly unnecessary for easy tasks, and in very difficult tasks the operator is so overloaded that he has to ignore the information. Rouse's experimental evidence tended to support this suggestion. It is evident from the foregoing that the effectiveness of the predictive display concept is a function of plant controllability. It remains to be established whether the relationships between other predictive display parameters hold for different values of plant gain.

3.5 Repetition Rate and Frequency of Updating

Repetition (refresh) rate of the display is the number of successive predictions displayed to the operator per unit of time. In theory for fast-time models it is determined by the prediction model time scale, the prediction span, and a negligible amount of time spent in updating or resetting the model. In practice the maximum repetition rate is determined by the limits of the computer one is using, and may be quite low, in the order of seconds. With low repetition rates the information conveyed by the predictor trace becomes more out of date as the cycle proceeds, and the predictive display itself acts as a sampled data system. Low repetition rates may also cause display flicker and visual fatigue problems for the operator as well as control difficulties. In general the required repetition rate increases as system response becomes more rapid.

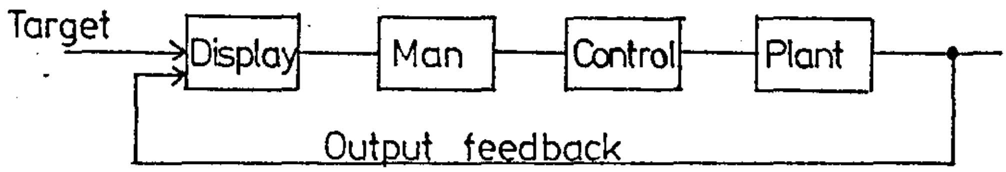
The frequency at which the predictor model is updated with fresh information is often identical to the repetition rate (being faster would be useless), in which case its effects are synonymous. When updating frequency is lower than the repetition rate the first prediction after updating will be the most accurate and each successive prediction will decrease in accuracy until the model is again updated. One solution to this dilemma is to update the predictor model artificially by extrapolating past sampled outputs of the plant over the updating period; another would be to let the predictor sample its own predictions and so update itself. For most applications repetition rate and frequency of updating are the same and, as they are predetermined by the computer system, are of rather academic interest. In any case their effect is likely to be slight. McCoy and Frost (1966) report that reducing the updating frequency of their predictor from continuous updating down to once every 50 seconds apparently made no difference to performance. The only practical significance is that prediction model inaccuracies can sometimes be offset by a high frequency of updating.

3.6 Modes of Control

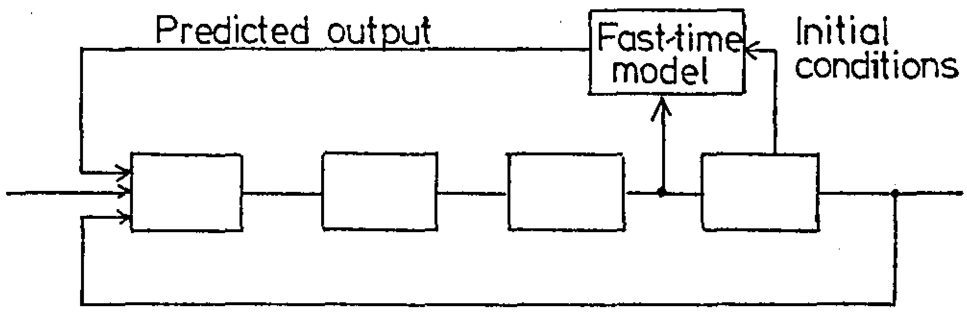
Two principal modes of control may be distinguished, depending on the philosophy behind the predictive display in use and the application for which it has been designed. These are on-line control and off-line control, of which category exploratory control and supervisory control (monitoring) are special cases. Figure 2 illustrates the main differences.

In on-line control the input to the prediction model is identical to the control input to the plant itself, so the operator sees a predicted path based on the assumption that he does not alter his control input. Any control change is immediately reflected on the predictive display. This mode of control is particularly suited to situations where an 'ideal' path or trajectory can be formulated,

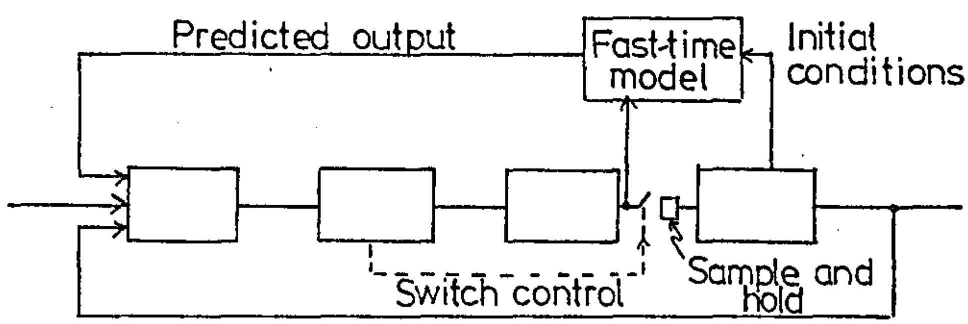
(a) Unaided Control



(b) On-line Prediction



(c) Exploratory Prediction



(d) Supervisory Prediction

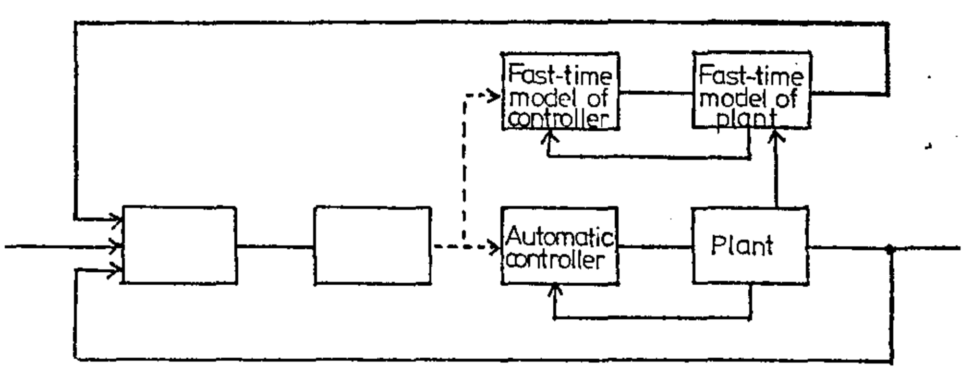


Figure 2: Possible modes of predictive control (after Warner, 1969).

e.g. aircraft landing. The pilot, via a continuous series of trial-and-error control actions in fast-time with real-time effects, is able to reduce the difference between actual and desired trajectories until his plane is on the runway glidepath.

In <u>off-line control</u> a hypothetical input is fed to the prediction model based on the assumption that the operator's control action will change during the predicted interval. The hypothetical input may take the form of sampled present control inputs or a complex pre-programmed sequence of control actions yielding a display of several different responses, the so-called 'multiple path prediction'. <u>Exploratory control</u> is a special case of off-line control. It differs from on-line control in that the operator's control actions are not input to the plant until he decides that the results of his choice of action, as reflected on the predictor display, constitute the optimum solution. In effect his control is directly coupled to the predictor display but only indirectly coupled to the plant, via an appropriate switch or sample and hold circuit. The selected control action may be the operator's most recent manipulation, or one that has previously been placed in 'storage' (Kelley, 1968). A variation of this technique is the case where the operator adjusts a hypothetical control program, building up a sequence of control actions, and only then does he command the actual controller to assume the form (in real-time, naturally) of the hypothetical program. Kelley, Beggs and Prosin (1973) have termed this flexible approach 'automanual control'.

It is evident that all forms of off-line control presuppose the luxury of sufficient time to explore the potential effects of alternative control actions. If appreciable searching is required before the best performance is reached, on-line control may be inadvisable as it will lead to substantially higher fuel consumption (McCoy and Frost 1966). Warner (1969) has shown that exploratory control is marginally better than on-line control (though not statistically different), so long as the required decision times are not short. Where control decisions are required immediately, however, on-line control is generally to be preferred. The additional control errors and use of fuel and resources attributable to an on-line mode of control are usually negligible when compared to the consequences of a long decision time.

The fourth mode of control, supervisory control, can be thought of as a further special case of off-line control and refers to those situations where the primary mode of control is automatic. The human functions in a system monitor capacity and may override the automatic system in cases of emergency, system failure, or for maintenance. Technically speaking, the entire control system is on-line, while the automatic and manual components are on-line and off-line respectively. The prediction model in this case also contains a fast-time model of the automatic controller. Two variations of supervisory control are possible, differing in the degree of 'pureness' of the off-line component In cases where the automatic control system is malfunctioning and manual back-up is essential, the operator may have little or not time to explore the utility of various control inputs. In this extreme he will be functioning in an on-line mode. In cases where automatic control malfunction occur but time is non-critical, or where the automatic system is functioning correctly but unanticipated events demand manual override, then the operator may be functioning predominantly in an off-line mode.

In the experimental Chapters of the present thesis, two control modes have been adopted corresponding to the distinction between discrete and continuous systems. Chapter 3, using Laios' (1975) Predictive Computer Display, is an example of exploratory prediction. The operator can try out various schedules off-line until he is required to implement one of them. In Chapters 5, 6 and 7 on-line prediction is employed to assist in the control of continuous chemical process plants. The time scale of the processes described in these later chapters is much shorter than for the scheduling applications of Chapters 3 and 4, so precluding the testing of alternative control strategies off-line. It should be noted, however, that the experimental tasks reported in this thesis have a substantial monitoring component.

4. REVIEW

Surveying the introductory Chapters, the following points have become apparent. In industrial process control, as with other control and decision-making applications, the human operator's ability to predict the consequences of his actions and to anticipate events is the underlying basis of skilled, 'open-loop' control. However, human predictive abilities are far from perfect, and whilst they may be adequate for the majority of everyday tasks, they cannot cope with abstract displays of complex processes without a lengthy training period or some form of artificial assistance. Predictive display systems are one of the most promising ventures into control and decision aiding, since they are directly geared to make up for what the human lacks.

Experimental evidence, largely from simulation studies of military and vehicular systems, suggests that substantial reductions in training times and a significant improvement in control and decision-making performance can be expected when predictive displays are employed. However, in spite of the dramatic improvements to be had, the natural conservatism of display designers has led to few real-world applications. In addition, the small amount of research that has been carried out into the design of predictive displays has been inadequate and the findings conflicting. There exists a pressing need to carry out an inter-related series of experiments with the object of evaluating the predictive display concept across a wide range of discrete and continuous systems, and to establish design guidelines for predictive display parameters in specific settings.

It is likely that knowledge of how people make use of predictive displays will shed some light on the way in which humans themselves structure information and anticipate consequences. A need exists from the theoretical point of view to unify the experimental findings from predictive display research within a common theoretical framework. Rejecting traditional control theory and decision theory models of the human operator as inadequate, the internal model approach seems to be the most widely applicable and promising modelling technique.

This thesis represents an initial attempt to fill some of the above gaps in our knowledge of predictive display systems.

PART II

EXPERIMENTAL PROGRAMME

INTRODUCTION TO EXPERIMENTAL PROGRAMME

1. AIMS

A review of the literature pertaining to predictive displays has shown that despite its potential, and despite the fact that the predictor instrument has been around since 1958:

1) there has been to date no comprehensive series of inter-related, multivariable experiments to determine the factors affecting predictive display performance, with the object of deriving optimum display configurations for specific situations.
2) not nearly enough studies have been conducted to evaluate the predictor technique across a variety of discrete and continuous tasks.

In addition, it has become evident that few operational applications of the predictor technique have been documented. The present thesis thus set out as an attempt to remedy these points, in the specific context of industrial process control, and in two parallel areas of application: scheduling tasks and continuous control tasks. The above broad aims were narrowed down to the following specific questions:

1) In a discrete laboratory scheduling task, how do variations in prediction span and the level of input uncertainty affect performance using a predictive computer display?
2) Are these results borne out using test data from an actual job-shop scheduling environment? (The author was not able to gain access to the job-shop itself.)

3) What are the potential benefits of introducing predictive displays in the control of a laboratory simulated continuous chemical plant? How sophisticated need the prediction model be? Do any benefits of training with a predictive display transfer to subsequent unaided control?

4) If the predictive approach should prove viable, how do variations in the display parameters and task characteristics affect performance? Specifically, using a laboratory dual-meter monitoring and control task what are the effects of adjustments in prediction span, variations in the level of input uncertainty, system gain, and the fidelity of the prediction model?

5) Are the potential benefits of the predictive approach borne out in an operational setting, namely in the control of a part-simulated, semi-batch chemical reactor with real plant and experienced operators, using a multipen predictive recorder?

6) What are the implications of the experimental studies for the design of predictive displays in discrete and continuous tasks, and for the ways in which process operators themselves control such tasks and structure information?

2. DESIGN AND ANALYSIS

The designs used in this experimental part have, with the exception of Chapter 4, been variations on the theme of multi-factor designs having repeated measures on one or more of the factors (Winer, 1971; Chapter 7). Kirk (1968) refers to this class of design as 'split-plot factorials'. These designs have the advantages of giving a tighter degree of experimental control than is possible with factorial designs (due to each subject acting as his own control), whilst at the same time allowing a realistic economy in the use of subjects. Potential disadvantages are that the confounding of the subjects factor with one or more of the experimental variables makes for a more complex analysis. Care is also needed to ensure that within-subjects phenomena such as sequence, practice and fatigue effects are minimised by counter-balancing/randomising the presentation order, thorough training, adequate rest periods, and in general by ensuring that the prerequisites of good experimental practice are met. Poulton (1974) has criticised all repeated measures designs for their proneness to asymmetrical transfer and range effects. However, it is not always possible to carry out independent group designs as Poulton advocates. The split-plot factorial is seen as a compromise solution. Given that its application can be justified and its limitations realised, this class of design has much to recommend it.

It should be noted that the many assumptions underlying the analysis of variance (ANOVA) approach were not tested; indeed the small sample sizes used in the present experiments precluded any such meaningful testing. (For a discussion of the principal assumptions underlying the ANOVA technique, reference may be made to Kirk, 1968, page 60.) For this reason the results of the statistical analyses should be regarded as indicative rather than conclusive. There is, however, much evidence that the F-distribution is quite robust with regard to minor violations of its assumptions (Cochran, 1947). As an additional safeguard, negatively biased or 'conservative' F ratios were used where possible. This approach adjusts the degrees of freedom in the F-ratio to present a test biassed against acceptance of the null hypothesis. The procedure (after Greenhouse and Geisser, 1959) is detailed in Winer (1971).

With the exception of the results from Chapter 4, the experimental data were analysed using the Biomedical Computer Programs package (Dixon, 1971). The variance analysis program BMD NO8V was employed specifying the nesting relationship for the design in question. All experimental effects were assumed to be fixed, whilst the subject factor was assumed to be random.

CHAPTER 3

AN EXPERIMENT TO EVALUATE THE EFFECT OF VARYING TASK CHARACTERISTICS AND PREDICTIVE DISPLAY PARAMETERS IN A SIMULATED SCHEDULING TASK.

1. OBJECT OF THE EXPERIMENT

This chapter continues the work of Laios (1975) by investigating the introduction of a wide range of input uncertainty and prediction span values, and their interaction, on a predictive scheduling aid designed specifically to cope with uncertain environments.

Several interesting questions follow on from Laios' work:

1) How are decision aids which have been designed specifically for uncertain environments affected by wide variations in the level of input uncertainty?
2) Since an adequate decision horizon appears to be vital for good performance, what effect does manipulating the prediction span (also termed the 'extrapolation interval' by some workers - the maximum distance to which planning may occur) have on performance?
   Is there, as some workers have found, an optimum distance for planning ahead in a given system?
3) How do non-specialist, non-mathematical users cope with predictive decision aids? (Laios used exclusively male post-graduate science and engineering students in his experiments).
4) Can performance studies with a predictive aid shed some light on an operator's strategies when carrying out a scheduling task?

2. METHOD

The laboratory simulated scheduling task used in this experiment was essentially that of Laios (1975), which in turn had been developed as a scaled-down version of Ketteringham and O'Brien's (1974) soaking pit scheduling problem. Complexity and learning times were set at a level that could be managed by naive student subjects. Certain important changes, notably display improvements, special purpose ergonomically designed keyboard and modified software, were made to

Laios' original simulation for the purposes of the present study. In all the task was considered to be a suitable vehicle on which to investigate the effect of internal and external parameter changes on discrete decision aids.

The following sections describe the problem environment, the simulation, experimental design, procedure and mode of data collection used in this study.

2.1 Description of Steel Manufacturing Process

Ketteringham et al.'s (1970) description of a typical steel plant producing 1.2 million tons of steel per year will be used to illustrate the main features of steel manufacture before and after the soaking pit area:

"Steel is made to a particular quantity in cast sizes of around 140 tons. The hot metal is then poured into moulds to make ingots of about 5 tons each. The steel is then allowed to cool, the length of the cooling time required before stripping being dependent on the steel quality. When the ingots have cooled sufficiently, they are stripped from the moulds and sent to the soaking pits for reheating. (This is the part of the process simulated.) The 20 pits vary in size and other characteristics but their common purpose is to heat the ingots to a temperature suitable for rolling. The length of time the ingots take to reach rolling termperature is dependent on a number of parameters, the most important of which is the length of time the ingot took to cool down. Other factors involved are the size and efficiency of the pit and the quality and amount of heating gas available. Clearly it is desirable that the ingots should spend the minimum amount of time in the pit provided that they are at the required rolling temperature when they are drawn out.

The scheduler's main criterion is to maintain a constant flow of hot ingots to the rolling mill, the operation of which is very expensive, around £600 per hour. The decisions taken by the scheduler are mainly concerned with the allocation of a cast or part of a cast to a soaking pit. Each pit has a different performance and each cast requires a different duration of soaking. Therefore with the time lags involved the scheduler is faced with the difficulty of predicting the long-term effects of his decisions. In the present situation he uses his experience and a few crude rules to satisfy short-term, cost-based objectives. He is incapable of calculating the long-term effects of his decisions which could adversely affect future mill operations".

2.2 Scheduling Problem

The experimental subjects' task was to schedule the utilisation of soaking pits simulated by computer. There were four soaking pits - denoted A, B, C and D - two with soaking times of 10 minutes (simulated time), and two with soaking times of 15 and 20 minutes respectively. The unloading of an ingot from each of these pits took a further three minutes before the pit could be used again. The simulation was comparable to 'real-life' in that pits would vary in their efficiency and hence in their soaking times, which would depend amongst other factors on the time since maintenance. Similarly, time would have to be allowed for ingots to be unloaded by crane before the next ingot could be loaded. However, in an actual scheduling task the operator would be required to control the output of 20-30 pits as opposed to the four pits of the simulated problem. Furthermore, the actual controller would have alternatives to loading the ingots into pits, for example, he could send ingots to a cold store.

As the controller of the soaking pits, the subjects' task was to assign ingots to pits so that a constant interval of five minutes, plus or minus one minute, was achieved between the times when ingots were ready for rolling (Figure 3). A variation of plus or minus one minute was considered reasonable. At the same time system constraints had to be satisfied by allocating ingots to pits as soon as possible, and by allowing three minutes after an ingot was ready before reloading that pit.

In principle the simulation was comparable with an actual process control task. However, it must be realised that in practice such a task would only constitute a part of the whole process control job. For example, an actual process controller might be required to carry out several other activities as part of his job, including communication with colleagues concerning aspects of the process, chart readings, report writing and so on. Long-term experience is known to be related to the social part of the job, as well as to the definition of the problem in unfamiliar situations.

2.3 The Display

The task was represented graphically on a PDP-12A computer screen (Figure 4). The left hand side of the display showed the four pits A, B, C and D. The simulation runs started at time 0, and as the simulation proceeded the time scale and pit contents moved to the left one minute at a time. Current time was represented by the left-most point on the time scale. For clarity, current time was also presented numerically at the bottom right of the screen.

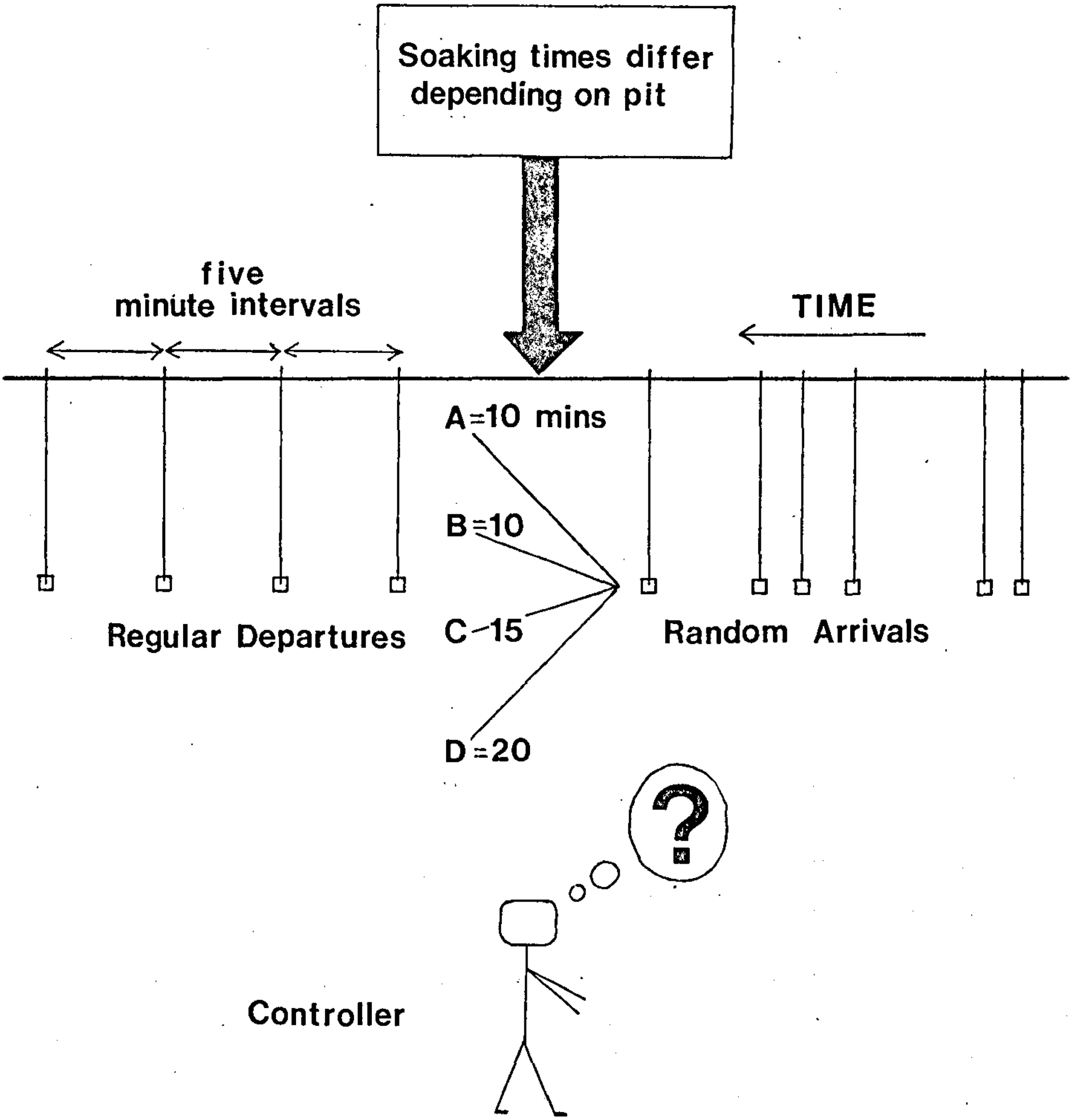


Figure 3: Summary of the Scheduling Task Problem

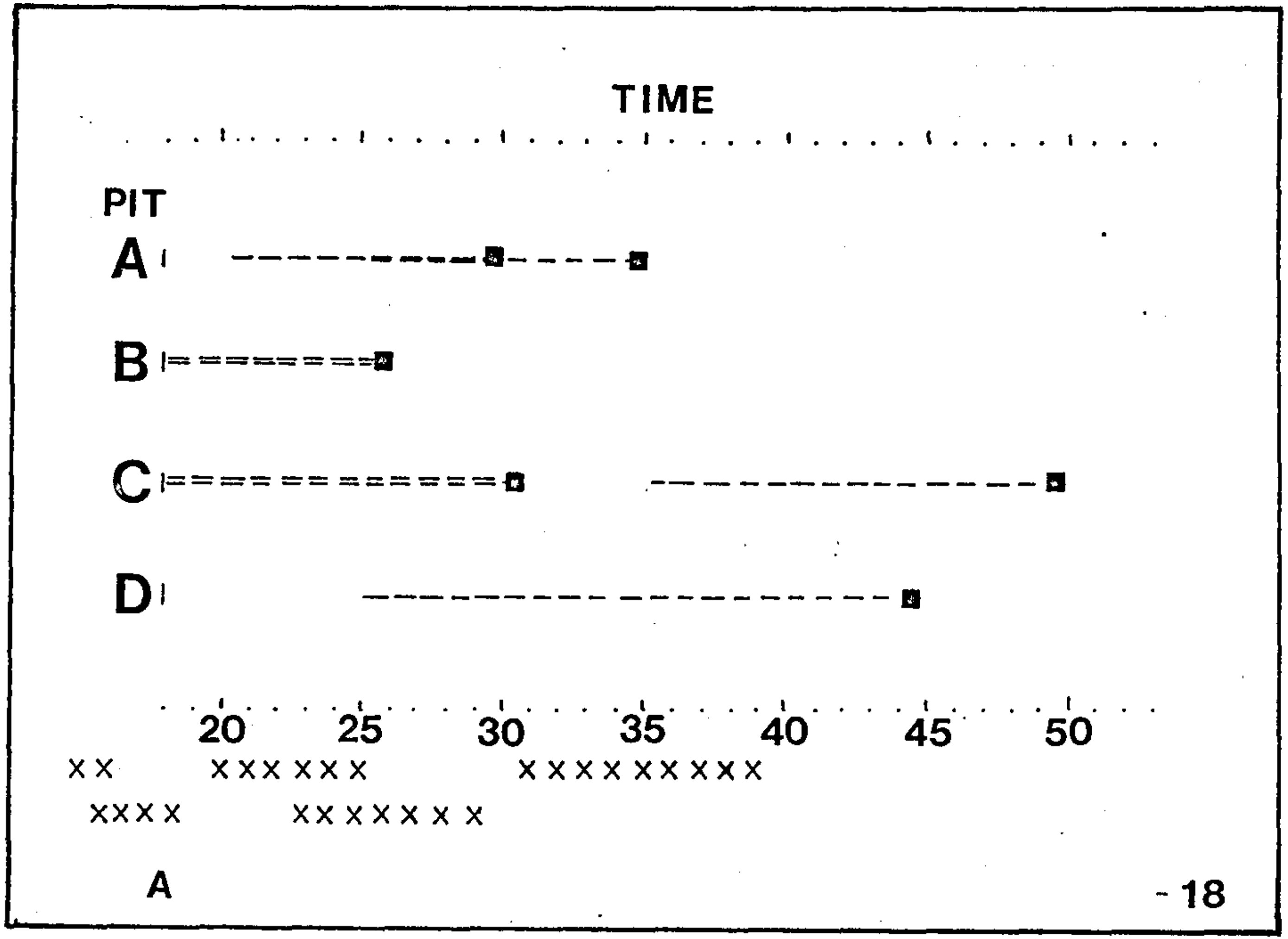


Figure 4: The Predictive Computer Display with full screen visible (35 minutes prediction span)

The display is shown at time 18 part-way through a trial. The four soaking pits A,B,C,D are denoted to the left of the display, with the time scale extending ahead to the right. The pit currently selected for test load purposes is indicated by a letter in the bottom left of the screen. Ingot arrivals are depicted by rows of crosses beneath the time scale. Single lines on the display represent test loads, of which there may be up to two in any pit (as in pit A). Double bars represent system loads, as in pits B and C. The squares at the end of each bar indicate the time at which the ingot

The time interval during which an ingot was due to arrive was shown directly beneath the lower time scale by a row of crosses. As in real-life, the exact arrival time of an ingot was indeterminate since it depended on a number of factors (delays, breakdowns etc.). In the experiment the uncertainty associated with ingot arrival times was represented by the width of the line of crosses along the time axis. An ingot would arrive at a time corresponding to one of the crosses in a given row, and although there was a greater tendency for the ingot to arrive at the centre of a given interval, it could arrive at any point. The time at which the ingot arrived during the displayed interval was determined by drawing randomly from a rectangular distribution, having location parameters equal to the actual ingot arrival times and scale parameters linearly decreasing as the actual arrival time approached (Laios, 1975). Subjects were therefore required to estimate on which cross an ingot would arrive, and base their schedules accordingly.

As the displayed arrival times (the rows of crosses beneath the bottom time scale) passed through current time and became historical information, they remained on the screen but to the left of the time scale until the entire set of arrivals information was updated. Updating occurred at times 10, 20, 30 and 40. At these times all historical arrivals information was deleted from the screen, those arrival intervals already advancing along the time scale became narrower, i.e. more accurate, and fresh arrival intervals were introduced at the extreme right of the time scale. Using the advance arrivals information subjects were able to 'game play' with the computer model and through the test load facility to plan their schedule of soaking pit utilisation ahead of current time. This process is described in detail in the next section.

2.4 The Controls

Interactive communication with the display was achieved by depressing push-buttons on a special purpose keyboard (Figure 5). Auditory and visual feedback to the subjects was provided through "click" push-buttons, which also illuminated when depressed. The keyboard comprised four 'TEST LOAD' buttons, corresponding to each of the four pits; ten 'AT TIME' buttons numbered 0 to 9; and four 'LOAD' buttons, again corresponding to each of the four pits. The 'TEST LOAD' and 'LOAD' buttons were colour coded orange and red respectively, and were also distinguishable by their positions. In addition there were four green coloured buttons marked 'ENTER', 'CLEAR', 'CANCEL' and 'CLEAR ALL'. The functional purpose of these buttons is best explained by means of examples.

1) If the subject anticipated an ingot arrival at time 20, he could test load (say) pit A at time 20 by pressing:

TEST LOAD	AT TIME ...		
$\boxed{A}$	$\boxed{2}$	$\boxed{0}$	$\boxed{\text{ENTER}}$

As the subject keyed in his choice, the program echoed his input in the extreme bottom left of the display. On pressing 'ENTER' a single bar appeared on the screen in pit A, starting at time 20 and extending 10 minutes ahead. This can be seen to correspond with the time that pit would require to heat an ingot. The simulation contained a heating model of the pits, so that if pit B had instead been test loaded the bar would also have extended 10 minutes ahead, but for pits C and D it would have extended by 15 and 20 minutes

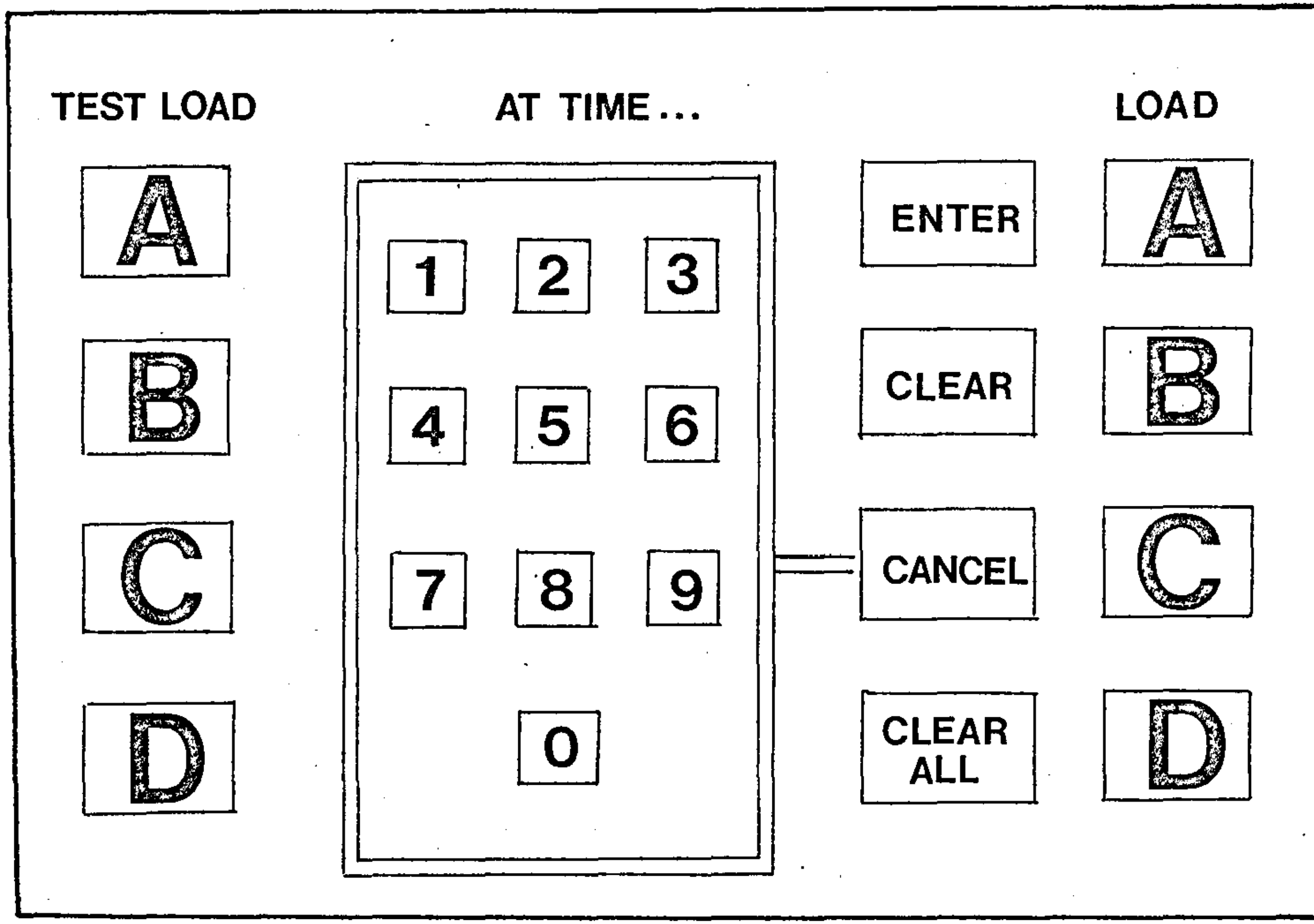


Figure 5: Special Purpose Keyboard

TEST LOAD keys are grouped together on the left of the keyboard. Used in conjunction with the AT TIME.. and ENTER keys, they permit test loading of any pit at a chosen time. Used in conjunction with the CLEAR key, they permit clearing of the test loads in specified pits. The CANCEL key allows a mis-keyed AT TIME.. number to be deleted. System LOAD keys are grouped on the right of the keyboard.

respectively. A square at the end of each bar indicated when the pit would be ready to empty, though the three minute unloading period was not included.

Constraints to the test load facility were that no more than two test loads could be entered into a given pit (Figure 4 shows a situation where the maximum of two test loads has been entered into pit A at times 20 and 25). Beyond this number the computer displayed an error message. In addition, test loads could not be made at times less than the current time, beyond 35 minutes ahead of current time, or into a pit for times when the pit was already occupied with a syste: load. In each case the computer displayed an appropriate error message, and the subject was required to depress any 'TEST LOAD' butto: to return to the scheduling display.

2) To clear pit A of test loads (single bars), the subject pressed:

TEST LOAD

$\boxed{\text{A}}$ $\boxed{\text{CLEAR}}$

He was then free to try out new entries in pit A.

3) To cancel a wrongly keyed number, the subject pressed:

$\boxed{\text{CANCEL}}$

then continued keying in the correct number. For clarity the 'AT TIME' keys were grouped together and linked by lines to the 'CANCEL' button (Figure 5).

4) To clear all test loads from the screen, the subject pressed:

CLEAR ALL

- actual system loads, of course, remained on the screen.

5) When an ingot actually arrived, a bell rang and the message 'CAST ARRIVED' flashed on the display. At this point the operator was required to load the ingot into a free pit at current time, and depending on his calculated strategy could press:

LOAD

A

B

C

D

A double bar then appeared on the screen commencing at current time to indicate that the chosen pit had been loaded. The internal heating model again ensured that the length of the double bar corresponded to the soaking time for that particular pit. The decision actually to load a pit was irrevocable, as such system loads could not then be deleted. Should the controller have attempted to load a pit which was either loaded or in the process of unloading, he would have been given an appropriate error message by the computer. If his chosen pit was already loaded, the controller was obliged to select an alternative pit. If it was in the process of unloading, he could either violate the three minute rule by depressing the 'LOAD' button a second time to load that pit, or he could select another pit.

2.5 Predictive and System Modes of Operation

It will be evident from the foregoing description that two distinct modes of operation were present: a predictive ('TEST LOAD') mode where the controller could try out various allocation schedules by 'game playing' with the computer model ahead of current time; and a system ('LOAD') mode to implement his final choice of schedule. By relating the squares at the end of each bar to the time scale, the controller could judge whether the main performance objective of a steady output flow of heated casts to the rolling mills was likely to be achieved with his present schedule.

2.6 Experimental Design

The two independent variables considered in this study were:

1) the level of uncertainty associated with ingot arrivals, manipulated by varying the width of the displayed arrival intervals;
2) the prediction span, manipulated by restricting the amount of display on the screen visible to the subject with a piece of cardboard.

In addition, subjects' academic background was considered as a subsidiary factor. A within-subjects design for uncertainty and a between-subjects design for prediction span was employed, as this arrangement permitted direct comparison with Laios' findings. Three levels of uncertainty corresponding to 1, 3 and 5 bits of uncertainty at an arbitrary time 25 minutes ahead were chosen. Three values of prediction span: full screen = 35 time units ahead visible, half screen = 20 units ahead, and quarter screen = 10 units ahead, were also considered. Six subjects were initially used for each of the three prediction span conditions. It was found however that individual variations in performance tended to obscure any experimental effects.

In order to obtain representative results an additional six subjects were run in each prediction span condition, making a total of thirty-six subjects in all. The design may be represented as follows:

	UNCERTAINTY		
	LOW	MEDIUM	HIGH
FULL SCREEN	G1	G1	G1
HALF SCREEN	G2	G2	G2
QUARTER SCREEN	G3	G3	G3

where G1, G2 and G3 represent independent groups of 12 subjects who underwent all three levels of uncertainty in a balanced order.

A wide range of science- and arts-based subjects were recruited so that the general applicability of the predictive aid could be tested. Subjects were trained under a level of uncertainty between the Low and Medium uncertainty conditions. It had been found that those subjects able to master the task at all reached satisfactory standard after three training trials. In addition, all subjects were trained with the full display visible. The advantage of using a common training regime for all three prediction span groups was that a separate test of location (Kruskal Wallis one way ANOVA by ranks) could be performed between the group training scores, a non-significant result indicating that subjects in the different groups were of comparable ability. This in fact was found to be the case. A second, more practical, reason was that an independent study on the validity of selection tests for process controllers was to be incorporated with the training programme (see Burton, 1976). A possible disadvantage of using a common training regime is that the Medium uncertainty/Full screen training combination may have resulted in a bias during the experimental trials. This was to some extent allowed for by using different representative ingot arrival patterns for the training trials, and for the Low uncertainty, Medium uncertainty and High uncertainty tasks. A successful training strategy could not therefore be transferred directly across to the experimental trials. This point will be discussed later, but at this stage it is sufficient to report that no consistent bias could be found.

2.7 Procedure

The experiment was conducted on a PDP-12 computer with associated VDU screen, teletype and magnetic tape units for program storage and data collection. On arrival, subjects were shown to the computer room and the nature of the task demonstrated. A detailed instructions sheet (see Appendix 3.1) was given to each subject and the contents repeated verbally to ensure their comprehension. Subjects were informed that 8 ingots would arrive in each training trial, 9 ingots in each experimental trial, and that the arrivals patterns in the experimental trials would be different. The broad heuristic that loading the pits with longer soaking times first frequently resulted in a more flexible schedule was also alluded to. No detailed feedback was given to the subjects at the end of each trial, though most subjects had a good idea of their performance level.

The three training trials were first run, followed by a 15 minute break for coffee and then the three experimental trials. Subjects comments were noted throughout and they were given ample opportunity to ask questions. An informal de-briefing session was held at the end of the experimental trials, at which subjects' general impressions were noted.

2.8 Subjects

Thirty-six male and female science-based and arts-based students were randomly selected from the student population by advertising and personal contact. Their specialities ranged from engineering and ergonomics at the one extreme, to librarianship and social psychology at the other. Subjects were randomly assigned between the experimental conditions, subject only to the need to distribute maths and non-maths backgrounds as evenly as possible. Subjects were paid £1 per hour for taking part.

2.9 Data Collection

An automatic data capture program logged every system input made by the subject onto magnetic tape. A Fortran analysis program was later run off-line and various performance measures calculated. These could then be correlated with subjects' opinions and subjective comments from the task. A typical result printout is given in Figure 6.

Three performance measures were computed for each trial.

1) Scheduling Errors - a relative measure of the average of the absolute differences between the actual ingot output intervals and the ideal target interval of 5 minutes, corrected by the total of the ideal differences from that target interval. This measure is taken from Laios (1975).

$$\text{Scheduling Errors (SE)} = \frac{\left\{ \sum_{i=1}^{(n-1)} \left| \left(t_{i+1} - t_i \right) - 5 \right| \right\} - I}{(n - 1)}$$

```
P.C.D. RESULTS PROGRAM V02 WRITTEN BY TONY OLIVER MARCH 1976

DATA FILE BLOCK NUMBER bbb
SUBJECT NAME nnnnn
MINS   SECS   INPUT

PIT LOADING SEQUENCE
       PIT    AT TIME
        LD        12
        LA        13
        LB        18
        LC        21
        LA        30
        LB        37
        LD        36
        LA        43

PIT EMPTYING SEQUENCE
       PIT    AT TIME
        LA        23
        LB        28
        LD        32
        LC        36
        LA        40
        LB        47
        LA        53
        LD        56

PERFORMANCE INDEX    0.00  Scheduling Errors SE

ACTIVITY ANALYSIS
PIT   NUMBER OF TIMES TEST LOADED
 A         9
 B         7
 C         7
 D         3
AVERAGE    6.50  Predictive Activity PA

DECISION HORIZONS
TIME INTERVAL     STAGES AHEAD PREDICTED
    1-10              1.90
   11-20              2.60
   21-30              3.60
   31-40              2.40
AVERAGE    2.72  Decision Horizon DH
```

Figure 6: Typical printout from results program

Scheduling Errors (SE) is a relative measure of the deviation from the target output interval of 5 minutes. Predictive Activity (PA) is the total number of test loads averaged over the four soaking pits. Decision Horizon (DH) is a measure of the number of test loads on the screen during each minute, obtained by calculating the mean value per minute during the time periods 1-10, 11-20, 21-30, 31-40 and then taking the average value.

where t_i is the unloading time of the ith ingot. 5 is the ideal target output interval, I is the sum of the ideal absolute differences from that interval, and n is the number of ingots per trial. Dividin by (n-1) gives a standardised relative measure of the number of scheduling errors. If the subject failed to allocate an ingot, n decreased accordingly.

2) Predictive Activity - a measure of the average use of the test load facility per pit.

$$\text{Predictive Activity (PA)} = \frac{TL_A + TL_B + TL_C + TL_D}{4}$$

where TL_A is the total number of test loads made in pit A, TL_B in pit B, etc., and the division by 4 gives an average measure of predictive activity per pit.

3) Decision Horizon - a measure of the number of test loads present on the screen at any one time.

$$\text{Decision Horizon (DH)} = \frac{\sum_{t=1}^{10} \frac{TL_t}{10} + \sum_{t=11}^{20} \frac{TL_t}{10} + \sum_{t=21}^{30} \frac{TL_t}{10} + \sum_{t=31}^{40} \frac{TL_t}{10}}{4}$$

where TL_t is the number of test loads on the screen at the same time during the t th minute. Since each pit will only accept a maximum of 2 test loads at one time, TL_t has a maximum value of 8 and a minimum value of 0. The four terms of the numerator represent average values of TL during the first ten minute period, second ten minute period and so forth. Dividing by 4 gives an average measure of the decision horizon per minute over the trial. At first sight one might expect a correlation to exist between predictive activity and decision horizon, since both are functions of the number of test loads, and since when PA = 0, DH must also = 0. However, consider the hypothetical situation where a subject makes a test load once every minute during the first 40 minutes of a trial, clearing his previous test load before making a new one. In this case, according to the formulae $PA = \frac{40}{4} = 10$, but $DH = \frac{1 + 1 + 1 + 1}{4} = 1$. In other words, although a substantial number of test loads has been made, the subject is effectively planning only one stage ahead throughout. It can therefore be seen that whereas predictive activity is a measure of the effort put into a search, decision horizon is a measure of how far ahead the subject is planning. Though the scheduling errors score was used as the main performance measure, an interesting indication of the 'search pattern' could be inferred from a plot of predictive activity vs. decision horizon. These are discussed separately under 4.4.

3. RESULTS AND STATISTICS

Graphs of scheduling errors, predictive activity and decision horizon averaged over each group of twelve subjects for the various levels of uncertainty and prediction span are given in Figures 7, 8 and 9. Raw performance scores can be found in Appendix 3.4-3.6.

The expected large variations between individuals scores, particularly in the case of the scheduling error data, has doubtless tended to obscure any differences due to experimental conditions. In spite of this the graphs show a clear effect of both uncertainty and prediction span on the performance measures. To test whether these differences were statistically significant, separate ANOVA's were performed on the scheduling errors, predictive activity and decision horizon data. The ANOVA model chosen was appropriate to the split-plot factorial design employed. As has previously been mentioned, tests on the group training scores had shown the three prediction span groups to be of comparable ability, and so the confounding of subjects with prediction span was in this case justified. Each ANOVA was followed by a Newman-Keuls multiple comparison test to determine which particular conditions had most contributed to any overall significant difference. Summary ANOVA tables are given in Tables 1, 2 and 3.

A separate analysis (not shown) had demonstrated presentation order effects to be statistically insignificant. As a check against range effects due to the within subjects component of the design, a separate analysis of the scores from the first experimental trials run by each subject was also performed. Mean performance scores were found to conform to the broad patterns shown in Figures 7, 8 and 9.

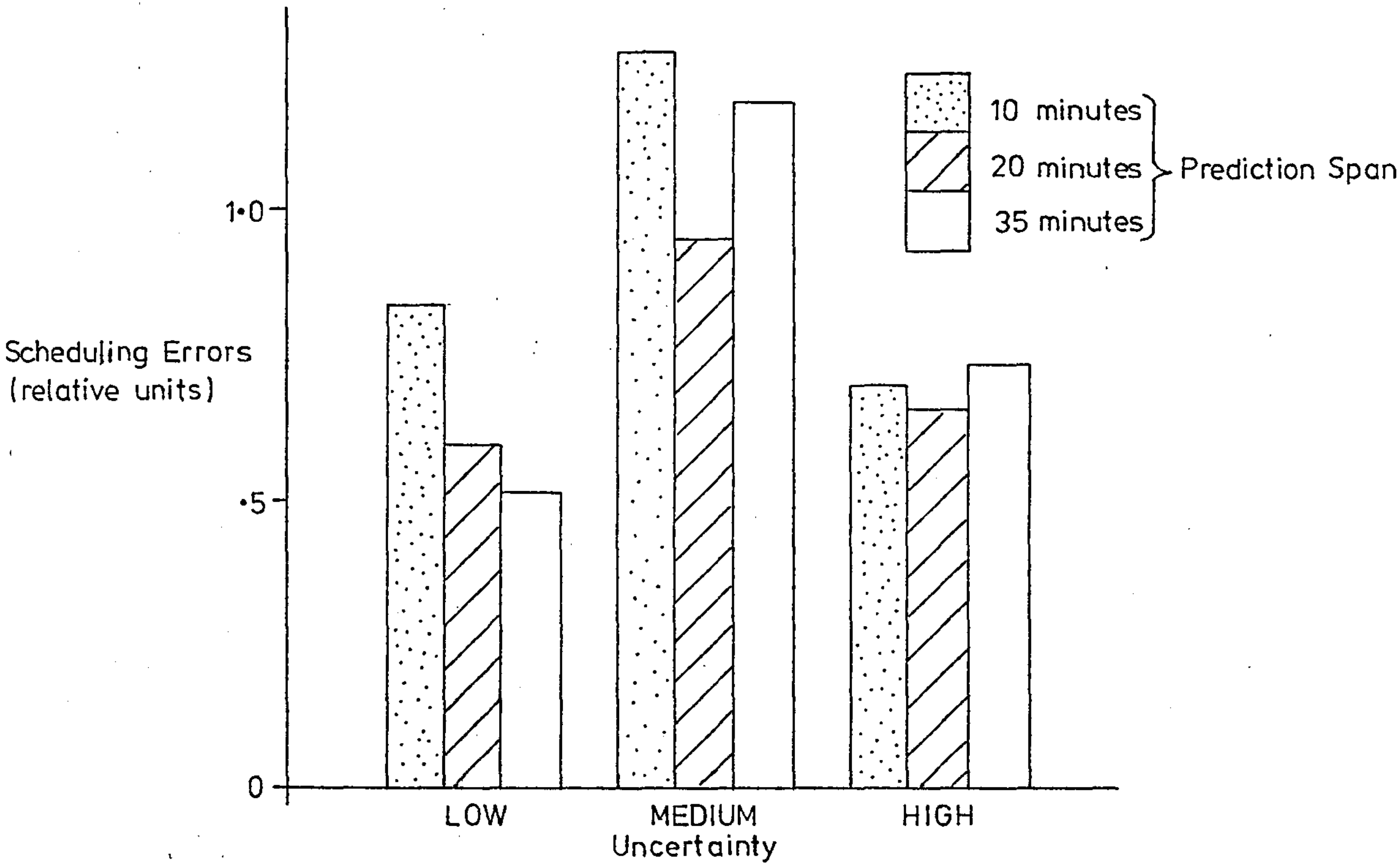


Figure 7: Scheduling Error Scores

TABLE 1: Summary ANOVA for Scheduling Error Scores

Source	Sum of Squares	df	Variance Estimate	'F'	Significance Level
Between Subjects					
Prediction Span	.75	2	.38	.81	- (df 2,33)
Subjects within groups	15.42	33	.47		
Within Subjects					
Uncertainty	5.27	2	2.63	5.69	5% (df 1,33)
Uncertainty x Prediction Span	.65	4	.16	.35	- (df 2,33)
Uncertainty x Subjects within groups	30.53	66	.46		

Conservative Test

Newman-Keuls Multiple Comparison Test

a) Uncertainty

	Low	High	Medium
Low	-	-	1%
High		-	1%
Medium			-

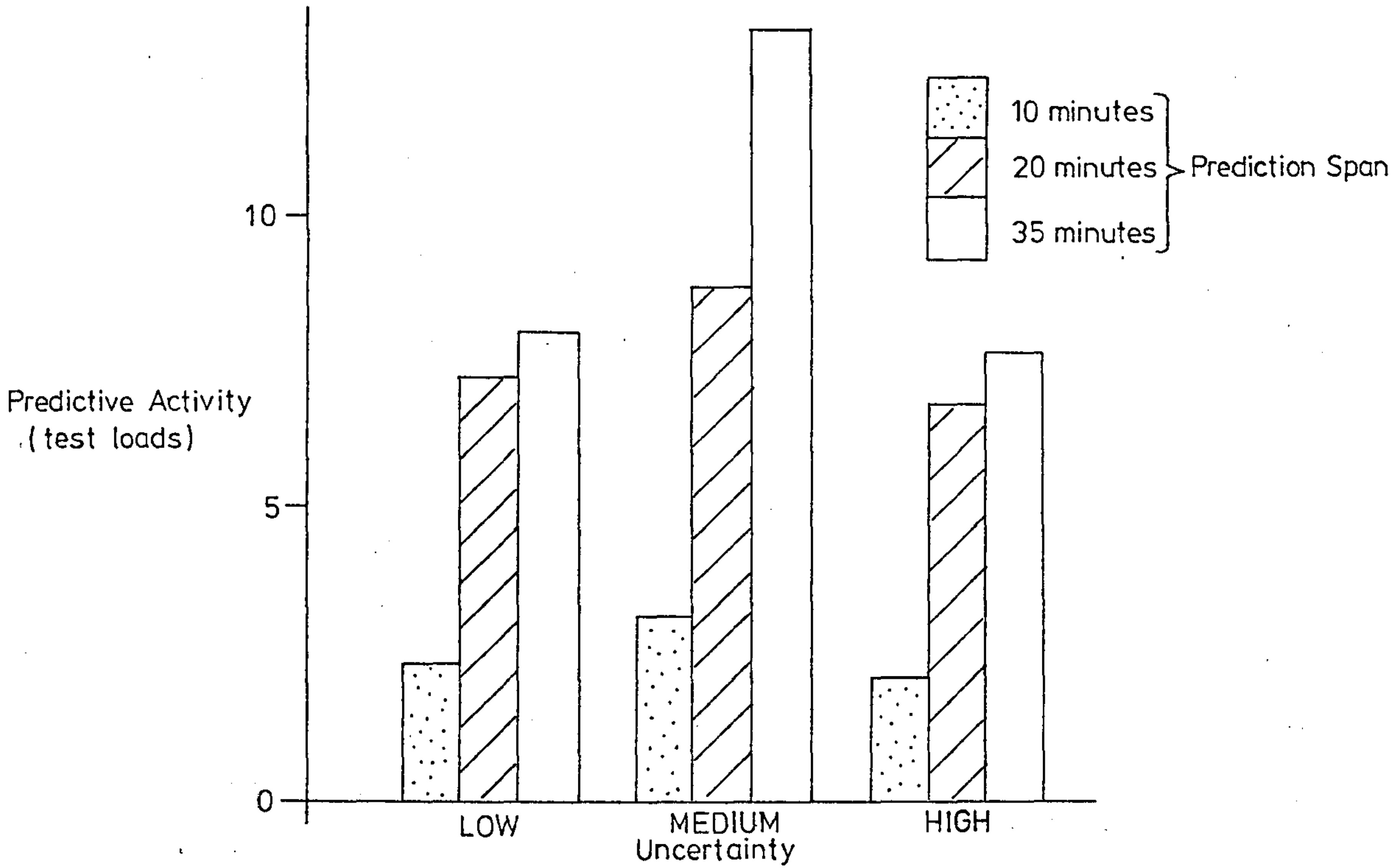


Figure 8: Predictive Activity Data

TABLE 2: Summary ANOVA for Predictive Activity Data

Source	Sum of Squares	df	Variance Estimate	'F'	Significance level
Between Subjects					
Prediction Span	967.77	2	483.88	20.77	0.1% (df 2,33)
Subjects within groups	768.8	33	23.3		
Within Subjects					
Uncertainty	174.45	2	87.23	20.96	0.1% (df 1,33)
Uncertainty x Prediction Span	89.84	4	22.46	5.4	1% (df 2,33)
Uncertainty x Subjects within groups	274.63	66	4.16		

Conservative Test

Newman-Keuls Multiple Comparison Test

a) Uncertainty		High	Low	Medium
	High	-	-	1%
	Low		-	1%
	Medium			-

b) Prediction Span		Quarter	Half	Full screen
	Quarter	-	1%	1%
	Half		-	-
	Full			-

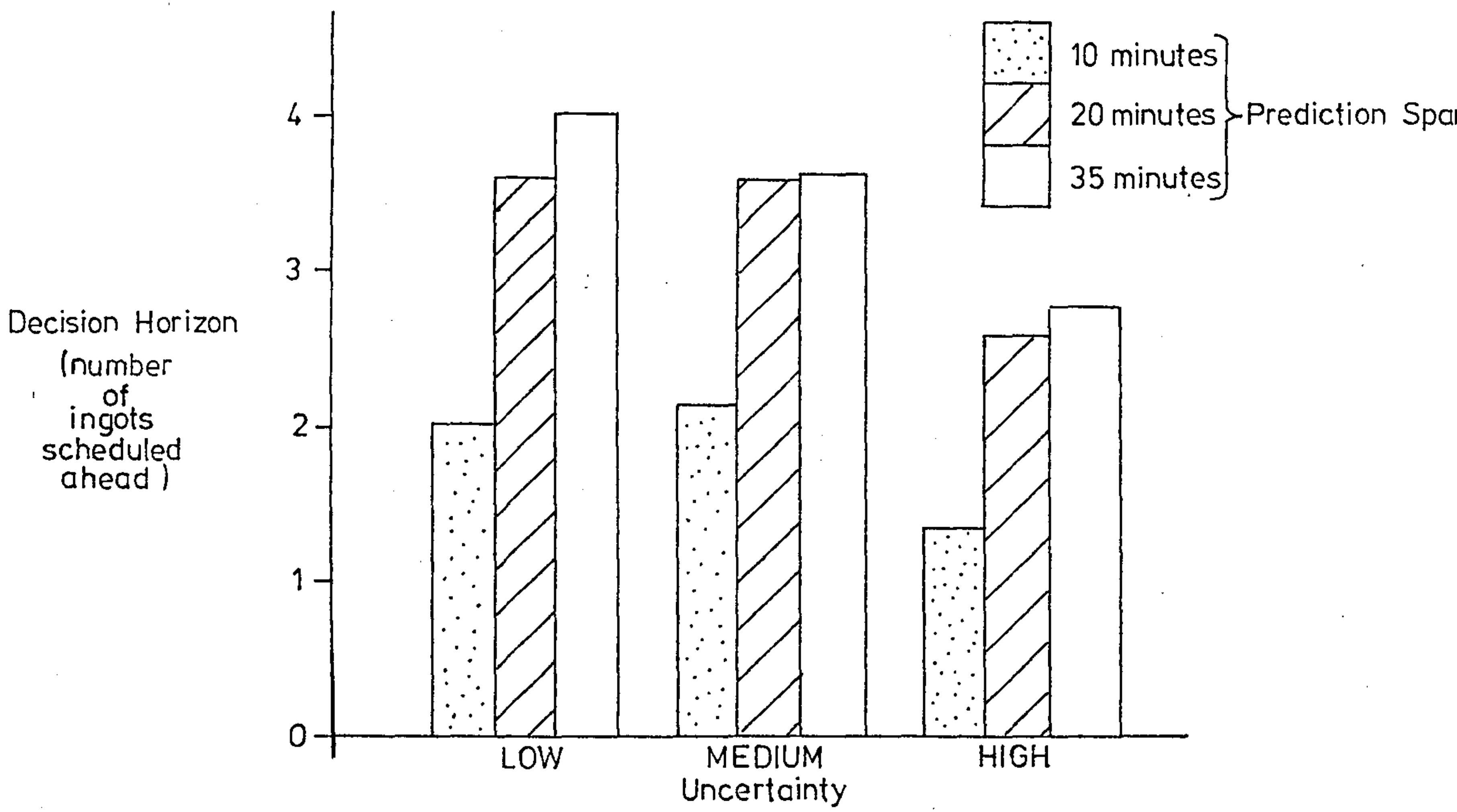


Figure 9: Decision Horizon Results

TABLE 3: Summary ANOVA for Decision Horizon Data

Source	Sum of Squares	df	Variance estimate	'F'	Significance level
Between Subjects					
Prediction Span	56.18	2	28.09	10.43	0.1% (df 2,33)
Subjects within groups	88.9	33	2.69		
Within Subjects					
Uncertainty	20.02	2	10.01	11.63	1% (df 1,33)
Uncertainty x Prediction Span	1.42	4	.36	.41	- (df 2,33)
Uncertainty x Subjects within groups	56.81	66	.86		

Conservative Test

Newman-Keuls Multiple Comparison Test

a) Uncertainty

	High	Medium	Low
High	-	1%	1%
Medium		-	-
Low			-

b) Prediction Span

	Quarter	Half	Full screen
Quarter	-	1%	1%
Half		-	-
Full			-

4. DISCUSSION

4.1 Scheduling Errors

Figure 7 indicates that decision performance was affected by uncertainty, and the ANOVA shows this effect to be significant at beyond the 5% level (Table 1). Closer inspection of the data reveals this result to be mainly due to the difference between uncertainties in the full screen condition. Following on from the ANOVA, the Newman-Keuls multiple comparison test (Table 1a) shows that the performance deterioration from Low to Medium uncertainty, and the subsequent improvement from Medium to High uncertainty, were both statistically significant at beyond the 1% level. However, no difference was found between the Low and High uncertainty conditions. It would be possible to invoke a Poulton (1974) -type range effect to explain the performance under Medium uncertainty - subjects could be transferring a successful (though subsequently inappropriate) schedule from the training trials to the experimental trial most resembling them. Subjects were, however, informed that the different trials contained different arrivals patterns, so this explanation seems implausible. Analysis of the first experimental trial scores also indicated a worsened performance under Medium uncertainty.

The scheduling error pattern of the subjects when using the predictive aid was adversely affected by uncertainty in much the same way as when no such predictive information is available. Laios (1976) found that unaided scheduling performance deteriorated when uncertainty in the form of inaccurate arrival estimates was first introduced, but reported a small improvement in performance at high uncertainty levels. Laios (1975) obtained similar results when a predictive facility was present. It seems that imprecise information degrades decision performance, but if such information is so inaccurate as to render it virtually useless, people resort to internal control models which entirely ignore the imprecise information. Under certain circumstances, e.g. when tasks are not exceedingly demanding or when an aid is available which facilitates the use of such models, performance may actually improve. The results from this study seem to affirm this view.

Close examination of peoples' activity patterns in this task has indicated that good performance was achieved by four subjects (S's 30, 32, 33 and 36) who hardly used the predictive display in its intended form except perhaps as a stateboard or memory aid. Rather, these subjects devised conditional rules and made their schedules accordingly, e.g. "if an ingot arrives soon, I could load it into C, but if it doesn't arrive for 5 minutes it would fit more neatly into B". In fact the task was of sufficient simplicity for adequate performance to be achieved by planning 2-3 stages ahead, and these few subjects could apparently manage the procedure in their heads. It would be highly interesting to speculate how the same subjects would have coped had the number of pits been greater (e.g. eight or ten) - in other words when the information needs to satisfy long term objectives exceeded subjects' short term memory limitations. Even taking aside the above four subjects, the performance of the remaining subjects under High uncertainty remains considerably better than their performance under Medium uncertaint It seems that it still was possible to use the same kind of conditional rules together with high utilisation of the predictive facility.

Returning to Table 1, the ANOVA further shows, somewhat unexpectedly, that the amount of screen visible to the subjects did not have a statistically significant effect on scheduling errors. Closer inspection of the data reveals that the only statistically significant effect (at beyond the 5% level) occurred in the Low uncertainty condition. Here, it seems that the further ahead subjects were able to plan, the lower their scheduling error score. Under such deterministic (certain) conditions the experimental controllers frequently scheduled all the visible ingot arrivals using the full extent of the display visible to them, while the corresponding decision horizon scores indicated them to be extrapolating in some cases beyond the visible screen.

Under Medium and High uncertainty conditions, however, the lowest mean scheduling error scores were achieved when only half the display was visible. This finding is in accordance with other investigators (Rouse, 1970; Bernotat, 1972) who have put forward the concept of an 'optimum prediction span'. It has already been noted that adequate performance was possible through planning 2-3 stages ahead, and this amount of planning was dictated approximately by the half screen (20 minutes) prediction span. It is tempting to suggest that a deterioration in performance may result equally from an inadequate or an excessive prediction span under uncertainty. This was reflected in the averaged values, though not statistically in the ANOVA - perhaps due to the high level of subject variability.

4.2 Predictive Activity

Figure 8 indicates that the amount of predictive activity is affected by uncertainty, and the ANOVA (Table 2) shows this effect to be highly significant at beyond the 0.1% level. The Newman-Keuls multiple comparison test (Table 2a) further shows that the increase in the use made of the test load facility from Low to Medium uncertainty, and the corresponding decrease from Medium to High uncertainty, were both statistically significant at beyond the 1% level. No significant difference was found between the Low and High conditions. It is interesting to note from Table 2 that an Uncertainty x Prediction Span interaction term is also present. Closer inspection of the data suggests that uncertainty exerted its largest effect on predictive activity in the Full screen condition, and that its effect diminished through the Half and Quarter screen conditions.

The following explanation of the effect uncertainty had on predictive activity is offered. When the information on which decisions are made is virtually certain, a schedule once made is unlikely to require updating. With the introduction of a Medium level of uncertainty, however, the allocation of ingots to pits is likely to require changing as updated information arrives - hence the increase in the use made of the test load facility. And when uncertainty is increased to such a level that rational choices cannot be made on the basis of the displayed information, a different policy is adopted requiring less search. In the latter condition, the display is being used more as a memory aid or stateboard rather than a planning aid. Laios (1976) reported a significant increase in the amount of on-line activity with the introductio of uncertainty in unaided scheduling, but on-line activity was not affected by a further increase in uncertainty level. On the other hand, when Laios (1975) introduced a predictive facility to the proceedings, he found that uncertainty had no apparent effect on on-line activity. It is interesting to note that no monotonic function was found in the present study as uncertainty was increased, rather a general increase from Low to Medium uncertainty, followed by a subsequent decrease as High uncertainty was introduced. It is possible that a range effect (Poulton, 1974) could again be the cuprit here, subjects being lured into making more test loads by the recognition of a familiar number of crosses. The present author believes this explanation to be unlikely, however, as analysis of the first experimental trials revealed a similar relationship to that shown in Figure 8. The discrepancy with some of Laios' results may be partly explained in the measures used, since Laios' definition of on-line activity was not specifically related to the use made of the predictive facility but included the calling of other displays. It is interesting to note that the average effect of uncertainty on predictive activity would appear to be the mirror image of its effect on scheduling errors - a high usage of the test load facility under Medium uncertainty actually being accompanied by an increase in scheduling errors. This could well be an artefact of the averages, however, as Figure 11 makes clear.

Turning to the effect of manipulating the amount of display visible on predictive activity, the ANOVA (Table 2) shows this effect to be highly significant at beyond the 0.1% level. The Newman-Keuls test (Table 2b) further indicates a fall-off in the use made of the test load facility between the Half screen and Quarter screen conditions, statistically significant at beyond the 1% level, though no significant fall-off occurred between the Full and Half screen conditions. Apparently reducing the amount of display visible by one half has little effect on overt prediction activity, but beyond this point halving screen size yet again does produce a deleterious effect. There may be a critical point for such a reduction, between 20 and 10 minutes prediction span.

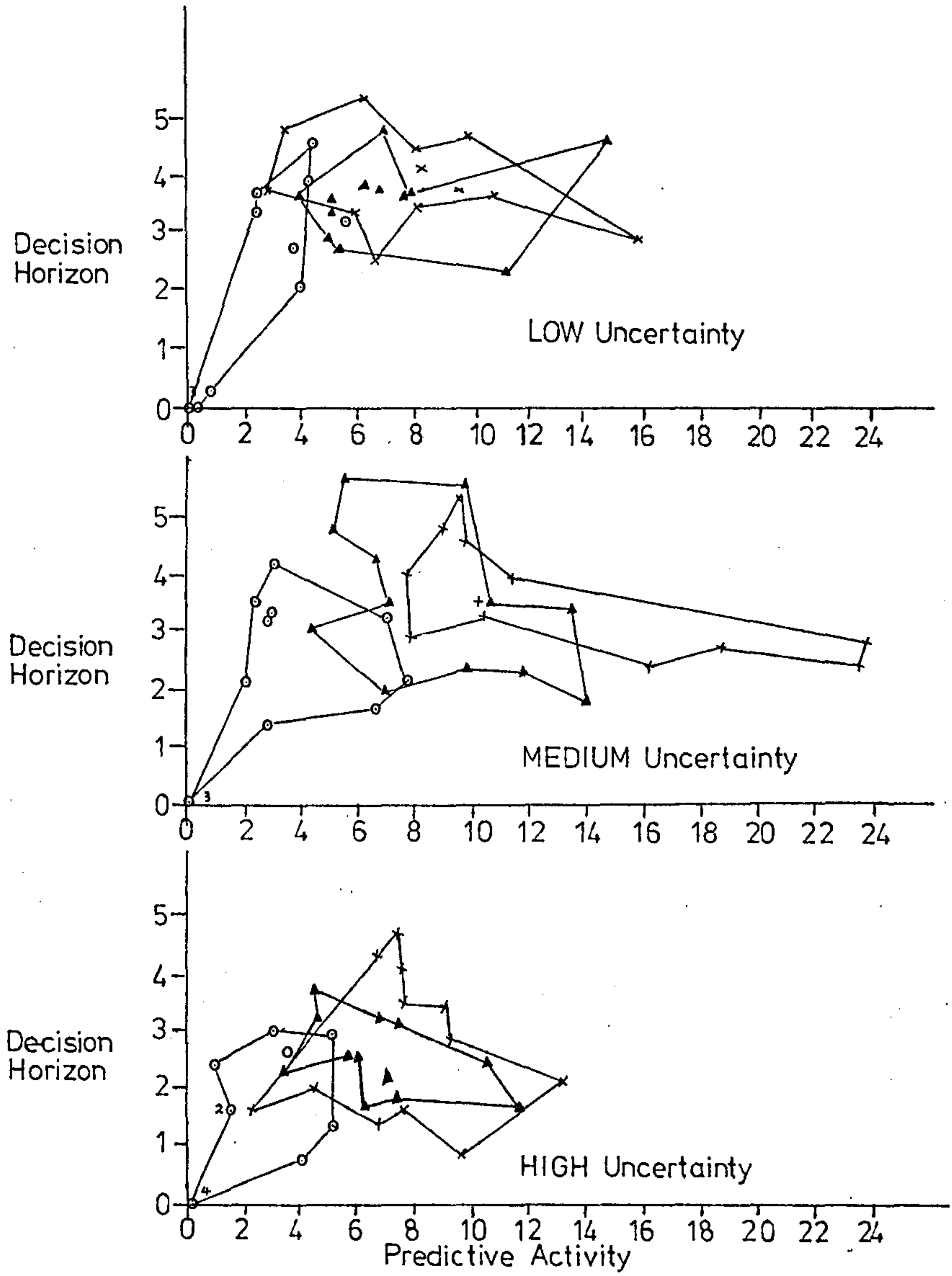
4.3 Decision Horizons

It is clear from Figure 9 that the decision horizon is adversely affected by uncertainty, and the ANOVA (Table 3) shows this effect to be significant at beyond the 1% level. It is evident that under High uncertainty the mean decision horizons are considerably shorter than the decision horizons under Low or Medium uncertainty, and the Newman-Keuls test (Table 3a) confirms that the fall-off is statistically significant at beyond the 1% level. It seems that in highly uncertain environments, people become more conservative in their willingness or ability to plan far ahead. Laios (1975) also found a consistent reduction in subjects' decision horizons under uncertainty.

As would perhaps be expected, prediction span has an effect on subjects' decision horizons. The ANOVA (Table 3) shows this effect to be statistically significant at beyond the 0.1% level. The multiple comparison test (Table 3b) gives a statistical difference significant at beyond the 1% level between the Quarter and Half screen conditions, but no difference between the Full and Half screen conditions. The critical point for reduction of decision horizons would appear to be between 20 and 10 minutes prediction span, a similar pattern to the predictive activity scores. In the Quarter screen condition, both predictive activity and decision horizon are significantly reduced. An analogy may be drawn with motorway driving - when visibility is reduced due to fog on the carriageways, planning ahead is restricted and speed is reduced.

4.4 Search Patterns

The combined effect of the experimental variables on predictive activity and decision horizon (the 'search pattern') is shown in Figure 10 overleaf. A characteristic pattern can be seen to emerge. As



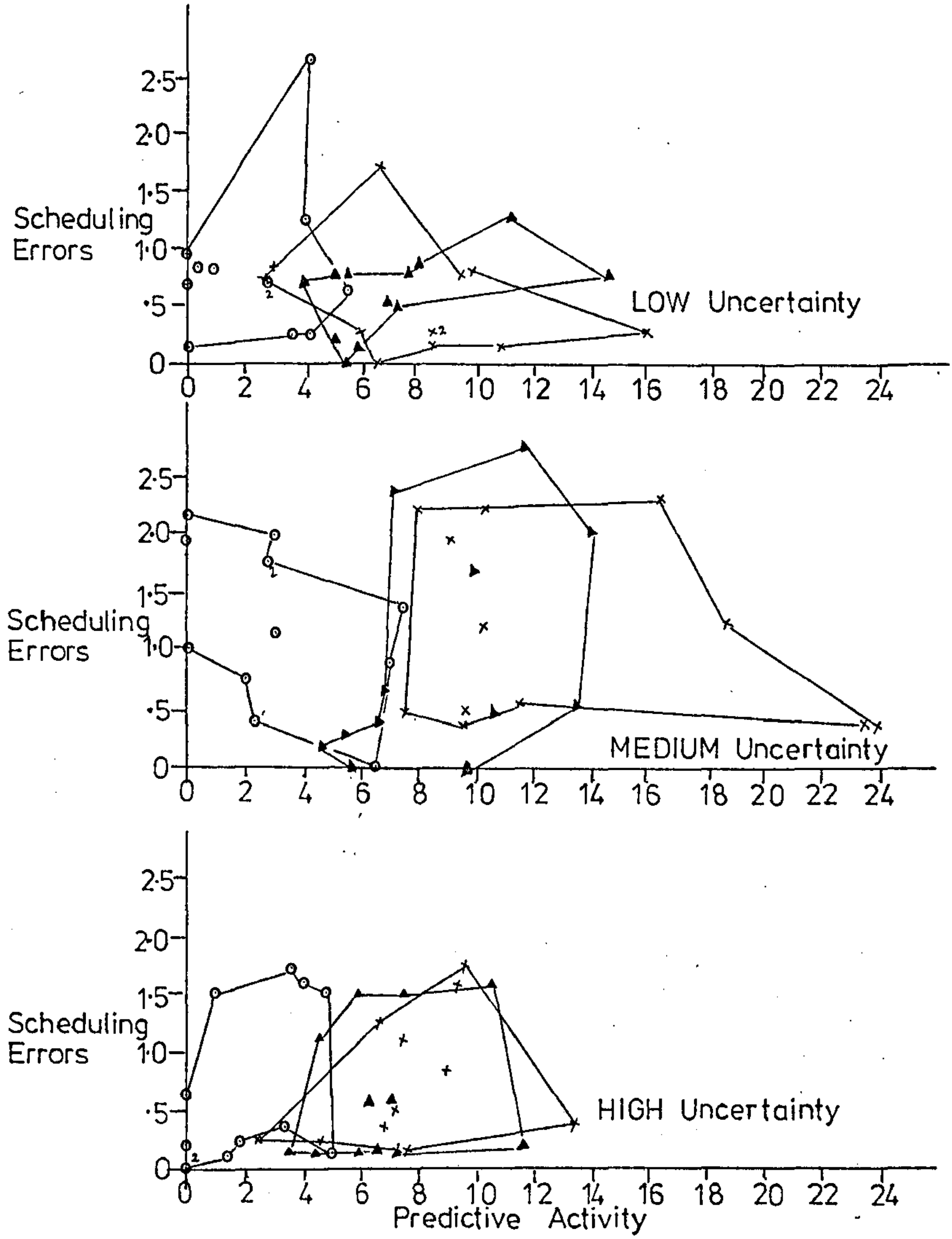
Key: Search pattern for 10 minutes prediction span
Search pattern for 20 minutes prediction span
Search pattern for 35 minutes prediction span

Figure 10: Search patterns (Decision Horizons vs. Predictive Activity) as a function of Uncertainty level and Prediction span.

uncertainty is initially increased from Low to Medium levels the search pattern broadens. In particular there is a notable increase in the amount of predictive activity for roughly the same decision horizon. With the introduction of a High level of uncertainty the search pattern can be seen to narrow, in terms of both the number of stages ahead which were explored and the amount of test loads used in exploring them. As far as prediction span is concerned, there appears to be little change in the search pattern as the span is reduced from Full to Half screen (except under Medium uncertainty where a slight reduction in predictive activity occurs). As prediction span is further reduced to Quarter screen size, however, a dramatic narrowing of the search pattern takes place. At an extreme of this trend several subjects ceased to use the predictive facility altogether. The combined effect on predictive activity and decision horizon thus revealed should help in clarifying the previous discussion.

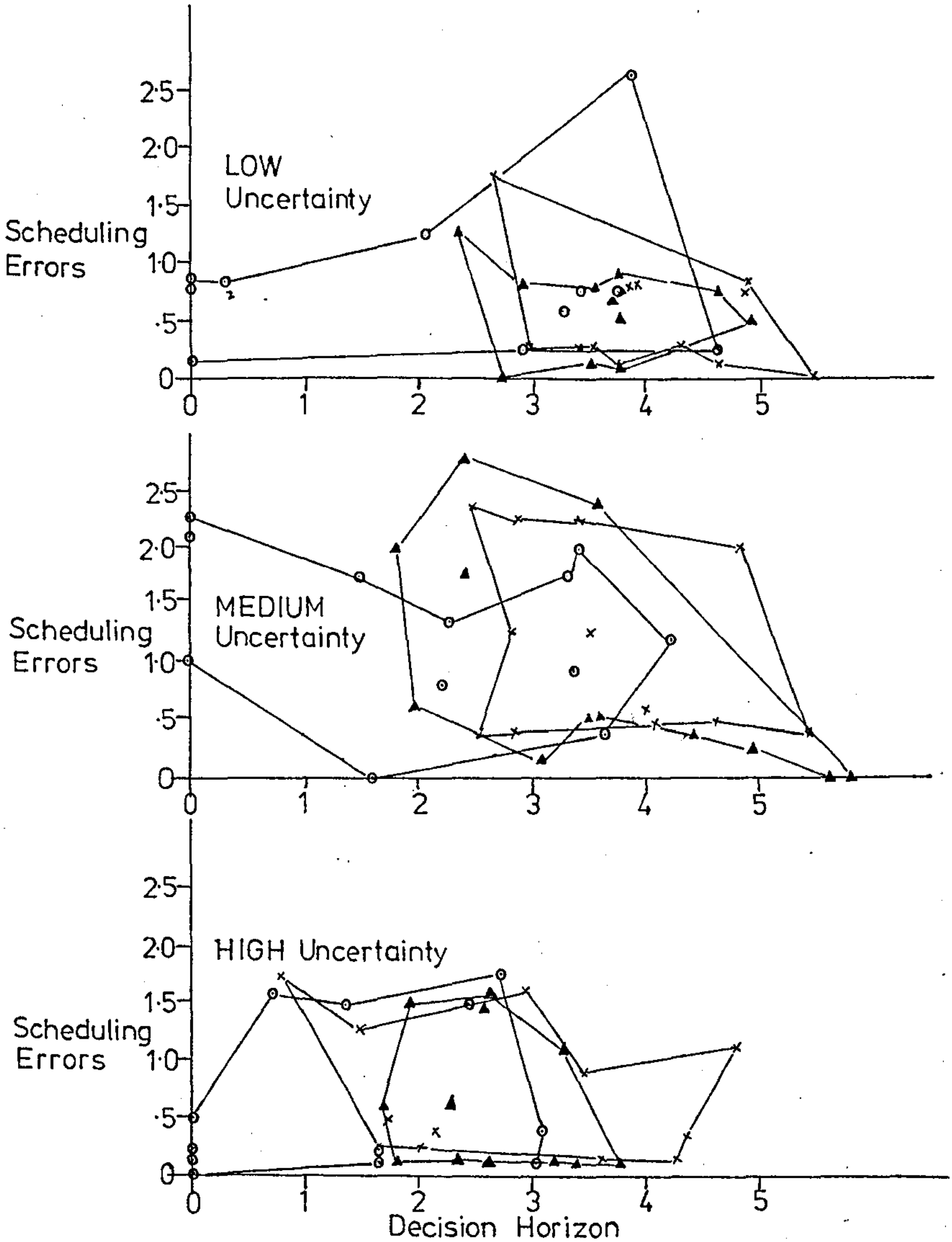
4.5 Relationship between scheduling performance and predictive activity, scheduling performance and decision horizon

Unfortunately, the results of the present investigation were not sufficient to examine the relationship between scheduling errors and predictive activity or decision horizon. As Figures 11 and 12 reveal, any relations are hidden because of the large individual differences between subjects and the different strategies used. There is little indication from Figure 11 that subjects necessarily performed better by making more use of the test load facility. Indeed according to the average scores an inverse relationship holds.



Key: Data for 10 minutes prediction span
Data for 20 minutes prediction span
Data for 35 minutes prediction span

Figure 11: Scheduling Errors vs. Predictive Activity as a function of Uncertainty level and Prediction span



Key: Data for 10 minutes prediction span
Data for 20 minutes prediction span
Data for 35 minutes prediction span

Figure 12: Scheduling Errors vs. Decision Horizon as a function of Uncertainty level and Prediction span

As far as decision horizons are concerned, it seems that the average decision horizon in all except one case exceeded 2 stages. (In the Quarter screen condition of Figure 9 the true decision horizon is actually understated since zero decision horizon was recorded for the four subjects who did not use the predictive facility and this has decreased the average.) With a decision horizon of more than 2 stages, an appreciable reduction in the number of scheduling errors should have been possible. However, no such reduction is evident from Figure 12. The failure to find any relationship may have been a by-product of the present experimental set-up. In retrospect, it might have been better to control directly the stages ahead the subjects could plan rather than indirectly by the amount of display visible.

4.6 Specialist vs. Non-specialist Users

A Mann-Whitney test was conducted between the scheduling error scores summed over the Low + Medium + High uncertainty conditions for specialist, science/maths-based subjects (n=24) vs. non-specialist, arts-based subjects (n=12). No significant difference was found.

It would seem that the organisational ability necessary for reasonable decision performance on this task is not a function of academic discipline, but of higher cognitive factors. Some people are inherently good schedulers, and some are not. It is worth noting that of the four subjects recruited but subsequently excluded from the analysis as not having mastered the task within the training time allowed, all were arts-oriented students. Similarly, an interest in games such as chess, which requires a considerable degree of forward planning, was expressed by several subjects achieving consistently low scheduling error scores. As learning proceeded there was also some evidence of the more able subjects violating secondary objectives (by delaying ingots before loading) in order to obtain smoother output flows. Laios (1975) first noticed this trading of objectives effect.

4.7 Subjective Comments

Comments elicited from students during and after the experiment show that there were really as many ways of tackling the problem as there were individuals. Some common threads did emerge, however, and are worth noting here.

Very few subjects noticed that the three training trials were identical - a different starting combination gave a fresh problem from the subjects' point of view. In the experimental trials subjects often reported that as the level of uncertainty increased, their amount of planning activity went down. Or as one subject put it: "In the Low uncertainty condition I could work out the intervals, with Medium uncertainty I could juggle the times but under High uncertainty I just had to wait and see".

Reported decision horizons were also longest in the Low uncertainty condition and fell-off with increasing uncertainty. In view of this it is hardly surprising that most students rated the Low uncertainty condition the easiest and High uncertainty the most difficult. A minority found the Medium condition the easiest, lending support to Poulton's middle-of-the-range phenomenon, with the Low condition boring and the High condition challenging. Yet another sub-group said they did best with a little (Low uncertainty) or a lot (High uncertainty) of displayed arrivals information, and found the Medium condition peculiarly confusing! This may have been due to the retention of an inappropriate scheduling acquired during the training trials.

On restricting the prediction span to Half screen size subjects reported using all the screen up to the card, and many extrapolated mentally a bit beyond. When only a quarter of the screen was visible, the card was generally held to be frustrating as "you had to do the sums in your head". Indeed, some subjects could see no point in using the display by this stage and did all calculations mentally, based on visible arrivals information and which pits were free. The latter subjects had presumably formed some internal model of the pit soaking times from experience and were able to use this in conjunction with pit status and arrivals information derived from the display.

As far as strategies were concerned, some students had a marked preference for one pit or another, others aimed to keep certain pits free, some claimed to have found a general purpose "right answer" (although none existed), others had no particular plan in mind but worked on the situation prevailing at the time. A few subjects even tried the heuristic rules hinted at in the instructions.

A surprising number found the scheduling problem enjoyable, if somewhat taxing. Many requested a pencil and paper as a memory aid. General complaints related to screen flicker which often resulted in eye strain, and the confusing effect produced when the ingot arrival times were updated at times 10, 20, 30 and 40. It is interesting to note that subjects varied widely in their ability to verbalise what their strategies had been in carrying out the problem. Some students were naturally reticent , others had to verbalise what they were doing "to help me think straight".

On the basis of this limited protocol evidence, subjects appeared to be setting up sub-goals (particular ingot arrivals, or groups of these) within the main task objectives of achieving a constant output flow, loading as soon as possible after an ingot arrival, and maintaining a three minute gap for pit emptyings. Much time was spent considering the effect of minor schedule adjustments on performance. The evidence here tends to support Bainbridge's (1974) loose, hierarchical, goal-directed model in decision-making tasks. (This question will be discussed further in Chapter 9.)

In summary, subjects' comments when coupled with the search patterns derived from the schedules gave a valuable insight into their thought processes during the scheduling task.

5. CONCLUSIONS

The main finding from this study is that decision aids which have been designed specifically to cope with uncertain environments are still adversely affected by the introduction of uncertainty. Scheduling performance was observed to breakdown as information uncertainty was initially introduced but subsequently to improve as the level of uncertainty was further increased. The first statement is self-explanatory. The second finding can be explained by the types of strategy adopted by the schedulers. There was frequently less overt search, the display being used as a memory rather than a planning aid. The use made of the predictive test load facility similarly rose and then fell away as uncertainty was initially introduced and then increased. A consistent reduction in the scheduler's decision horizon was also demonstrated when information uncertainty was increased beyond a critical point.

Prediction span was found to have an effect on scheduling performance only in deterministic conditions, where longer spans led to lower scheduling errors. This highlights the importance of advanced planning in scheduling. There was some suggestion of an optimum prediction span under uncertainty, but this was not supported by the statistical analysis. Reducing the prediction span beyond a critical point did, however, lead to a reduction in both decision horizons and the amount of predictive activity.

As far as academic background is concerned, non-specialist, non-mathematical users appeared to be as capable of using the predictive aid as specialist users with a mathematical background, having first got over the initial hurdle of understanding what the task entailed.

On the basis of subjective comments and search patterns produced from the schedules, some insight has been gained into operator strategies during scheduling tasks. In common with most other work in this area, a high level of individual variability was found, and this has made it difficult to draw many general conclusions. Plotting of individual search patterns revealed an interesting relationship between predictive activity and decision horizon. Limited verbal protocol evidence lends support to Bainbridge's concept of a loose, hierarchical goal-directed model in decision-making tasks.

CHAPTER 4

AN EXPERIMENT TO VALIDATE THE EFFECT OF VARYING PREDICTION SPAN IN A SCHEDULING TASK USING TEST DATA FROM AN OPERATIONAL JOB-SHOP.

1. OBJECT OF THE EXPERIMENT

The previous chapter has demonstrated the effect of uncertainty and prediction span on performance using a predictive scheduling aid designed for operation in certain environments. It remains to be seen how these important factors are dealt with in the real-world, or at least using test data from an actual industrial scheduling system. As the effect of inaccurate input information has to some extent already been explored in a real-life situation (Bibby, 1974; McEwing, 1977) it was decided to restrict the present experiment to the effect of variations in prediction span.

This chapter therefore employs the 'scheduler's abacus' coloured wooden-block scheduling system designed by Laios and Gibson (1976), with test data obtained from an operational job-shop, in an attempt to verify the effect of prediction span on scheduling performance under deterministic conditions. The author was, unfortunately, not able to gain research access to the job-shop itself. In the absence of direct access to an operational job-shop, the use of the 'scheduler's abacus' in conjunction with test data obtained from an operational job-shop was held to be an acceptable compromise for the purposes of the present study.

2. METHOD

The simulated representation of a job-shop scheduling environment used in this experiment was essentially the 'scheduler's abacus' improved machines-by-time configuration, described in detail elsewhere (Gibson and Laios, 1978). The problem was such that it could be managed by student subjects after a reasonable training period, since the abacus facilitated the perceptual representation and solution of an otherwise complex numerical task. Although the abacus is a manual system its results are relevant to both manual and man-computer systems.

The following sections describe the problem environment, the simulation, experimental design, procedure and mode of data collection used in this study.

2.1 Description of the Problem Environment

The task presented to the controller of a job-shop in an engineering works may be stated as follows. A large number of jobs, each requiring a given sequence of operations, compete for time on common machine facilities. Typical machine operations would be milling, boring, drilling, turning, etc. Each operation occupies a given machine for a specific time, since a machine can only process one job at a time. The controller's problem is to schedule the passage of the jobs through his machine shop, so that the jobs are finished in priority order by the required due dates, whilst at the same time minimising inefficient use of his workforce and machines. The latter objectives are frequently conflicting, so that without some form of control aid the controller has difficulty in working out the schedule best satisfying the numerous performance criteria. In addition, the combinatorial nature of the problem means that with J jobs and M machines there will be $(J!)^M$ possible schedules to consider. This means that computers alone also cannot handle problems of such magnitude. Interactive job-shop scheduling systems therefore take advantage of man's flexibility and pattern recognition abilities allied to the computer's capacity for rapid calculations and its built-in consistency.

The job-shop scheduling problem accounts for considerable wastage of resources and productive capacity, but is often unrealised as management attempts to bury the problem by providing excessive shop capacity, safety stocks or working on low profit margins. Pounds (1963) summarises industry's attitude: "There is no scheduling problem for them (factory schedulers) because the organisation which surrounds the schedulers reacts to protect them from strongly interdependent sequencing problems". Clearly any aid which enables the job-shop controller rapidly to interpret the current state of his machines and to test alternative schedules is likely to help improve productivity through a more even and efficient work flow.

2.2 The Simulation

The scheduler's abacus apparatus used for this experiment is illustrated in Figure 13 (top). The design of the abacus is documented fully elsewhere (Gibson and Laios, 1978) so a brief description will suffice here. Basically the scheduler's abacus comprised thirteen horizontal slotted channels mounted on a wooden board. Each channel represented a particular machine, with different types of machine being represented by different colour codes. Two yellow and two blue channels were available, in addition to one each of red, purple, mauve, plain, light blue, green, light green, black and lastly turquoise. Fifty-five representative jobs were available to be scheduled on the machines.

Each job consisted of between one and four machine operations, represented by coloured wooden blocks of different lengths. The length of a block was proportional to the time it would require on the machine. The different coloured blocks of different lengths representing one job through different machine operations all had the same job number written on their top right-hand corner. The component parts of a given job had to be processed in a specified order, e.g. a job might have to be formed before trimming off the rough edges; this sequence was shown by serial numbers 1, 2, 3 etc., on the top left-hand corner of the coloured blocks. A listing of the fifty-five jobs is given in Appendix 4.5.

Figure 13: The scheduler's abacus (above) at the start of a trial showing the jobs in progress, and (below) showing a completed schedule.

2.3 The Task

The subject's task was to schedule as many of the jobs as he could, making the fullest use of the machines, in as short a time as possible. This he did by selecting a number of jobs and fitting them onto the appropriate coloured channels on the board, arranging the job so there were no conflicts, e.g. a job could not be on two machines at the same time, and then adding, removing, replacing and rearranging jobs to get the best schedule. A completed schedule is shown in Figure 13 (bottom).

2.4 Experimental Design

A between subjects, independent groups design was used with 5 subjects randomly assigned per condition. The only experimental variable considered was prediction span, or the time interval through which the subject could plan his schedule. The latter was manipulated using a large piece of white card to restrict the abacus surface visible to the subject. Three levels of prediction span were considered: Full board (40 hours scheduled at a time), Half board (20 hours scheduled at a time) and Quarter board (10 hours scheduled at a time).

The first group of five subjects were asked to make their schedule using the full board. Five subjects were asked to schedule up to the first 20 hours, then the remainder of the board. Another five subjects scheduled the board 10 hours at a time. In the latter two conditions, when the subjects were satisfied with their part-schedule, their efforts were covered over with a sheet of transparent perspex, as a safeguard against later alteration, and the next section of the board was made available for scheduling.

2.5 Procedure

On arrival in the experimental room, subjects were presented with a standard set of instructions (see Appendix 4.1) describing the abacus and the scheduling problem. The instructions were repeated verbally to ensure comprehension. Subjects were then asked to make a practice schedule using jobs 1, 5, 10, 15, 20, 26 and 30 only, to check whether the instructions had been understood. The trial jobs were then replaced and each subject requested to make a full schedule from a selection of the 55 waiting jobs. Subjects were asked to continue scheduling until they were satisfied that their schedule was the best compromise between the performance criteria. Subjects were free throughout to ask questions or make comments. Subjects' opinions were noted, as was the time they took to complete their schedules.

If constraint violations were discovered in a subject's final schedule it was altered to show the actual pattern of work that would have resulted from the schedule, before performance measures were computed.

2.6 Subjects

All subjects were undergraduate students with some mathematical background. They were paid £1 for taking part.

2.7 Data Collection

In addition to the scheduling time in minutes, percentage machine utilisation and percentage of jobs unfinished scores were computed from the finished schedules. The following formulae were used:

$$\text{Percentage machine utilisation (MU)} = \frac{(494 - \text{Machine idle time}) \times 100}{494}$$

where 494 represented the total machine capacity in hours. (It was not possible to achieve 100% machine utilisation. Given the jobs available, the maximum possible machine utilisation was around 80%.)

$$\text{Percentage of jobs unfinished (JU)} = \frac{\text{Jobs not completed} \times 100\%}{55},$$

where 'jobs not completed' includes those jobs not attempted, or only part completed, at the end of the two day (40 hour) scheduling period.

3. RESULTS AND STATISTICS

The performance measures used in the evaluation of the completed schedules are shown in Figure 14 averaged for the three experimental conditions. The detailed scores for machine utilisation, jobs unfinishe(and scheduling times can be found in Appendix 4.4.

Separate single factor ANOVA's appropriate to the simple randomised design employed were performed on the percentage machine utilisation score percentage jobs unfinished data, and scheduling times in minutes. Summary ANOVA tables are given in Table 4. If statistically significant the ANOVA was followed by the appropriate Newman-Keuls multiple comparison test.

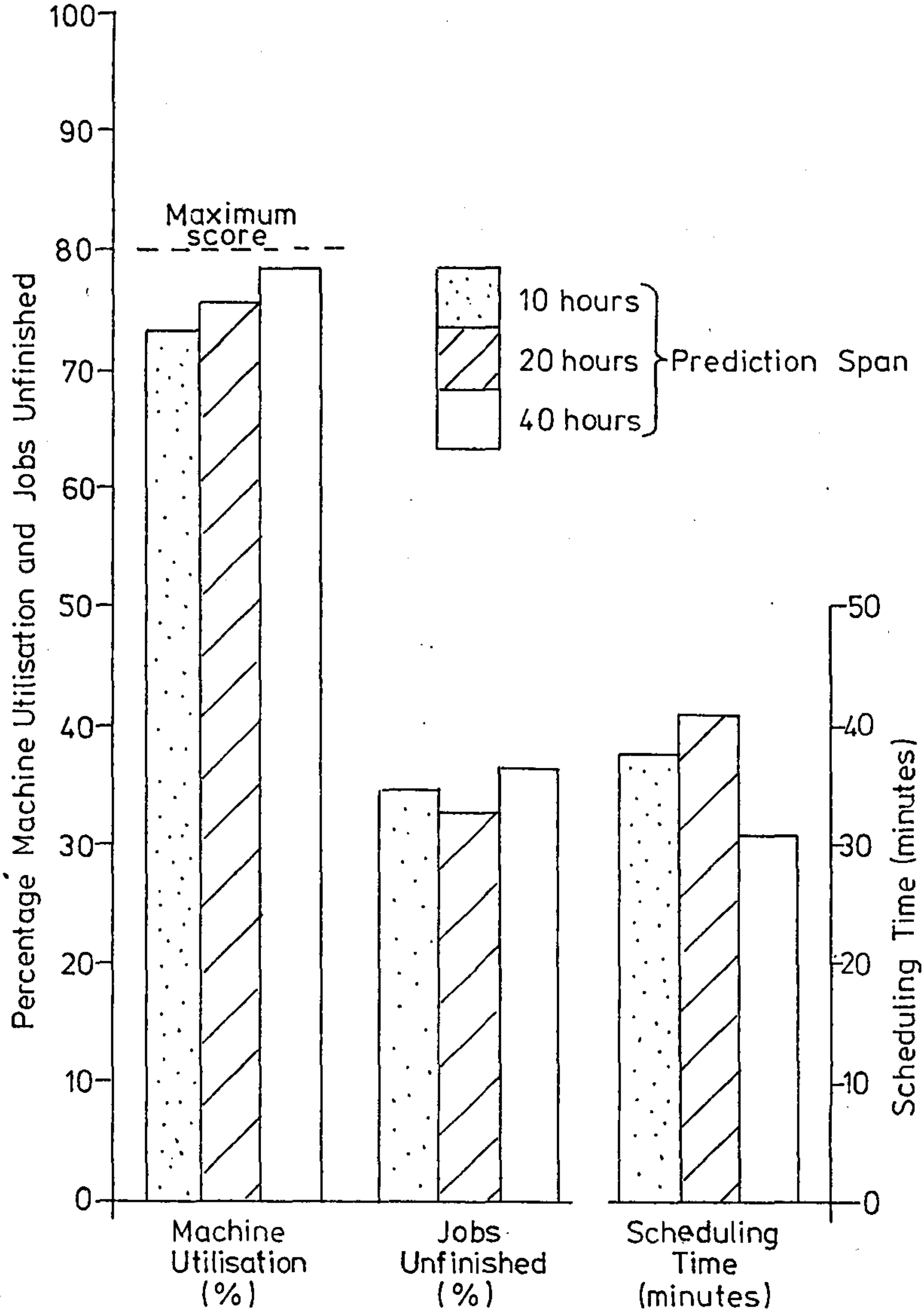


Figure 14: Machine utilisation, jobs unfinished and scheduling error scores for different prediction spans.

TABLE 4: Summary ANOVA tables for Machine Utilisation, Jobs Unfinished and Scheduling times

a) Machine Utilisation (MU)

Source		SS	df	Variance Estimate	'F'	Significance level
MU	(Between S's)	58.905	2	29.453	11.083	1% (df 2,12)
Error	(Within S's)	31.888	12	2.657		

Newman-Keuls Multiple Comparison Test

	Quarter	Half	Full board
Quarter	-	-	1%
Half		-	5%
Full			-

b) Jobs Unfinished (JU)

Source		SS	df	Variance Estimate	'F'	Significance level
JU	(Between S's)	26.689	2	13.345	0.275	- (df 2,12)
Error	(Within S's)	582.388	12	48.532		

c) Scheduling times (ST)

Source		SS	df	Variance Estimate	'F'	Significance level
ST	(Between S's)	260.8	2	130.4	1.043	- (df 2,12)
Error	(Within S's)	1500.8	12	125.067		

4. DISCUSSION

4.1 Performance Measures

The mean data scores plotted in Figure 14 indicate that prediction span has an effect on the performance measures, and the ANOVA (Table 4a) shows that in the case of the machine utilisation scores its effect was statistically significant at beyond the 1% level. This would appear to confirm the findings of the previous chapter regarding scheduling performance under deterministic (certain) conditions - scheduling performance deteriorates as prediction span is reduced. The Newman-Keuls multiple comparison test (Table 4a) further indicates that there is a critical point for such a reduction. Full board machine utilisation scores were found to be reliably different from the Half board (5% significance) and Quarter board (1% significance) conditions, though no difference could be detected between the Quarter and Half board conditions. Adequate machine utilisation in this study could thus be achieved planning at least 20 hours ahead. Though the improvement in machine utilisation is only in the order of a few percentage points, such an improvement can result in considerable cost savings when compounded over an entire factory.

The ANOVA (Table 4 b,c) further indicates that no significant effect of prediction span on jobs unfinished or scheduling times could be found. Figure 14 suggests that the lowest jobs unfinished scores were achieved in the Half board condition, where scheduling times were also longest, but it is likely that individual strategies have served to mask this effect. Achieving maximum machine utilisation together with minimum jobs unfinished were to some extent conflicting criteria, and this is reflected in closer inspection of the performance measures (Appendix 4.4) by some evidence of a trade-off between them, notably in the Full board condition (Figure 15).

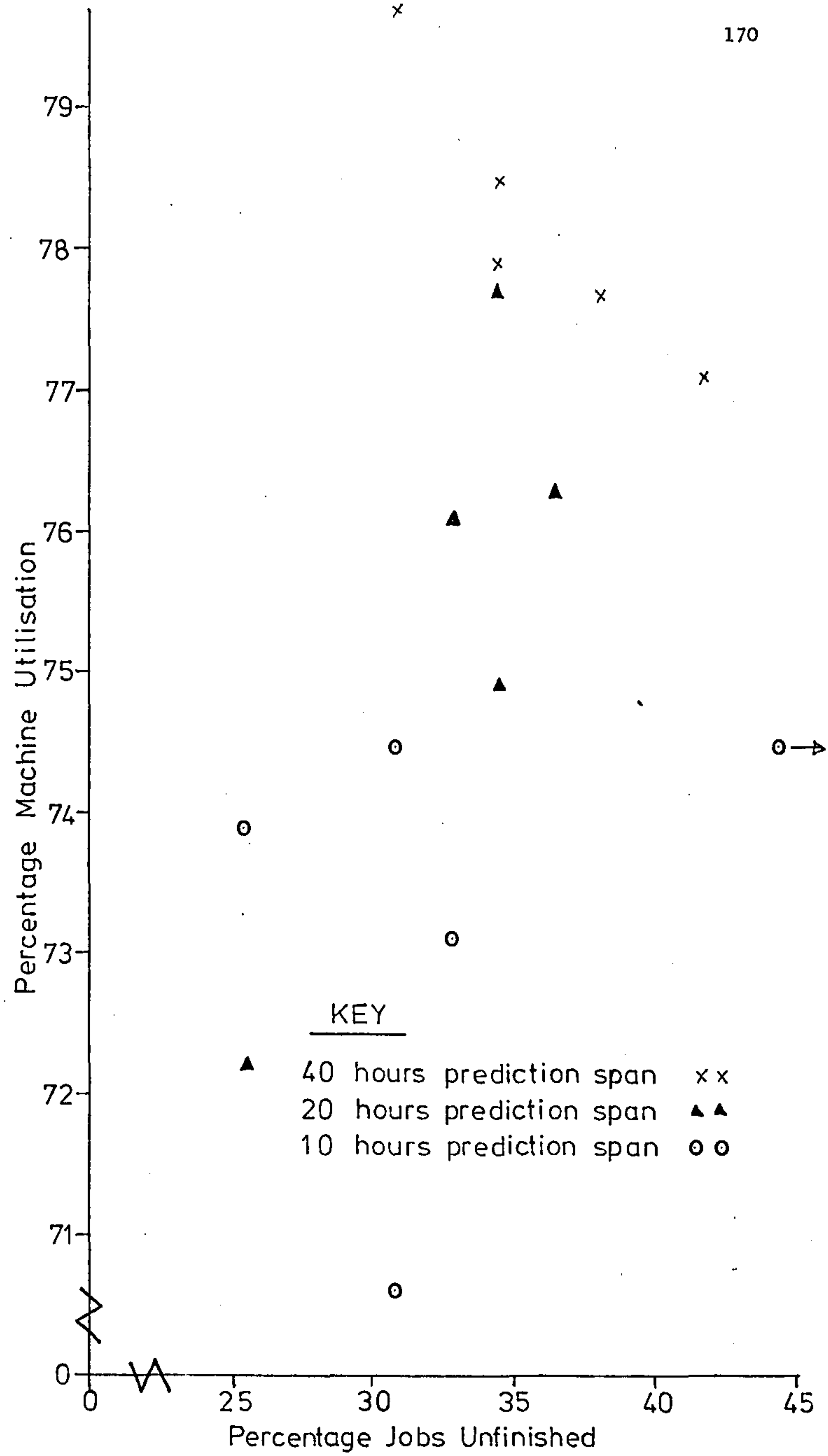


Figure 15: Machine utilisation vs. jobs unfinished scores for individual subjects

4.2 Subjective Comments

Remarks made by subjects during and at the end of the experimental trials served to reinforce the impressions gained from the objective performance measures. Subjects having the Full board available to them seemed able to maintain adequate decision horizons, and their high degree of forward planning led to superior machine utilisation scores. Those subjects restricted to scheduling half the board at a time commented on the desirability of being able to plan further ahead, or at least being able to alter one's past schedule to alleviate subsequent allocation problems. Only one subject judged the Half board condition to be useful in that "it got me into a mood of compactness right from the beginning".

Subjects restricted to scheduling the board 10 hours at a time tended to use up the shorter jobs first and leave the longer jobs unscheduled. In fact some particularly long jobs (Job 54 in Appendix A4.5 for example) appeared to be critical to good performance, but with only the Quarter board available considerable foresight or experience would have been needed to fit these in initially in preference to the shorter jobs. Subjects in the Quarter conditions universally commented on the need to be able to plan further ahead.

As in the previous chapter, some subjects found it helpful to verbalise their thought processes as they went along. On the basis of such limited protocol evidence, subjects appeared to be setting up sub-goals (particular jobs, or groups of jobs) within the main objective of maximising machine utilisation and minimising jobs unfinished. Much time was spent checking for conflicts between consecutive operations as well as in testing the consequences of slight adjustments within a main schedule. The evidence again tends to lend support to Bainbridge's (1974) concept of a loose, hierarchical goal-directed model in decision-making tasks. (This point will be discussed further in Chapter 9.)

4.3 Uncertain Environments

It would be interesting to see whether the notion of an optimum prediction span in uncertain conditions, as suggested in the last chapter and by other workers in this field, is borne out using the scheduler's abacus system. However, it is not a simple matter to incorporate uncertainty into a wooden block system, and the testing will come with its computer-based implementation. Field trials of such a system are currently in progress as part of a related study at Loughborough (Gibson, 1978). It has already been shown in another context (McEwing, 1977) that in real-life situations the uncertainty associated with information on which schedules are based results in a performance considerably worse than in a near-deterministic laboratory simulation. On the basis of previous work one might perhaps expect to find an optimum prediction span when the scheduler's abacus is taken into the field, as the real-world uncertainty associated with the information on which decisions are based makes it impossible to plan so far ahead.

5. CONCLUSION

The main conclusion from this chapter has been to confirm, using test data from an operational job-shop, the finding of the previous chapter concerning prediction spans. Under deterministic conditions, longer prediction spans lead to improved scheduling performance. A critical value of prediction span was found below which machine utilisation performance deteriorated, stressing the importance of adequate look-ahead or 'predictive' information in deterministic manual scheduling tasks.

No attempt was made to verify the effect of prediction span on scheduling performance under uncertainty. However, an experiment is currently in progress as part of a related project at Loughborough to implement a computer-based version of the scheduler's abacus and to test it in an operational setting.

In common with the previous chapter, limited protocol evidence from the present study seemed to favour a loose, goal-directed hierarchical model of decision-making skills.

CHAPTER 5

A PRELIMINARY INVESTIGATION OF PREDICTIVE TECHNIQUES IN CONTINUOUS CHEMICAL PROCESS CONTROL

1. OBJECT OF THE EXPERIMENT

The previous two chapters have been concerned with parameter variations in discrete aids for scheduling problems. The predictor display concept was, of course, originally developed for continuous control applications, though its application in continuous industrial environments has so far been limited.

This chapter describes a preliminary experiment to investigate the possible uses of predictive displays in the area of continuous chemical process control, and especially their effect on process operator training.

2. METHOD

A simulated task was sought which would be within the capabilities of student subjects, and yet which would be representative of the type of process found in a chemical plant. The simulated continuous stirred tank reactor (CSTR) was judged to meet both these requirements.

Subsequent sections describe the simulation, subject's task, experimental design, procedure and mode of data collection used in this study.

2.1 The Simulation

A hypothetical chemical industry control task, that of a simplified continuous stirred tank reactor (CSTR), was used in this study. The simulation was adapted from Luyben (1973). The process is shown diagramatically in Figure 16. Basically 'Kettle', as the simulated process has come to be known, comprised a lagged tank with input and output flows of liquid, and an independent heating/cooling flow through the lagging jacket. Following a chemical reaction in the tank, the

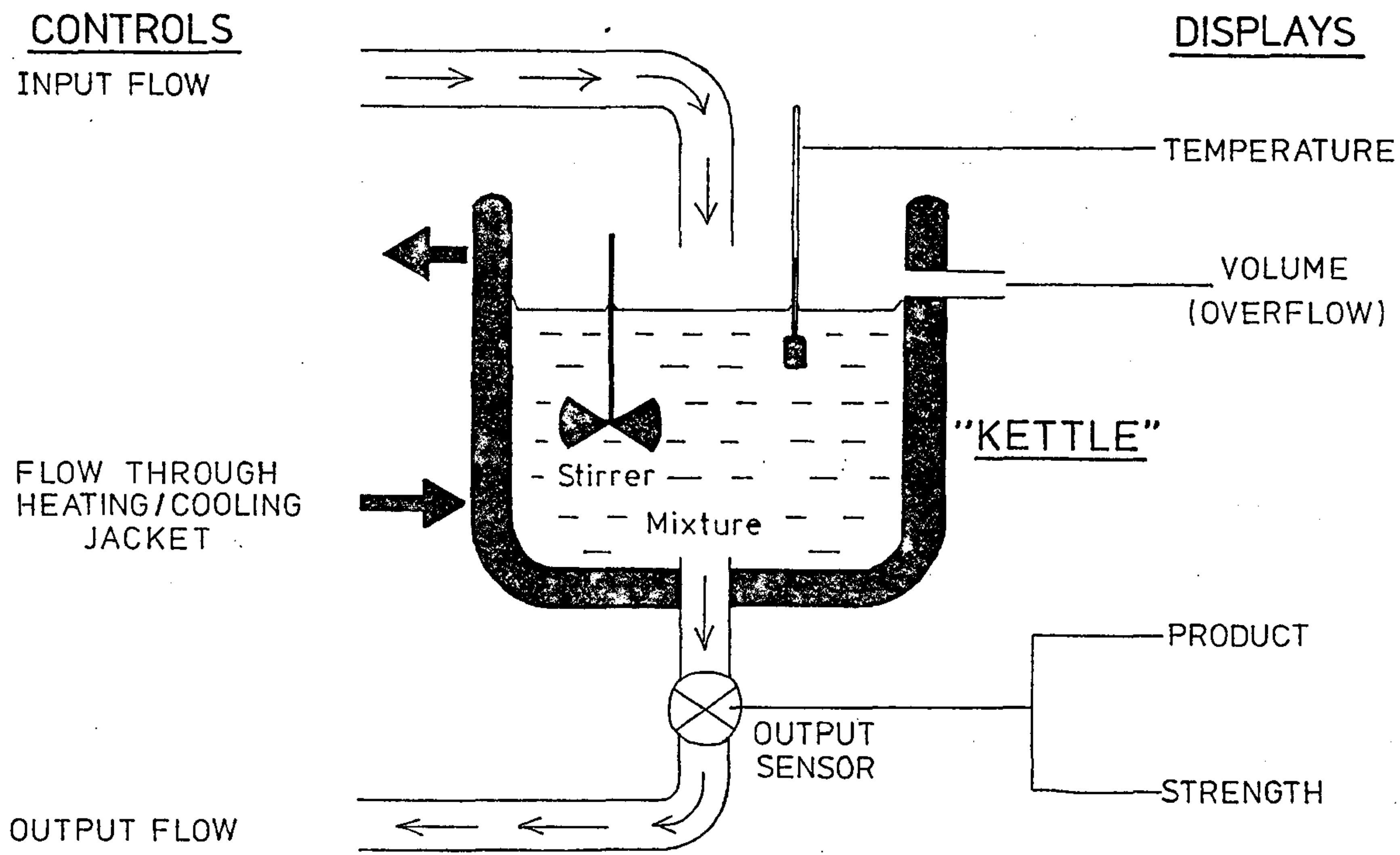


Figure 16: The simulated process - continuous stirred-tank reactor (CSTR)

output flow contained a certain concentration of finished product. The reaction curves were arranged such that maximum production and concentration were achieved when the tank was full without overflowing, and the temperature steady at 50°. Production tailed off as the tank volume decreased, or as the temperature of the mixture fell away either side of 50°. Production ceased when the temperature fell below 25° or rose above 75°. It can be seen that with tank volume held constant, production figures were thus dependent on temperature. The reaction characteristics can be expressed as follows:

$$\text{Temperature} = \frac{\text{Total Heat}}{\text{Total Volume}}$$

where Total Heat was a function of the Flow In temperature, existing heat inside the vessel, and rate of reaction.

$$\text{Volume} = \int (\text{Flow In} - \text{Flow Out}).\ dt$$

where Flow In ceased automatically if the vessel overflowed, by means of an overflow valve.

$$\text{Strength} = \frac{(100 - \text{Volume of Reactant})}{\text{Total Volume}}$$

and Product = Flow Out.(Strength - 50)

where production started only when Strength had risen above 50, since saleable product was assumed to require a minimum strength of 50%.

2.2 The Task

The experimental subjects' task was to maximise the amount and concentration of product contained in the output of the CSTR. This primary objective was in turn achieved by fulfilling the two secondary objectives of maintaining the tank full without over-flowing and at the same time keeping the temperature as near as possible to the 50° mark. The time taken initially to fill the vessel and start the reaction was also an important factor in determining the final amount of product made. In order to enhance subjects' interest in this otherwise, perhaps, unappealing problem, they were informed that the process they were to control formed part of a new design of distillery manufacturing a certain alcoholic product.

The display presented to the subjects is shown in Figure 17. The display comprised four vertical meters indicating temperature inside the vessel (with an alarm indication if the temperature rose above 90°), kettle volume (with overflow warning), amount of product manufactured and its strength. All meters were calibrated from 0 to 100 with a pointer moving up and down at the right of each scale. For the temperature meter an option was provided of a computer prediction of how the temperature would vary during the next minute. It has already been noted that with volume held constant, production figures were dependent on temperature. Two bases for the prediction were available: an approximate, unsmoothed Taylor series extrapolation (Tay) which effectively fitted a curve through the last three temperature values, or a 'Perfect Predictor' model (PPM) which was in fact the simulation itself run in fast-time. The Taylor series extrapolation model took the form:

$$y\,(t+\Theta) = y(t) + \dot{y}\,(t).\Theta + \ddot{y}\,(t).\Theta^{2}/_{2!} + \dddot{y}\,(t).\Theta^{3}/_{3!}$$

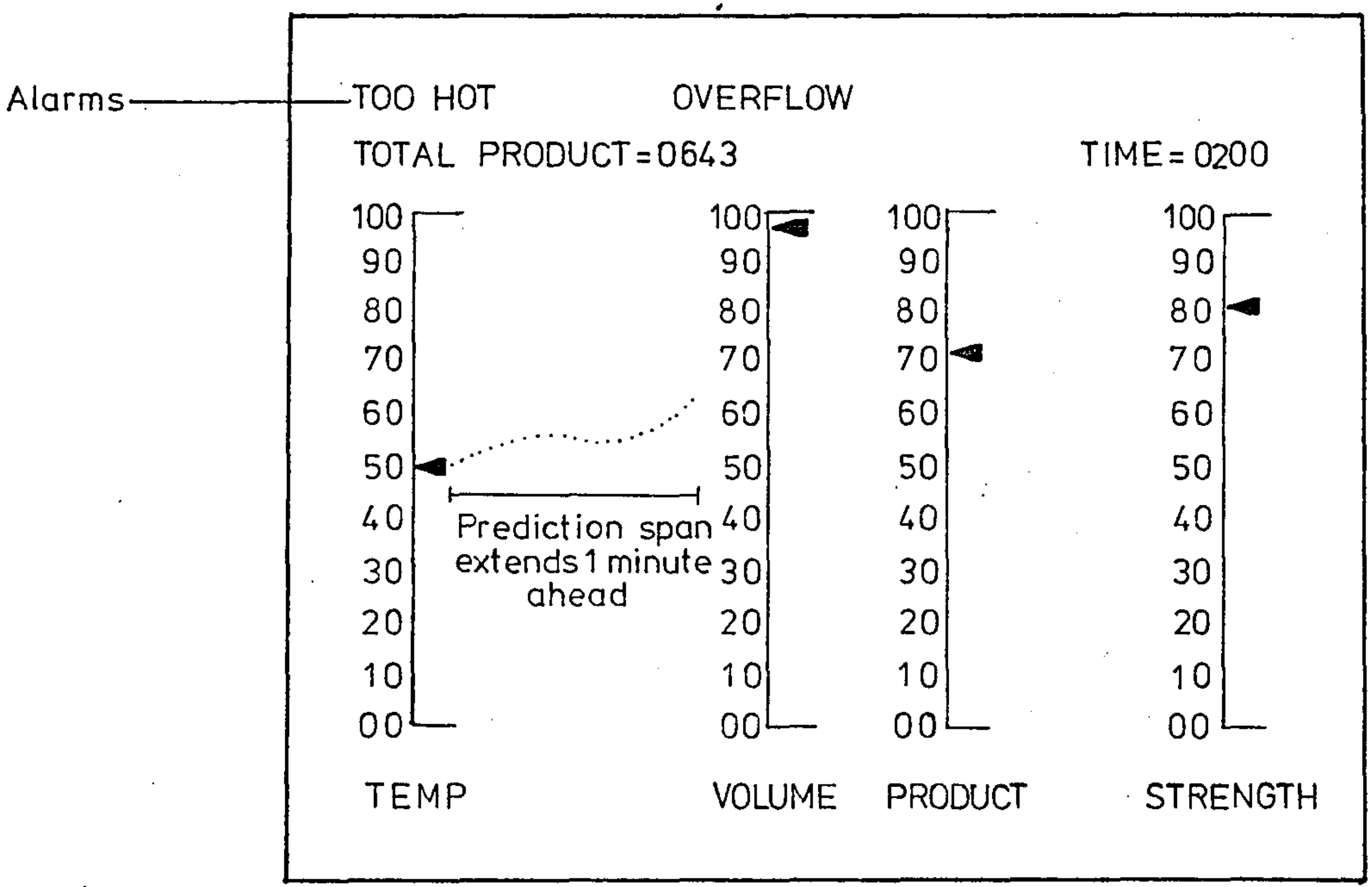


Figure 17: The display with predictive display option part-way through a trial

where $y(t+\Theta)$ is the predicted value of y at Θ seconds ahead of current time t, and $\dot{y}(t)$, $\ddot{y}(t)$, $\dddot{y}(t)$ are respectively first, second and third order derivative terms. In addition to this information a seconds counter was provided in the top right of the display showing elapsed time since the start of the trial, and a total product counter in the top left giving an approximate indication of integrated product over time. The special purpose control unit is illustrated in Figure 18. Subjects could control the flow into and the flow out of the tank, and the heating/cooling supply through the jacket, by means of slider potentiometers.

Initial conditions for the simulation were keyed in from a teletype by the experimenter at the start of each trial. System variables included the option of a predictor trace, whether prediction was based on a Taylor series (Tay) or Perfect Predictor model (PPM), and whether input uncertainty was present. In the 'with uncertainty' condition a random walk within the limits $\pm 10^{\circ}$ was superimposed onto the input flow temperature, corresponding to the noise frequently associated with measurements from actual plant. At the end of each trial the total amount of product manufactured was printed on the teletype as feedback to the subject.

2.3 Experimental Design

Factors tested in this study were whether or not training had been carried out with a predictor trace, whether or not uncertainty was present and which type of predictor (No Predictor, Taylor series extrapolation, or Perfect Predictor model) had been used in the experimental trials. A between subjects design was used for the training and uncertainty factors, coupled to a within subjects design for the type of predictor (if any) available.

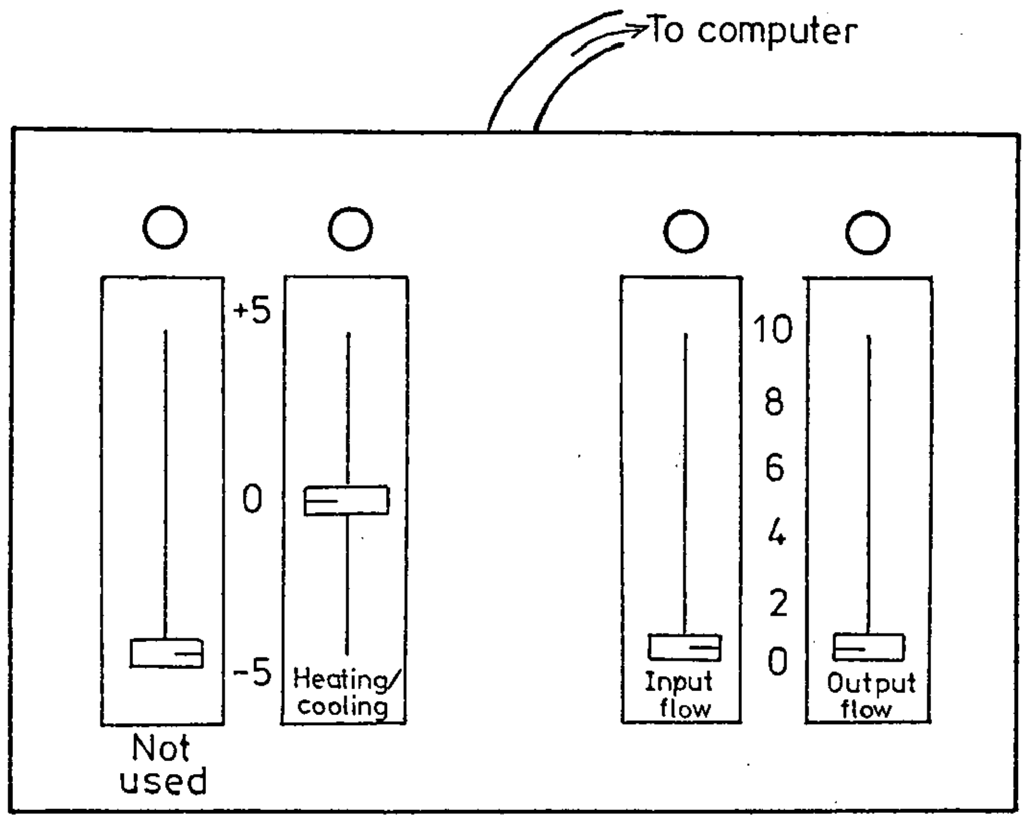


Figure 18: Special purpose control unit

Separate groups of eight randomly assigned subjects were trained with or without a predictor trace respectively, and within these main groups two sub-groups each of four subjects undertook all their trials with or without uncertainty present. Each subject was tested under all three predictor conditions in a part-balanced order - half the subjects encountered the No Predictor condition first, and half encountered the Perfect Predictor condition first. The design may be represented as follows:

		NP	Tay	PPM
Trained with Predictor	Uncertainty	G1	G1	G1
	No "	G2	G2	G2
Trained without Predictor	Uncertainty	G3	G3	G3
	No "	G4	G4	G4

where G1, G2, G3, G4 represent independent groups of four subjects who underwent all NP, Tay and PPM conditions in a part-balanced order.

2.4 Procedure

The experiments were carried out on the Departmental PDP-12 computer. On arrival, subjects were given a written set of instructions to read (Appendix 5.1 - 5.2) and the nature of the task was demonstrated. The contents of the instructions sheet were repeated verbally to ensure subjects understood them. Six training trials were carried out to ensure familiarity with the problem. Half the subjects were trained with a predictor trace and half were trained without. The experimental trials then followed. Each trial lasted for 5 minutes with a brief interval between trials. During this time the subjects' comments were elicited on the trial they had just run. After all the trials had been run, the subjects' overall impressions were noted.

2.5 Subjects

The subjects used in this study were students or research workers at Loughborough. All had some mathematical background. They were paid £1 per hour for taking part.

2.6 Data Collection

Total product manufactured scores were obtained by integrating the actual product meter readings over the 5 minute period of each trial, and were used as the main performance measure.

3. RESULTS AND STATISTICS

Group averages of the total amount of product manufactured in the different experimental conditions are plotted in Figure 19. Original data scores can be found in Appendix 5.3.

To test for statistical significance, a multi-factor ANOVA was carried out on the total product manufactured scores, the model chosen being appropriate to designs of this type having some repeated measures. A summary ANOVA table is given in Table 5. This was followed by a Newman-Keuls multiple comparison test where statistical significance of the main effect had been found.

4. DISCUSSION

4.1 Total Product Manufactured

It can be seen from Figure 19 that a clear effect due to the type of predictor model was present, and this was confirmed by the ANOVA as statistically significant at beyond the 1% level (Table 5). Following on from the ANOVA, the Newman-Keuls multiple comparison test (Table 5a) failed to detect any significant difference between the No Predictor and Perfect Predictor conditions, whereas the Taylor series model was

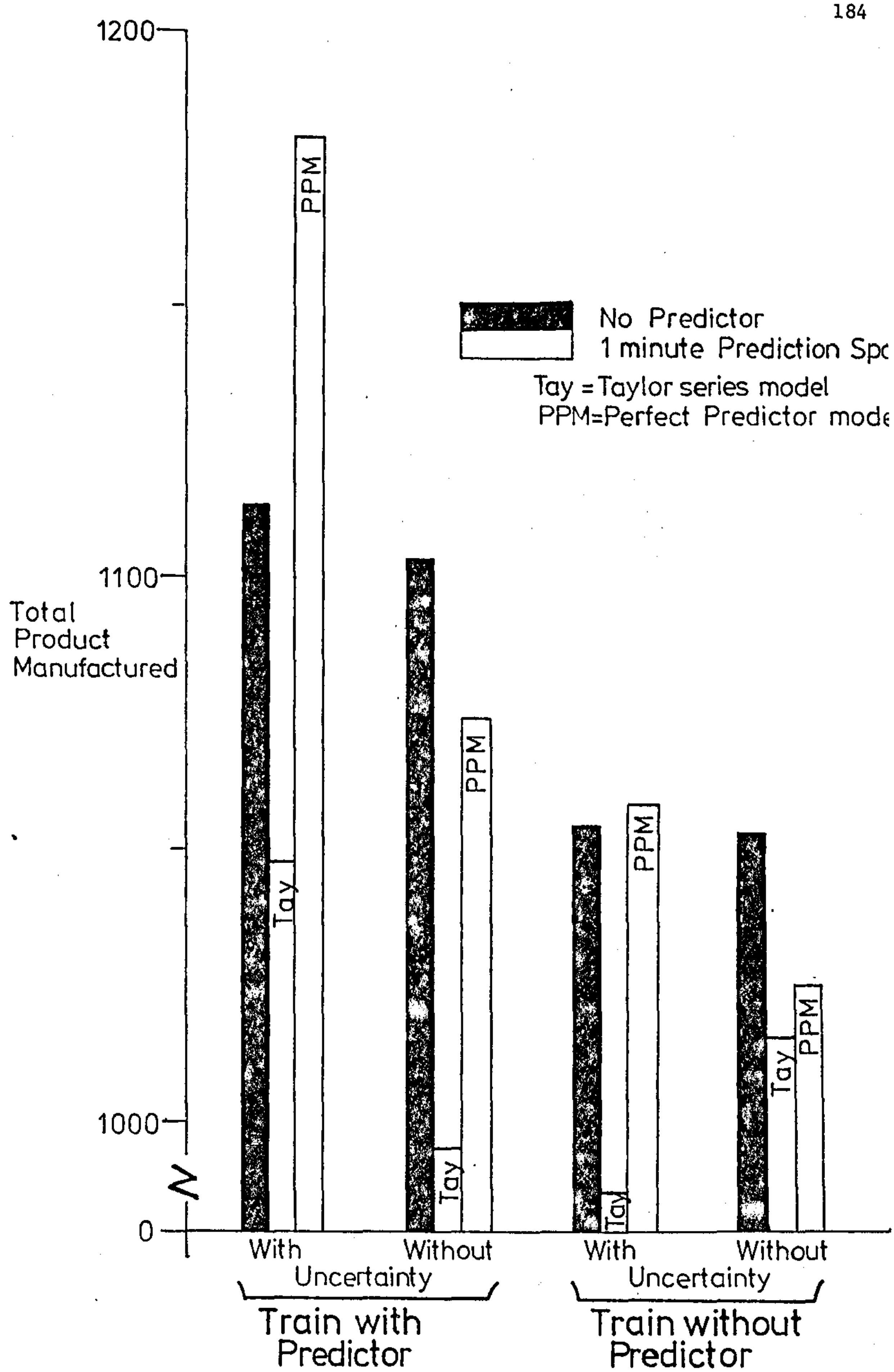


Figure 19: Total product manufactured during the 5 minute trials.

TABLE 5: Summary ANOVA table for total amount of product manufactured data

Source	Sum of Squares	df	Variance Estimate	'F'	Significance Level
Between Subjects					
Training	38,278.76	1	38,278.76	3.31	- (df 1,12)
Uncertainty	12,208.13	1	12,208.13	1.05	- (df 1,12)
Training x Uncertainty	10,755.05	1	10,755.05	0.93	- (df 1,12)
Subjects within Groups	138,990.9	12	11,582.58		
Within Subjects					
Model (NP/Tay/PPM)	58,246.64	2	29,123.32	13.13	1% (df 1,12)
Training x Model	8,896.39	2	4,448.19	2.01	- (df 1,12)
Uncertainty x Model	8,547.51	2	4,273.76	1.93	- (df 1,12)
Training x Uncertainty x Model	1,791.59	2	895.8	0.404	
Model x Subjects within Groups	53,218.37	24	2,217.43		

Conservative Test

Newman-Keuls multiple comparison test

a) Prediction model

	Tay	PPM	NP
Tay	-	1%	1%
PPM		-	-
NP			-

found to be significantly worse than either of these (significant at beyond the 1% level). Figure 19, however, suggests that this effect was dependent on uncertainty. According to Figure 19, the Perfect Predictor model gave superior scores when uncertainty was present, but in the absence of uncertainty the No Predictor condition was to be preferred. In all cases the Taylor series model apparently gave the lowest production figures. One would perhaps expect any practice or order effect to have accrued in favour of the Taylor series predictor since this condition was always encountered as the second or third experimental task. Even with this possible bias, however, the Taylor series predictor was consistently worse than the other predictor conditions. In addition, as might be expected, those subjects trained with the predictor trace attained higher production scores in all the experimental trials than did those trained without. The value of predictive displays may thus be in helping subjects to learn - or form an accurate internal model of - the plant dynamics during the training period. Unfortunately, however, this effect just misses statistical significance in the ANOVA (Table 5).

Some words of explanation are necessary to interpret the Taylor series results and the differential effect of the Perfect Predictor with uncertainty. Problems were encountered with the Taylor series extrapolation model due to inadequate smoothing. The Taylor series trace was a basic, unsmoothed predictor and as such proved to be overly susceptible to minor fluctuations in past values of temperature, causing the trace to oscillate about its mean predicted path and so reducing its credibility to the subjects. Any future testing of the Taylor series extrapolation technique on noisy data would have to incorporate some form of exponential or moving average smoothing of the raw scores if the approach is to be given a fair trial.

Concerning the differential effect of uncertainty on predictor effectiveness, the benefits from using the Perfect Predictor were most marked when uncertainty was present, and particularly where subjects had specifically been trained using the predictor. Its effectiveness diminished when subjects were not so trained. It may well have been that the nature of the task without uncertainty - the temperature tended to drop gradually towards the 50 mark of its own accord once the reaction had started - was such that the system was not of sufficient complexity to warrant any form of predictive assistance. Several subjects commented to this effect. Rouse (1970) has also suggested that predictor displays are of little benefit in simple tasks. Once random disturbances in the form of uncertainty on the input temperature had been introduced, however, the task became sufficiently challenging to warrant a predictive aid. A further complicating effect of introducing uncertainty was that the random disturbances tended to 'throw' the temperature upwards at the start of a trial, causing the reaction to start more quickly and resulting in a higher overall product score. This accounts for the slightly higher production figures in both 'with uncertainty' conditions, though the effect is not statistically significant.

In general it would seem that the problems encountered were largely a by-product of attempting to devise a simplified laboratory analogue of an actual chemical industry task. No laboratory simulation can reproduce in full the complexity and the stresses of the real-world. Two alternative courses are open for future work in this area: either to dispense with the trappings of an actual task and to devise a simple, uncomplicated laboratory problem on which to carry out an in-depth investigation of predictor characteristics; or to take the predictor concept into the field and test it in an actual industrial setting. An ideal solution would perhaps be to adopt both alternatives.

4.2 Subjective Comments

It is clear from the subjects' comments that the task could be divided into two distinct phases: filling the kettle as quickly as possible without allowing the vessel to overheat; and secondly conducting the reaction itself. Some subjects found the predictor trace very useful in the initial phase as it resulted in fewer "overheat" errors. Others found the predictive aid useful during the secondary phase of bringing the temperature down to 50 and controlling the reaction. Subjects' ratings of their preference for the No Predictor, Taylor series and Perfect Predictor models reflected the condition under which they had been trained, and tended to confirm the impressions gained from the scores of total product manufactured.

The Taylor series extrapolation was universally disliked as it "swung about erratically and served only to confuse". One subject reported he was able to overcome this problem by mentally fitting a trend line between the extremes of oscillation. It is to be expected that with adequate smoothing of the data points used as the basis for prediction this problem would be overcome. Overall, eleven subjects (including all subjects trained in this condition) rated the Perfect Predictor model as the easiest to control with, leaving five subjects who preferred the No Predictor condition. The latter tended to anticipat the plant response internally. Of those subjects who preferred the predictor, one failed to detect any difference between the two prediction models. Several subjects noticed and made use of a side-effect of the Perfect Predictor trace. The simulation was such that the trace became slightly uneven just before the kettle overflowed, a consequence of reciprocation in the input flow valve. Thus overflow and temperature information could be obtained from the same trace.

It is interesting to note that two subjects requested a predictor trace be available to show rate of product manufacture, i.e. changes in the main performance criterion. (Unfortunately the design of the system would not permit this.) Their complaint with the temperature predictor was that it did not directly relate to the overall objective of the task: to manufacture as much product as possible in the time given. In the absence of a rate-of-manufacture predictor these subjects experimented with different settings of temperature and volume in an attempt to find a trade-off between the two secondary criteria which would result in the maximum rate of climb of the product counter. From this anecdotal evidence it seems that predictor displays are potentially of most benefit if they relate directly to the overall performance criterion.

5. CONCLUSIONS

This study should be regarded as a preliminary investigation into the use of the predictive display concept in continuous chemical process control. It has shown that given the right conditions, i.e. adequate training, a credible predictor and a sufficiently demanding problem, the application of predictive displays can lead to an improved performance. Problems were encountered through trying to simulate a representative chemical industry task with naive student subjects in the laboratory. It is recommended that future work in this area should first involve a much-simplified laboratory problem on which an in-depth study of predictive display characteristics could be conducted. A second phase would comprise a full-scale field study to test the best overall design in an actual industrial setting. This approach is followed in Chapters 6 and 7.

CHAPTER 6

AN EXPERIMENT TO EVALUATE THE EFFECT OF VARYING TASK CHARACTERISTICS AND PREDICTIVE DISPLAY PARAMETERS IN A SIMULATED CONTINUOUS DUAL-METER MONITORING AND CONTROL TASK

1. OBJECT OF THE EXPERIMENT

The previous chapter has shown that given the right conditions, a predictor display can help in the control of a laboratory simulated, continuous chemical process. Before testing the concept in the field, however, it is necessary to determine how varying the parameters of such a display affects operator performance, with the object of arriving at recommended display configurations for particular situations.

The present chapter explores in more detail the characteristics which a predictor display for continuous process control tasks should possess, with particular regard to such factors as the effect of input uncertainty, prediction span, plant gain (K), and fidelity of the prediction model employed.

2. METHOD

A laboratory task was sought which could be generalised to a variety of monitoring and control situations, which would facilitate the detailed investigation of a number of predictor display design features, and yet at a level of complexity which could be mastered by naive subjects without the need for extensive training. The simulated dual-meter monitoring and control task used in this experiment was derived from experience gained in the CSTR simulation of the previous chapter. The latter simulation had suffered from two principal drawbacks:

1) It was a simplified version of a hypothetical chemical industry task (Luyben, 1973).
2) The temperature predictor tested, whilst of indirect benefit, was not directly related to the main system objective of profit maximisation.

These factors combined to render the predictor not as operationally significant as it might otherwise have been. Therefore in the present study any similarity with a specific chemical process was removed, and the task became one of monitoring the state of two process meters and keeping the movement of their pointers simultaneously within prescribed limits. Such a procedure is at the heart of many process, and other, control problems.

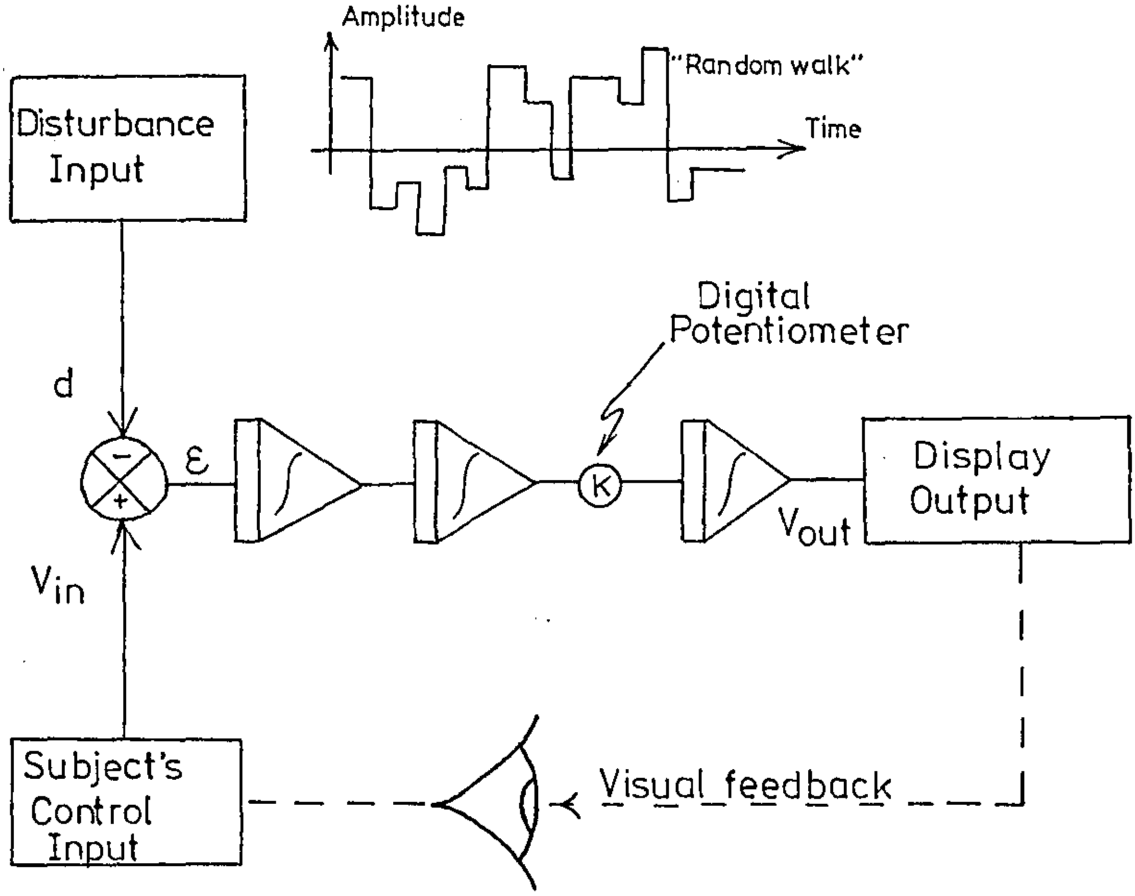
Subsequent sections describe the simulation, subject's task, experimental design, procedure and mode of data collection used in this study.

2.1 The Simulation

The simulation consisted of two independent unstable channels having identical third-order dynamics (Figure 20). Each channel comprised three integrations in series with a digital potentiometer, and was driven by an error signal (ε) derived from the subject's control input (Vin) minus a disturbance level (d). The disturbance level varied randomly in its magnitude, duration and direction, but on average changed once every 10 seconds or so. The output (Vout) from each channel controlled the position of a pointer set against the calibrated scale of a vertical process-type meter.

2.2 The Task

The experimental subject's task was to anticipate the path of two pointers moving against vertical scales calibrated 0-100 and by his control actions to compensate for their movement away from the 50 mark, keeping both pointers simultaneously between 45 and 55 on the scales. The display arrangement is shown in Figure 21, and the



$$Vout = {}^{K}/_{S^3}\ \mathcal{E}$$

where $\mathcal{E}$ = (Vin - d)

S^3 represents a third-order system and
K is the plant gain.

Figure 20: Identical dynamics of the two simulated channels

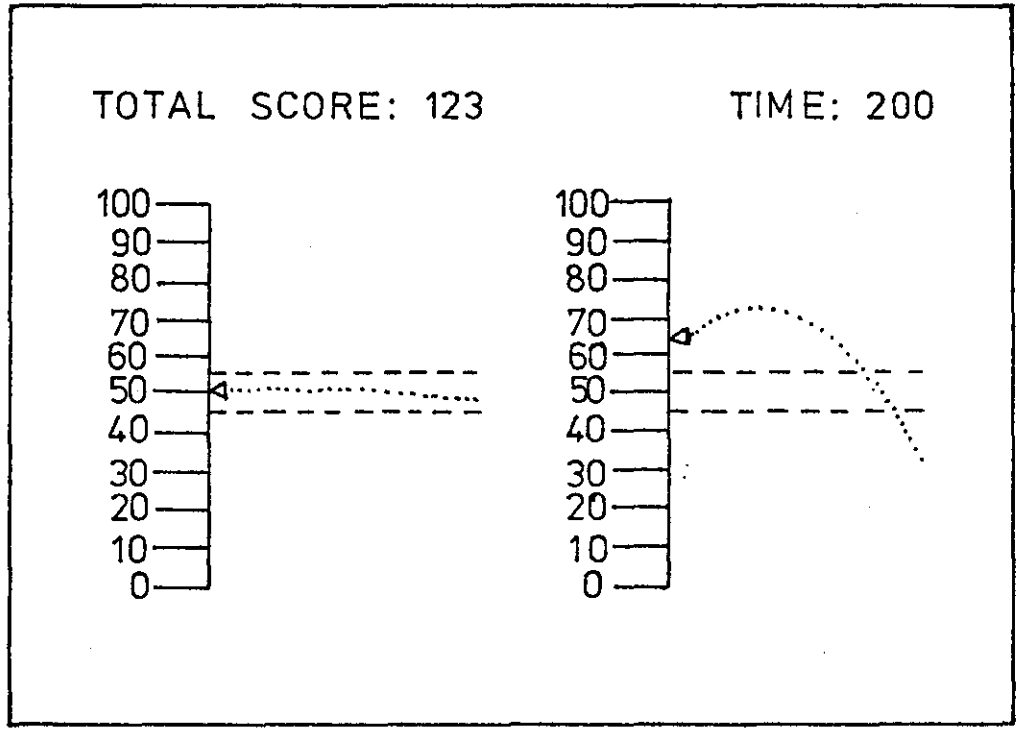


Figure 21: The display part-way through a trial with 30 seconds prediction span. (The predictor traces indicate the left-hand pointer to be relatively stable, whereas the right-hand pointer will pass through the limits and continue downwards. Note that only one of the pointers can be viewed at a given instance.)

control unit in Figure 22. A seconds counter was provided on the display, together with an indication of how many seconds the pointers had been simultaneously within limits.

The pointers could not be viewed simultaneously but were instead selected by pressing a button above the appropriate slider on the control unit. To assist control, the option of a predictor trace was provided extending to the right of both pointers. Prediction was based either on an approximately accurate Taylor series extrapolation model (Tay) using the three most recent data points, or on a 'Perfect predictor' model (PPM) based on the simulation itself run in fast-time. The Taylor series extrapolation model took the form:

$$y(t + \Theta) = y(t) + \dot{y}(t).\Theta + \ddot{y}(t).\Theta^2/2! + \dddot{y}(t).\Theta^3/3!$$

where $(t + \Theta)$ is the predicted value of y at Θ seconds ahead of current time t. $\dot{y}(t)$, $\ddot{y}(t)$, $\dddot{y}(t)$ are respectively first, second and third order derivative terms.

The determination of an appropriate predictor model is largely an engineering problem, and the two models used here were chosen as being representative of different extremes of computational power requirements. Smoothing problems encountered in the previous chapter with the Taylor series approach were overcome by generating derivative terms directly from the simulation, rather than from successive values of the output. In a practical application where one would be forced to use successive values of the output as the basis for calculating derivative terms, output smoothing could be achieved by established techniques (moving averages, digital filters etc.).

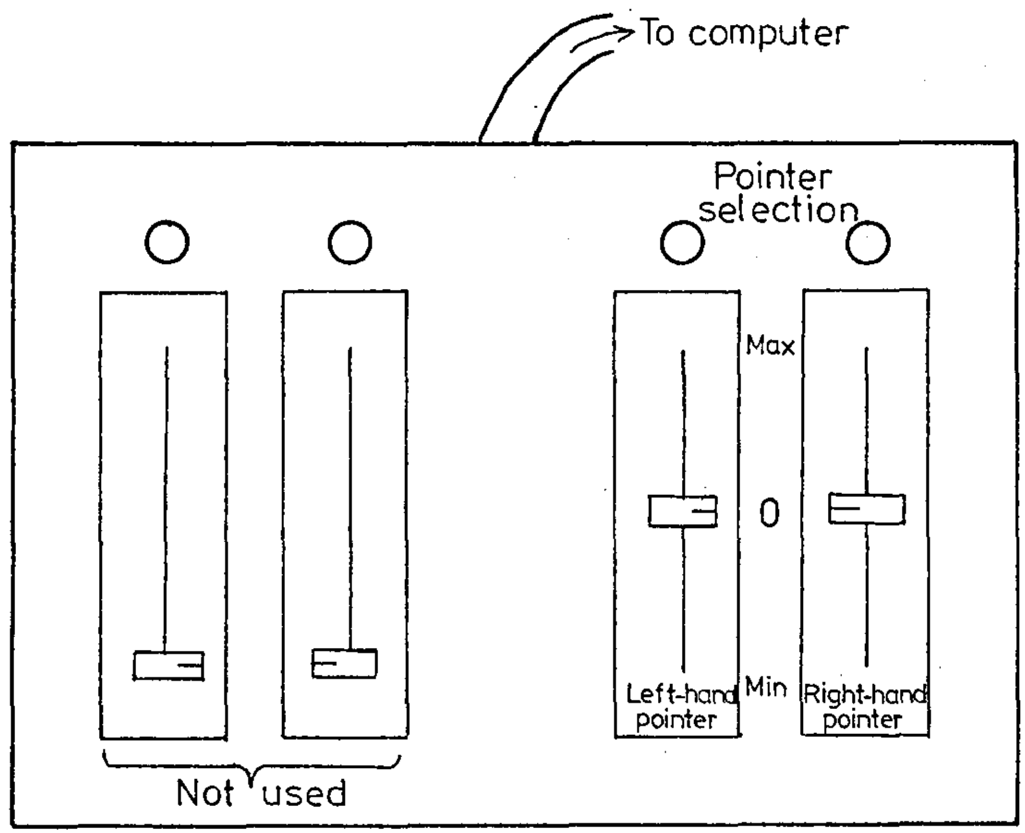


Figure 22: Subject's control panel

Initial conditions for the simulation were keyed in by the experimenter from a teletype at the beginning of each trial. Adjustable variables were pointer limits, plant gain (effectively system speed of response, achieved by adjusting the potentiometer value), level and timing of the random disturbances, type of prediction model used, prediction span (length of predictor trace from 0 up to 30 seconds), and trial length in seconds. At the end of each trial the total time within limits score was output on the display as feedback to the subjects.

2.3 Experimental Design

Factors examined in this study were the level of plant gain, the level of random disturbances (uncertainty), the prediction model used (Tay or PPM) and the prediction span. A between-subjects design was employed for plant gain, coupled to a within-subjects design for the remaining factors. Three separate groups each of 5 subjects were randomly assigned to one of three gain conditions (low, medium, high gain - corresponding to slow, moderate and fast system responsiveness). Each subject underwent three levels of uncertainty (disturbance levels $\pm 0^{\circ}$, $\pm 10^{\circ}$ and $\pm 20^{\circ}$), four levels of prediction span (0, 5, 15 and 30 seconds) and both types of prediction model (Taylor series and Perfect Predictor). The presentation order of the 21 trials was randomised to overcome sequence effects, whilst a thorough training schedule ensured that practice effects were minimal compared to the magnitude of the experimental effects.

The design may be represented as follows:

PREDICTION MODEL/
PREDICTION SPAN (seconds)

			Tay	PPM	Tay	PPM	Tay	PPM
		0	5	5	15	15	30	30
LOW GAIN	Low Uncertainty	G1	G1	G1	G1	G1	G1	G1
	Medium Uncertainty	G1	G1	G1	G1	G1	G1	G1
	High Uncertainty	G1	G1	G1	G1	G1	G1	G1
MEDIUM GAIN	Low Uncertainty	G2	G2	G2	G2	G2	G2	G2
	Medium Uncertainty	G2	G2	G2	G2	G2	G2	G2
	High Uncertainty	G2	G2	G2	G2	G2	G2	G2
HIGH GAIN	Low Uncertainty	G3	G3	G3	G3	G3	G3	G3
	Medium Uncertainty	G3	G3	G3	G3	G3	G3	G3
	High Uncertainty	G3	G3	G3	G3	G3	G3	G3

where G1, G2, G3 represent independent groups of five subjects who underwent all Uncertainty, Prediction model, and Prediction span conditions in a randomised order.

2.4 Procedure

The experiments were carried out on the Departmental PDP-12 computer. On arrival, subjects were given a written set of instructions to read (Appendix 6.1) and the nature of the task was demonstrated. The contents of the instructions sheet were repeated verbally to ensure comprehension. Six training trials were then carried out under a medium level of uncertainty ($\pm 10^{\circ}$ disturbance level). A standard training order was used: No Predictor (NP), Perfect Predictor Model (PPM), Taylor series extrapolation model (Tay) PPM, Tay, NP. An abbreviated results printout was obtained for each training trial.

The experimental trials then followed, their order being randomised. Each trial lasted for 5 minutes with a break of 4 minutes between trials. During this time subjects filled in a short questionnaire giving their subjective impressions of their last trial. A sample questionnaire is shown in Appendix 6.4. An extended coffee break was given midway through the experiment. After all the trials had been run, subjects' overall impressions were noted.

2.5 Subjects

The subjects used in this study were undergraduate students at Loughborough. All had some mathematical background and were paid £1 per hour for taking part.

2.6 Data Collection

An automatic data capture program logged every system input made by the subject, together with system states such as pointer positions, into store at 1 second intervals. At the end of each trial an abbreviated results printout could be obtained showing time within limits for each pointer and both pointers simultaneously, time spent looking at each pointer, integrated error scores for each channel plus histograms of control actions and pointer positions. In addition the option of a fuller printout giving control and pointer positions at 1 second intervals, disturbance levels and last predicted value displayed at 10 second intervals, and a breakdown of channel switchings could be selected. A comprehensive analysis of each trial was thus possible, which could then be matched to subjects' comments and results from the questionnaire analysis.

3. RESULTS AND STATISTICS

Time in seconds during which one or both pointers had been outside the prescribed limits was used as the main performance measure. (In practice, integrated absolute error scores - the total pointer deviations from the 50 scale marker - were found to give similar results, and so are not reported here.)

Group averages of the performance measure in the different experimental conditions have been plotted in Figures 23-26. Figure 23 shows the grand averages of the Taylor series extrapolation model and Perfect predictor model for the four prediction span conditions (each point on the graph is the average of three gains, three levels of uncertainty and five subjects). Figures 24-26 expand the basic information of Figure 23 to include the effects of different levels of uncertainty and plant gain, separate graphs being drawn for Low, Medium and High gain. Data scores for individual subjects can be found in Appendices 6.2-6.3. Much could also be learnt about individual subjects' control strategies from scrutiny of the control histograms for each trial, from subjective comments and from the completed questionnaires. These will be discussed in section 4.

Inspection of the time outside limits data showed it to be severely positively skewed. As is appropriate with severely skewed time data of this kind, the within-cell variances were first stabilised before ANOVA analysis by performing a logarithmic transform on the raw error scores, of the form:

$$X'_{ijk} = \log_{10} (X_{ijk} + 1)$$

The addition of 1 to each X_{ijk} term served to prevent the occurrence of $\log_{10}(0)$.

The transformed data were analysed in several different ways. A preliminary analysis (Table 6) was used to test for broad differences between the No Predictor, Taylor series extrapolation model (30 seconds prediction span) and Perfect Predictor model (also 30 seconds prediction span) conditions. More detailed analyses were also performed on the full set of Taylor series data (Table 7) and on the complete set of Perfect Predictor data (Table 8). The ANOVA model used was appropriate to multi-factor designs of this type containing some repeated measures, and was followed by tests of simple effects where a significant interaction term had been obtained.

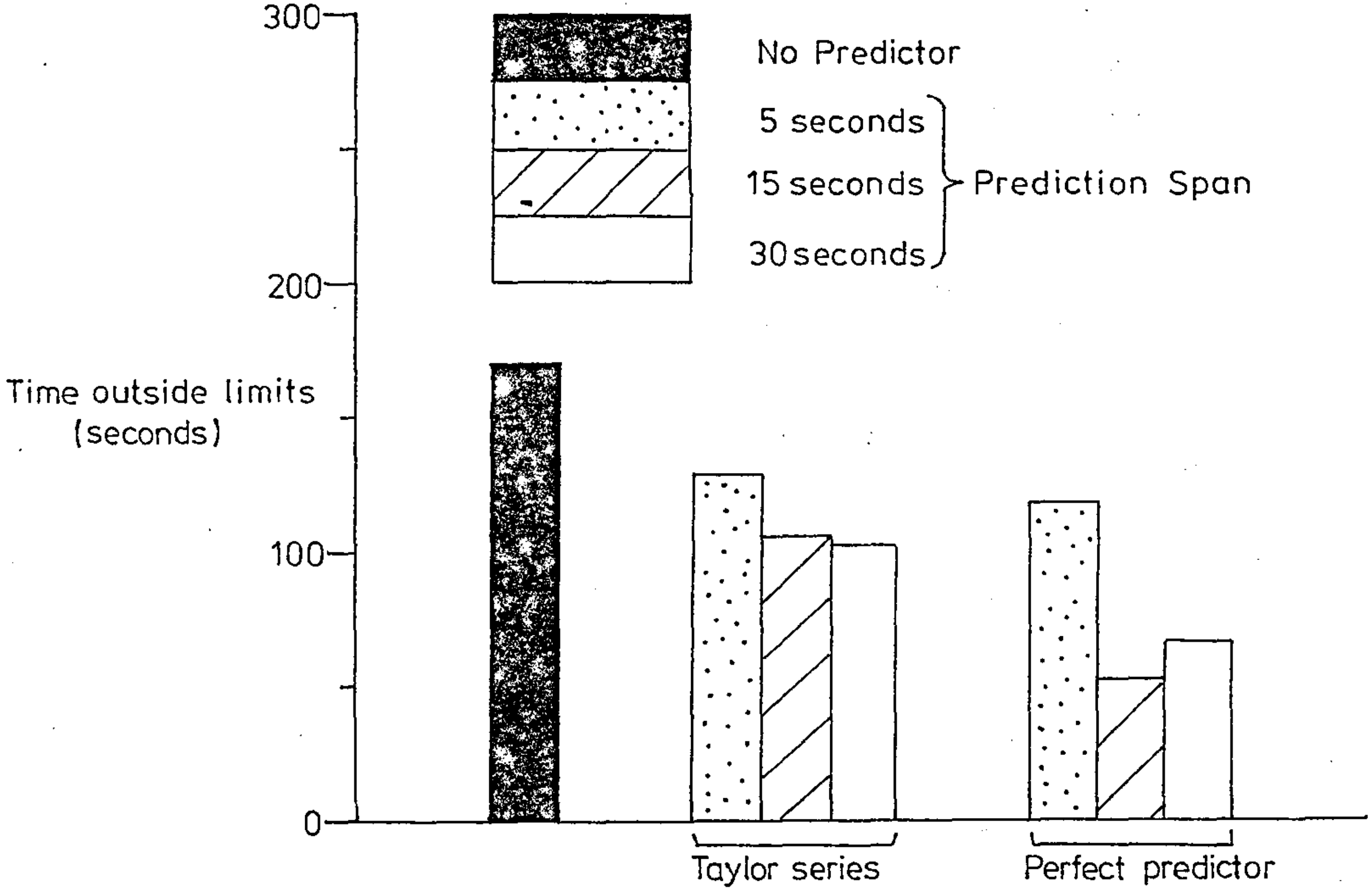


Figure 23: Grand averages of time outside limits scores

TABLE 6: Summary ANOVA for log transformed NP vs Tay (30 second) vs PPM (30 second) scores (time outside limits data)

Source	Sum of Squares	df	Variance Estimate	'F'	Significance level
Between Subjects					
Gain	59.274	2	29.637	100.24	0.1% (df 2,12)
Subjects within groups	3.548	12	0.296		
Within Subjects					
Uncertainty	4.757	2	2.379	25.82	0.1% (df 1,12)
Gain x Uncertainty	3.871	4	0.968	10.51	1% (df 2,12)
Uncertainty x S.w.g.	2.211	24	0.092		
Model (NP/Tay/PPM)	32.522	2	16.261	34.47	0.1% (df 1,12)
Gain x Model	9.657	4	2.414	5.12	5% (df 2,12)
Model x S.w.g.	11.321	24	0.472		
Uncertainty x Model	6.781	4	1.695	22.39	0.1% (df 1,12)
Gain x Uncertainty x Model	4.86	8	0.608	8.03	1% (df 2,12)
Uncertainty x Model x S.w.g.	3.633	48	0.0757		

Conservative Test

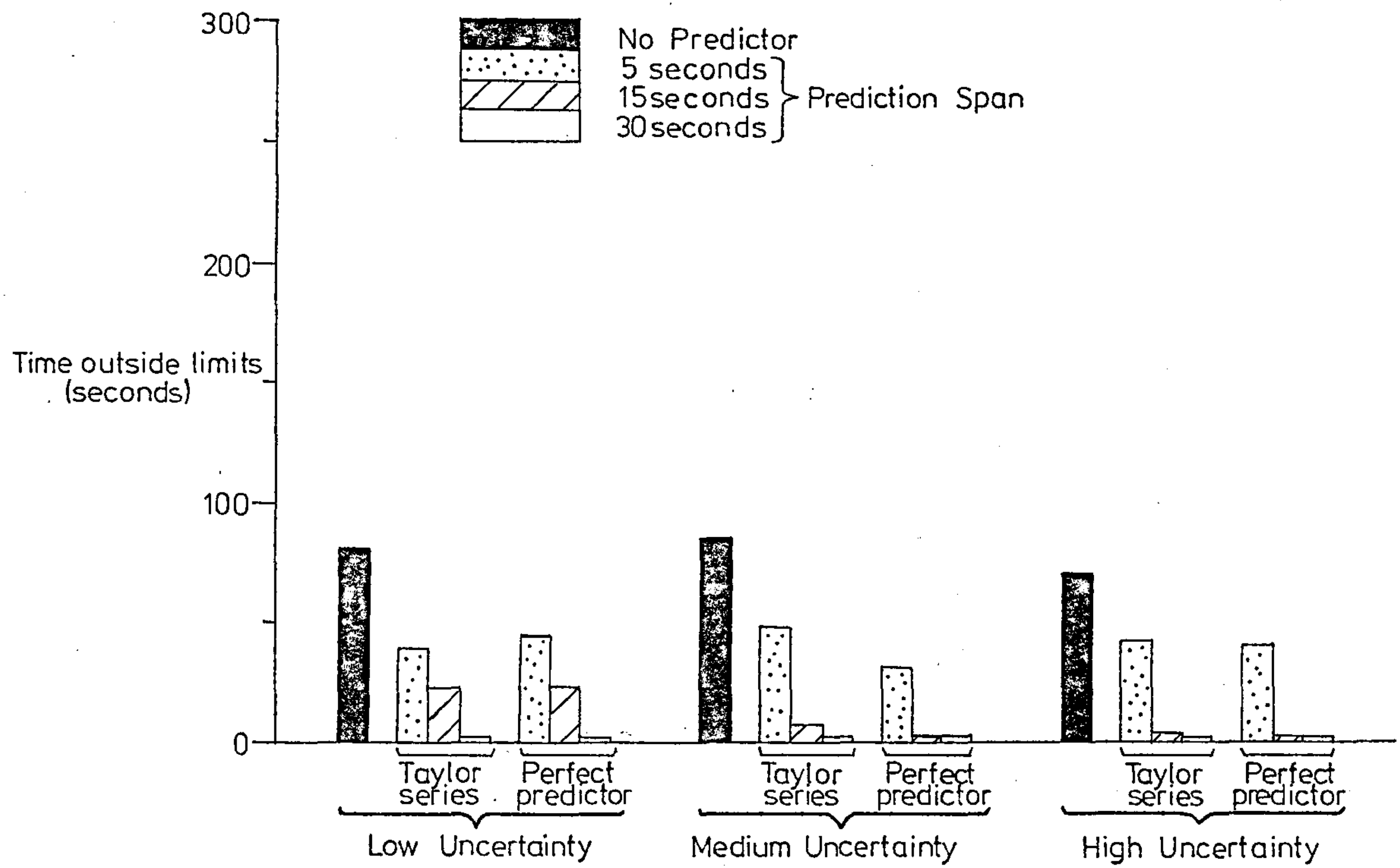


Figure 24: Time outside limits scores .. low gain (slow response time)

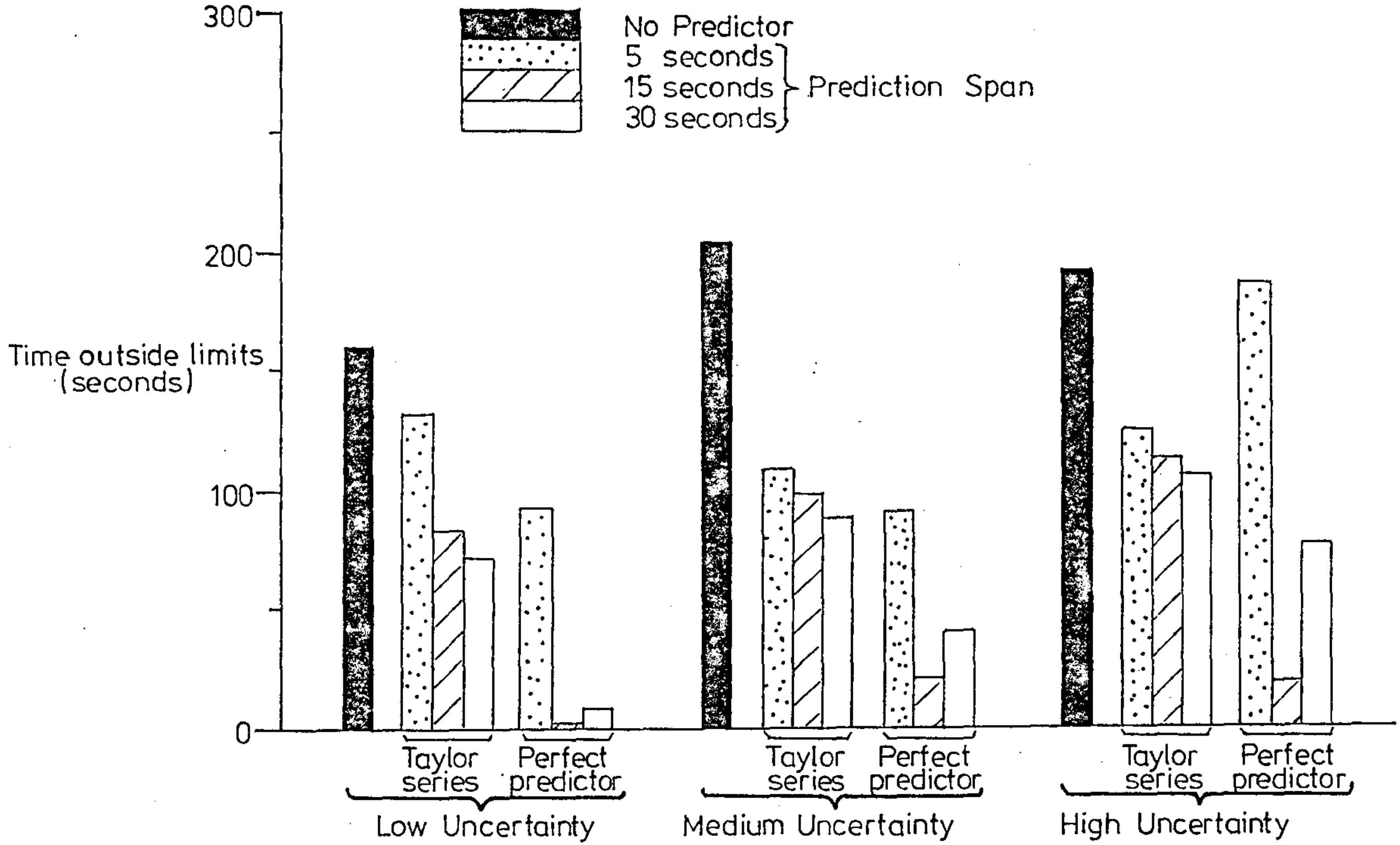


Figure 25: Time outside limits scores .. medium gain (moderate response time)

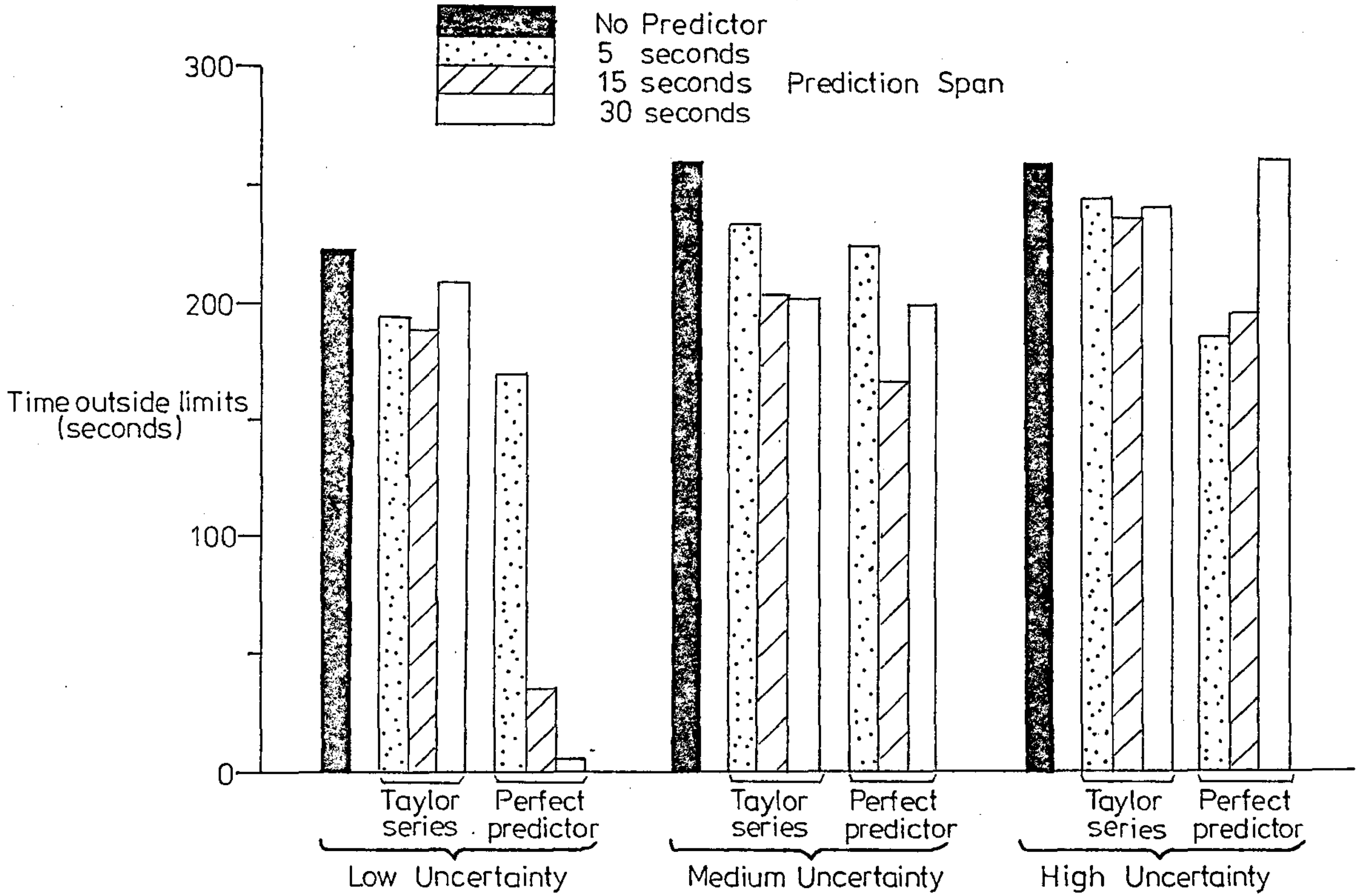


Figure 26: Time outside limits scores .. high gain (fast response time)

TABLE 7: Summary ANOVA for log transformed Taylor series scores (time outside limits data)

Source	Sum of Squares	df	Variance Estimate	'F'	Significance level
Between Subjects					
Gain	82.827	2	41.413	32.98	0.1% (df 2,12)
Subjects within groups	15.07	12	1.256		
Within Subjects					
Uncertainty	0.0475	2	0.0238	0.16	- (df 1,12)
Gain x Uncertainty	0.7887	4	0.1972	1.29	- (df 2,12)
Uncertainty x S.w.g.	3.665	24	0.1527		
Prediction Span	13.66	3	4.553	13.05	1% (df 1,12)
Gain x Prediction Span	9.988	6	1.665	4.77	5% (df 2,12)
Prediction Span x S.w.g.	12.557	36	0.349		
Uncertainty x Prediction Span	0.455	6	0.0758	0.59	- (df 1,12)
Gain x Uncertainty x Prediction Span	2.065	12	0.1721	1.35	- (df 2,12)
Uncertainty x Prediction Span x S.w.g.	9.193	72	0.1277		

Conservative Test

a) Tests on Simple Effects (Gain x Prediction Span Interaction)

Source	Significance
Between Gains at 0 seconds Prediction Span	5%
Between Gains at 5 seconds Prediction Span	0.1%
Between Gains at 15 seconds Prediction Span	0.1%
Between Gains at 30 seconds Prediction Span	0.1%
Between Prediction Spans at Low Gain	0.1%
Between Prediction Spans at Medium Gain	5%
Between Prediction Spans at High Gain	-

TABLE 8: Summary ANOVA for log transformed Perfect Predictor model scores (time outside limits data)

Source	Sum of Squares	df	Variance Estimate	'F'	Significance level
Between Subjects					
Gain	48.448	2	24.224	45.76	0.1% (df 2,12)
Subjects within groups	6.353	12	0.529		
Within Subjects					
Uncertainty	7.866	2	3.933	13.88	1% (df 1,12)
Gain x Uncertainty	11.024	4	2.756	9.73	1% (df 2,12)
Uncertainty x S.w.g.	6.801	24	0.283		
Prediction Span	49.706	3	16.569	37.92	0.1% (df 1,12)
Gain x Prediction Span	4.695	6	0.783	1.79	- (df 2,12)
Prediction Span x S.w.g.	15.73	36	0.437		
Uncertainty x Prediction Span	9.197	6	1.533	5.86	5% (df 1,12)
Gain x Uncertainty x Prediction Span	8.923	12	0.744	2.84	- (df 2,12)
Uncertainty x Prediction Span x S.w.g.	18.845	72	0.262		

Conservative Test

a) Tests on Simple Effects (Gain x Uncertainty Interaction)

Source	Significance
Between Gains at Low Uncertainty	-
Between Gains at Medium Uncertainty	0.1%
Between Gains at High Uncertainty	0.1%
Between Uncertainty at Low Gain	-
Between Uncertainty at Medium Gain	1%
Between Uncertainty at High Gain	0.1%

b) Tests on Simple Effects (Uncertainty x Prediction Span Interaction)

Source	Significance
Between Uncertainty at 0 seconds Prediction Span	-
Between Uncertainty at 5 seconds Prediction Span	-
Between Uncertainty at 15 seconds Prediction Span	0.1%
Between Uncertainty at 30 seconds Prediction Span	0.1%
Between Prediction Spans at Low Uncertainty	0.1%
Between Prediction Spans at Medium Uncertainty	0.1%
Between Prediction Spans at High Uncertainty	0.1% (just)

4. DISCUSSION

4.1 Time outside limits data

Considering first the preliminary ANOVA (Table 6) which excluded prediction span by comparing scores from the No Predictor condition with scores from the Taylor series and Perfect predictor models using the full 30 seconds prediction span, it can be seen that all the main effects (gain, uncertainty, prediction model) were highly statistically significant, with the complication of considerable interactions.

In general, time outside limits scores were found to increase with faster plant response and with increasing levels of uncertainty. The significant interaction term (gain x uncertainty) suggests that uncertainty had a differential effect depending on the system responsiveness. The third main effect - that of NP vs Taylor series vs Perfect Predictor models - was highly significant, coupled to a strong interaction with uncertainty (the Taylor series model was peculiarly immune to variations in uncertainty), and a lesser interaction with plant gain.

A major finding is that there was virtually no difference between the two prediction model scores in the Low plant gain condition (Figure 24). Inspection of the original data (Appendix 6.2-6.3) shows that near perfect within-limits performance was achieved using the Taylor series extrapolation model as well as with the Perfect predictor trace. This seems to reinforce Kelley's (1960a) and Bernotat's (1972) earlier findings of the effectiveness of simple prediction models. Figure 23 demonstrates that in broad terms, however, the Perfect Predictor was clearly superior to the Taylor series model, especially with longer prediction spans, though either prediction model was preferable to No

Predictor at all. It is also evident from Figure 23 that minimum error scores were achieved with the full 30 second prediction span for the Taylor series model, but with a prediction span of only 15 seconds for the Perfect Predictor trace.

In order to study the interactions with prediction span in more detail, two separate analyses were performed on the complete data - one analysis of the Taylor series scores and a separate analysis of the Perfect predictor model scores.

4.2 Interpretation of Taylor series data

The analysis of Table 7 indicates that for the Taylor series prediction model a strong effect due to plant gain (0.1% significance) was found, and a somewhat lesser effect (1% significance) due to prediction span. A slight interaction between these two variables was also present. No effect was discovered due to uncertainty, and it would seem to be one of the important features of using a Taylor series prediction model that no significant worsening in performance can be expected as the level of input disturbance rises. (The slow response time of such a model may well have served to act as a filter to the input noise.)

Because the gain x prediction span interaction was significant, tests on simple main effects were called for rather than further direct testing of the main effects. Results of such tests are given in Table 7a. Examining the interaction effect in more detail suggests that plant gain had an increasingly significant effect when any form of Taylor series predictor trace, however short, was introduced. Conversely, the effect of different prediction spans was most marked for slow plant response, longer spans resulting in lower error scores, but its effect lessened as plant gain was increased. No significant difference between different spans was found in the High gain condition.

Clearly the choice of prediction span using this type of extrapolation model will depend on the gain of the system concerned. For systems with slow or moderate response times the maxim "the longer the better" appears to be valid. For systems with very fast response times a slight reduction in prediction span may be advisable on practical grounds

4.3 Interpretation of Perfect predictor data

The analysis shown in Table 8 demonstrates that all the main effects (gain, uncertainty, prediction span) achieved a high degree of statistical significance, in addition to a strong gain x uncertainty interaction term (significant 1%), and a lesser uncertainty x prediction span term (significant 5%). It is evident that the Perfect predictor model reacted somewhat differently to changes in the experimental conditions than did its Taylor series counterpart. In both cases performance deteriorated as speed of plant response increased, (Figures 24-26) but in contrast to the Taylor series data the Perfect predictor was also adversely affected by increasing the level of uncertainty. This effect was somewhat dependent on the prediction span in use, a more marked deterioration in performance occurring for longer prediction spans. This point will be further discussed below.

Because the two interaction terms achieved significance, tests on simple main effects were again called for rather than further direct testing of the main effects. Findings from the analyses are summarised in Table 8 a,b. Considering first the gain x uncertainty interaction, this point is perhaps of rather academic interest as the analysis is in terms of means obtained by averaging over scores from the four prediction span conditions. The analysis reveals that no significant differencewas present between the different plant gains at Low uncertainty levels, but that a considerable difference existed under Medium and High uncertainty. Inspection of Figures 24-26 makes it evident that while this may have been so for the group means, the choice of prediction span to be used was a major complicating factor - indeed an uncertainty x prediction span interaction term was also found (see below). Similarly the analysis for averaged prediction spans indicated that there was no effect due to uncertainty for Low plant gains (Figure 24), but an increasingly significant effect as response speed increased through Medium (significant 1% - Figure 25) to High gains (significant 0.1% - Figure 26). Again it must be stressed that the nature of this part of the analysis excluded vital information associated with the different prediction spans.

Considering the uncertainty x prediction span interaction, tests for simple main effects showed no difference due to uncertainty for short prediction spans (0 and 5 seconds), but a highly significant effect (significant 0.1%) for prediction spans of 15 and 30 seconds. Figures 24-26 represent this interaction pictorially. In terms of prediction spans, though the effect due to different spans was highly significant for all levels of uncertainty, it was most pronounced at lower uncertainty levels. Inspection of Figures 24-26 suggests that an optimum prediction span existed for the Perfect predictor at higher levels of uncertainty. This point has not been revealed by the analysis so far, and so it was decided to carry out trend tests on each gain x uncertainty combination to explore the issue further. The test thought to be most appropriate was Page's L non-parametric test on trends, as this test is more powerful than the omnibus F-test or the equivalent Friedman test, and makes no assumptions about underlying distributions (Boersma et al., 1964). The predicted order amongst prediction spans tested was in accord with the cell averages for the Perfect predictor data shown in Figures 24-26 and the significance levels obtained are given below:

	LOW UNCERTAINTY	MEDIUM UNCERTAINTY	HIGH UNCERTAINTY
LOW GAIN (slow response)	0 5 15 30 ——→ predicted trend Sig. 1%	0 5 15 30 ——→ = Sig. 5%	0 5 15 30 ——→ = Sig. 1%
MEDIUM GAIN (moderate response)	0 5 15 30 ——→←—— Sig. 0.1%	0 5 15 30 ——→←—— Sig. 5%	0 5 15 30 ——→←—— Sig. 1%
HIGH GAIN (fast response)	0 5 15 30 ——→ = Sig. 0.1%	0 5 15 30 ——→←—— Sig. 0.1%	0 5 15 30 ——→ = ←—— Sig. 0.1%

Test for trends amongst prediction spans for Perfect predictor model data

It can be seen that the 'best' prediction span, i.e. that giving the lowest error scores, decreased as uncertainty increased. This effect was most noticeable for the High gain (fast response) condition (Figure 26) when the optimum span decreased from approximately 23 seconds, to 15 seconds, then to approximately 10 seconds as the level of input disturbance rose. It would seem that operators cannot make use of as much of the Perfect predictor trace due to the long-distance predictive information being rendered inaccurate by input uncertainty.

Given these interactions, it is clear to see why previous workers, as reviewed in Chapter 2, have come up with conflicting findings concerning optimum prediction spans. Contrasting the Perfect predictor scores with the Taylor series extrapolation model; only in the High gain condition was there any indication that a reduction in usable span to approximately 15 seconds occurred, but as Figure 26 suggests this effect was nowhere near as significant (Page's L significant at 5%) as for the corresponding Perfect predictor condition (Page's L significant at 0.1%).

4.4 Control Histograms

Inspection of the control histograms (Figure 27 gives examples) for each trial suggests that distinct patterns of control were present for the two prediction models, though of course variations did occur across subjects. Typically, control without any form of predictor was characterised by use of the extreme limits of control in 'bang-bang' fashion. With the introduction of a 5 second Taylor series trace control was still characterised by long periods spent at the extremes, but there was an additional distribution at the centre of the range corresponding to finer control adjustments. This central distribution typically spread out towards the extremes as prediction span was extended to 30 seconds.

In the case of the Perfect predictor model, control was characterised by a much smoother gaussian-type distribution. Though for short prediction spans some time was spent at the extreme limits of the controls, this component disappeared as prediction span was increased beyond 5 seconds, and control then consisted of very fine adjustments around the centre of the range. In other words, the faster response of the Perfect predictor model made immediately obvious the

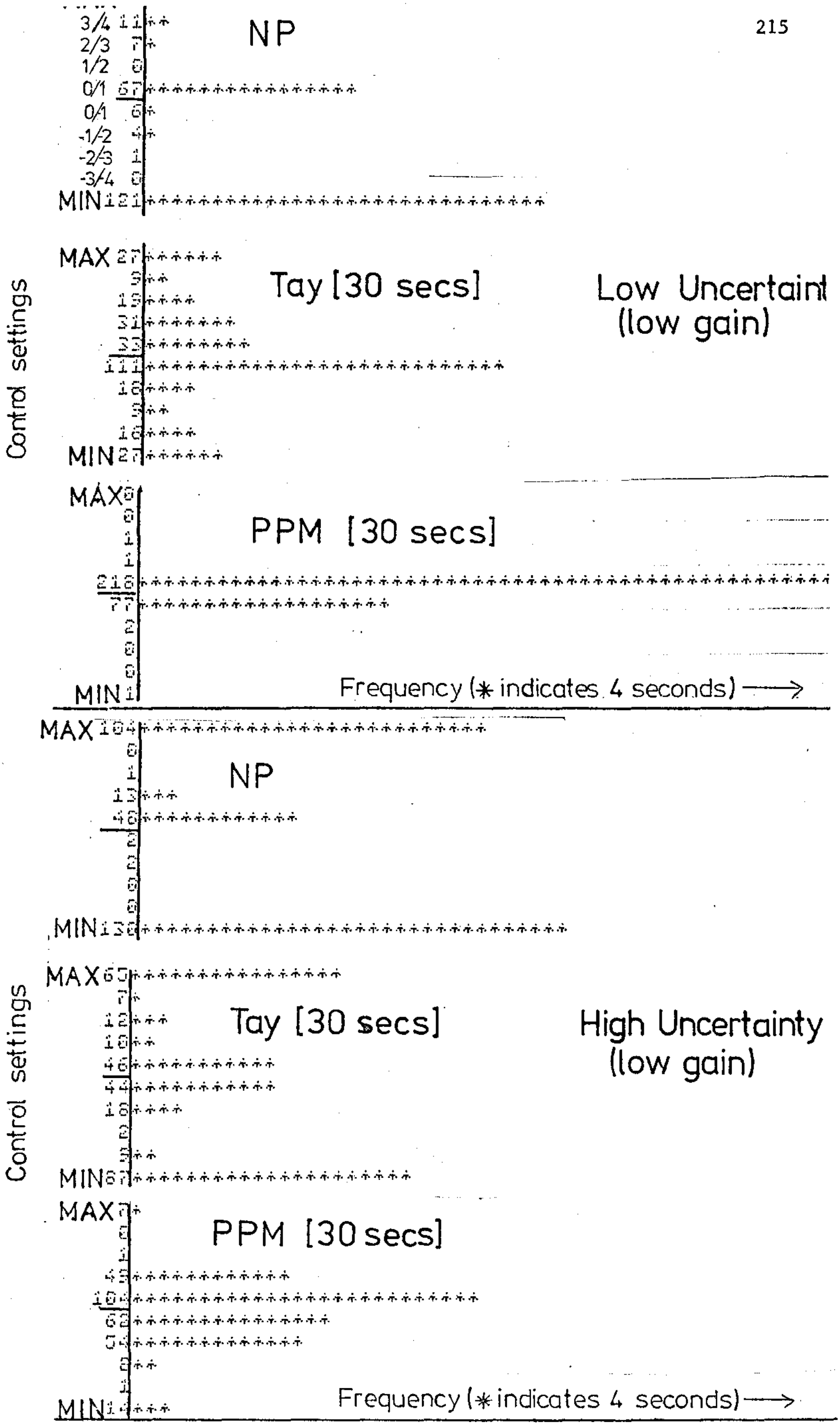


Figure 27: Control histograms - number of seconds during which control is within specified range settings

effect of a control action and resulted in a smoother pattern of control, whereas the Taylor series rate-of-change predictor required time for a control change to affect it and so control, even if smooth at the start of a trial, frequently ended up at the extremes as the system became progressively unstable.

The control histograms provided a useful extra indication of performance. In the Low gain condition, for example, although error scores were much the same for the two prediction models, scrutiny of the histograms suggests that with the Taylor series model this was achieved by considerably greater control effort. The Perfect predictor gave much smoother control than did the Taylor series equivalent, though it must be stressed that either predictor was preferable to none at all.

The effect of introducing input disturbances to the system was to spread the distribution of control actions towards the extremes, and a similar effect was found when increasing plant gain.

4.5 Display switching

Analysis of the display switching data also revealed some interesting variations in strategy. Display switching appeared to be most regular when the pointers were within the prescribed limits and under control, and appeared to increase as control of the pointer(s) was progressively lost: this was most likely to occur for short prediction spans and high gains. In addition, the bias in the time spent looking at the left- or right-hand meter shifted during the course of a trial, more time being spent attending to the pointer most out of control. In the stable equilibrium state, a switching rate of once every few seconds was not atypical. Unfortunately, however, switching rates faster than once a second were not capable of being recorded by the system, and this has precluded a more detailed analysis.

4.6 Subjective comments and questionnaire analysis

In general, analysis of subjective comments on the task tended to confirm the impressions gained from the objective analyses. Subjects all stated that they preferred using a predictor trace to control without a predictor, and this was reflected as the 'with predictor' trials being rated as easier to control. Opinions were divided for preference between the Taylor series model and the Perfect predictor model. Seven out of the fifteen subjects preferred the Perfect predictor trace in that it was a lot more accurate and gave immediate feedback as to the consequences of control changes. Five subjects preferred the Taylor series trace because its slow, rate-of-change response was easier to follow and gave more time for them to respond. Three subjects failed to detect any difference between the two prediction models. In all, the Perfect predictor model was rated as being more useful than its Taylor series counterpart for a given prediction span. Subjects also rated their control actions as being considerably smoother using the Perfect predictor trace, particularly with longer prediction spans.

On the question of prediction spans, subjects were equally divided in their preference for the longest possible prediction span (30 seconds) or a shorter span (e.g. 15 seconds). The 5 second span was universally disliked, but thought just possible to control with. On their ratings as to the usefulness of the predictors, subjects rated the 30 second span as being most useful in the Low gain condition but the 15 second span as being most useful in the faster responding Medium and High gain conditions. Most subjects failed to detect any variations in the level of input uncertainty, though several commented that the pointer appeared to disobey the controls or to move about of its own accord in some of the trials. It was thought more difficult to control these trials (High uncertainty), particularly in the High gain condition using long prediction spans based on the Perfect predictor model, as the variations due to uncertainty in the middle-to-end part of the trace were found misleading. It is interesting to note that with only one exception all subjects reported using the end segment of the trace for control - this finding is clearly at odds with reports from the Dunlap labs. mentioned earlier in Chapter 2. Smith and Kennedy (1975) had noted that their subjects used the first or central segment of a trace in order to effectively minimise the time to reach their desired trajectory. Clearly subjects make full use of the trace they are given.

It is apparent from analysis of the strategies reported by the subjects that anticipation of the pointers' movement played a vital part in control, especially for the No Predictor condition and to a lesser degree in the Taylor series condition. Use of a perfectly accurate prediction model effectively eliminated the need to anticipate the pointers' trajectory. In the No Predictor conditions, subjects followed the strategy of moving the controls to their extreme to compensate as soon as any perceived movement of the pointer was detected. With experience some subjects tried to anticipate the pointers' point of turn and to gradually reduce their control input beforehand. Again, with experience of the system dynamics (probably gained from the 'with predictor' conditions) some subjects restricted their range of control actions so as not to use the extreme positions, and made their actions consciously smoother.

As found in previous chapters, subjects varied widely in their ability to verbalise their control processes. One subject (a physicist) reported controlling on the theory that the system dynamics were analogous to simple harmonic motion, another evolved a yo-yo model of the process, yet another claimed the dynamics were equivalent to an inverted pendulum. It is clear from this admittedly anecdotal evidence that subjects' anticipations were based on some crude form of internal model of the process, through which predictions could be made in the absence of a computer-provided prediction. Using the Taylor series model predictor subjects frequently used the slope of the trace as the main criterion for the amount of compensation which they applied. Some anticipation was still required however, and the problem became one of keeping the predictor trace horizontal within the prescribed limits and with the pointer stationary. In the Perfect predictor model conditions, the problem was further simplified and became one of watching the end of the trace and compensating to keep it within limits and as near to the 50 mark as possible. A slightly different policy was adopted if a pointer drifted outside the limits - the object then became to get that pointer back within limits as quickly as possible, if possible keeping the end point of the trace between the limits as the pointer approached them. Protocol evidence from this study will be discussed further in Chapter 9.

5. CONCLUSIONS

A comprehensive design study has been carried out to investigate how variations in predictive display parameters and task characteristics affect operator performance in a generalised dual-meter monitoring and control task. Predictive displays were found to bring about an improvement in average time outside limits scores in all the experimental conditions. For systems with a slow speed of response, there was little to choose between a highly accurate and a relatively unsophisticate prediction model, given that adequate performance with the latter was achieved at the expense of greater control effort. For systems having moderate to fast response times, the more sophisticated prediction model was justified.

Recommendations can be made regarding the choice of an appropriate prediction span for simple and sophisticated prediction models under various levels of plant responsiveness and input uncertainty. With simple prediction models, which seemed relatively immune to uncertainty, prediction span was affected by plant gain. For systems with low to medium gains, the maxim "the longer the better" was appropriate, with perhaps a slight reduction in usable prediction span for high gain systems With a hypothetical Perfect prediction model, the optimum prediction span was reduced with the combined effect of increasing uncertainty level and increasing plant gain. Conflicting results of previous workers are explained in terms of the differing gains, levels of uncertainty and prediction models of the systems investigated.

Reported strategies from subjects in the present study suggest that the formation of a crude form of internal predictive model is an important part of unaided control. Subjects used the full extent of the predictor trace presented to them.

CHAPTER 7

AN EXPERIMENT TO VALIDATE THE USE OF PREDICTIVE TECHNIQUES IN THE CONTROL OF A PART-SIMULATED, SEMI-BATCH CHEMICAL REACTOR IN AN INDUSTRIAL SETTING.

1. OBJECT OF THE EXPERIMENT

Previous chapters have shown the potential benefits to be gained from the introduction of predictive techniques in the control of laboratory simulated continuous chemical processes. Chapter 6 in particular has investigated the characteristics that such displays should possess, and has shown that even a relatively unsophisticated prediction model can assist in the control of a process having a moderately slow speed of response.

The present experiment aims to validate the use of a simple prediction model, formulated as a multi-pen predictive recorder, in the control of a part-simulated, semi-batch chemical reactor employing real plant and experienced operators. The study set out with the object of testing the effect of the predictive display on temperature control, but the original brief was later widened to include pH prediction at the operators' request.

2. METHOD

A suitable plant was sought on which to test the feasibility of predictor displays in a field setting. The Control Division at Warren Spring Laboratory had developed a part-simulated, semi-batch chemical reactor (the 'batch kettle') initially to provide data for process operation research, and kindly agreed to provide collaborative facilities. Since the batch kettle not only comprised real plant, but was manned by experienced process operators, it was felt to provide a good opportunity to test the suitability of predictive aids in a realistic industrial setting.

2.1 The 'batch kettle' plant

The operation of the kettle is documented elsewhere (Cininas, 1975a; King, 1975; King and Cininas, 1978) so only a brief description is given here. Illustrations of the batch kettle plant and its control room can be found in Appendices 7.1-7.2. The simulated process (Figure 28) produced an acid product following an exothermic chemical reaction. To achieve this the 90 gallon kettle was first charged with 'feed', 'reactant' was added, and then the chemical reaction began. In order to avoid the cost of purchasing and disposing of chemicals and yet retain the semblance of a real plant, the process was partially simulated - water was used to represent the three chemical reagents and the reaction itself was computer modelled. The operators, however, treated the kettle as though it were a real process.

An Argus 500 computer measured the quantities of reagents added and calculated the amount of product formed and the heat of reaction, which in turn determined the control setting of a second 'reaction steam heating' coil, resulting in an increase in the kettle temperature. In addition the Argus calculated a pH change which had to be continuously neutralised by the manually controlled addition of 'caustic', again using a simulated relationship. The pH value was displayed in the control room depicted in Figure 29, together with indications of temperature, reagent and cooling flows and a schematic mimic diagram (Figure 30).

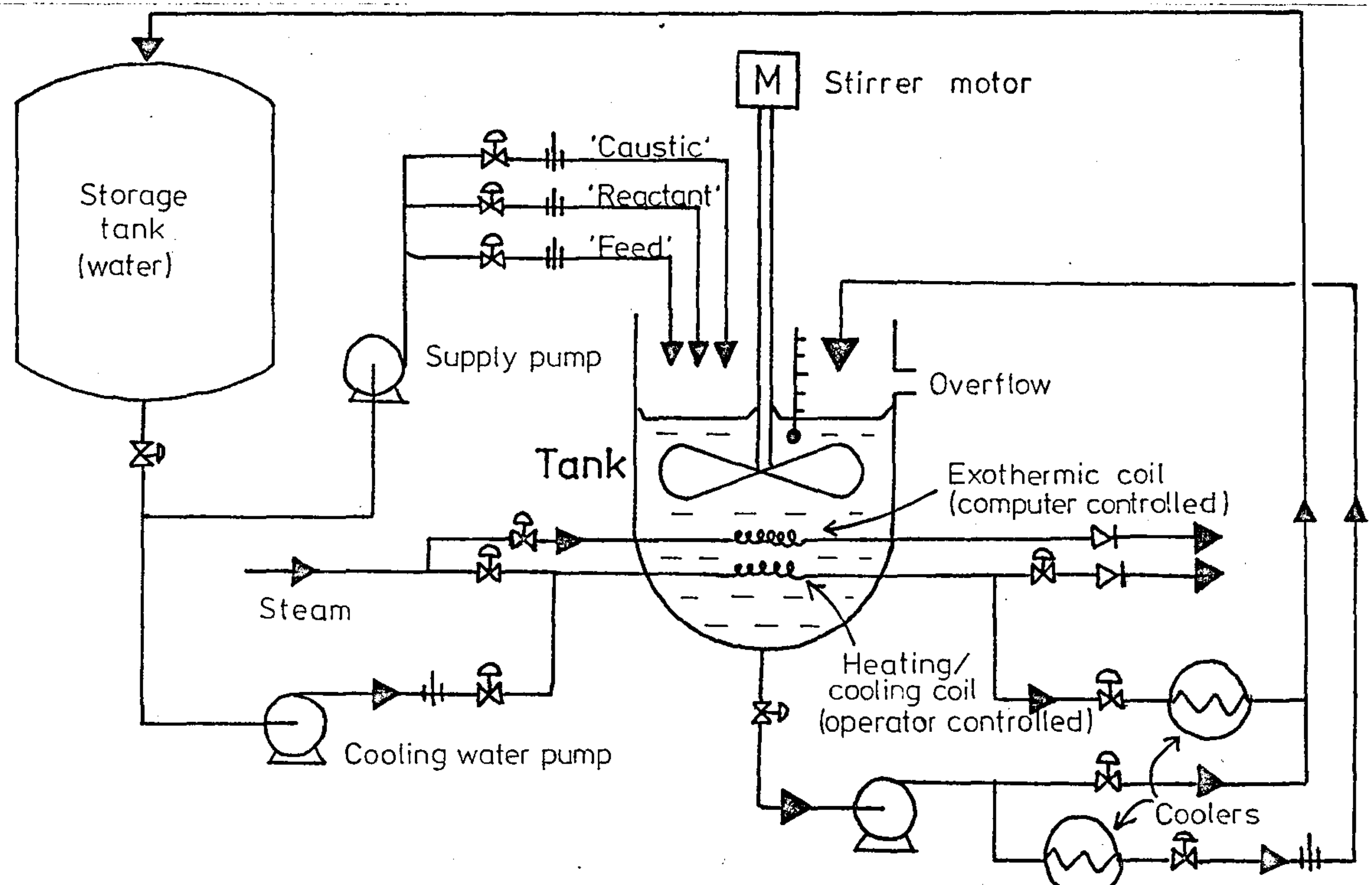


Figure 28: The part-simulated, semi-batch chemical reactor ('batch kettle')

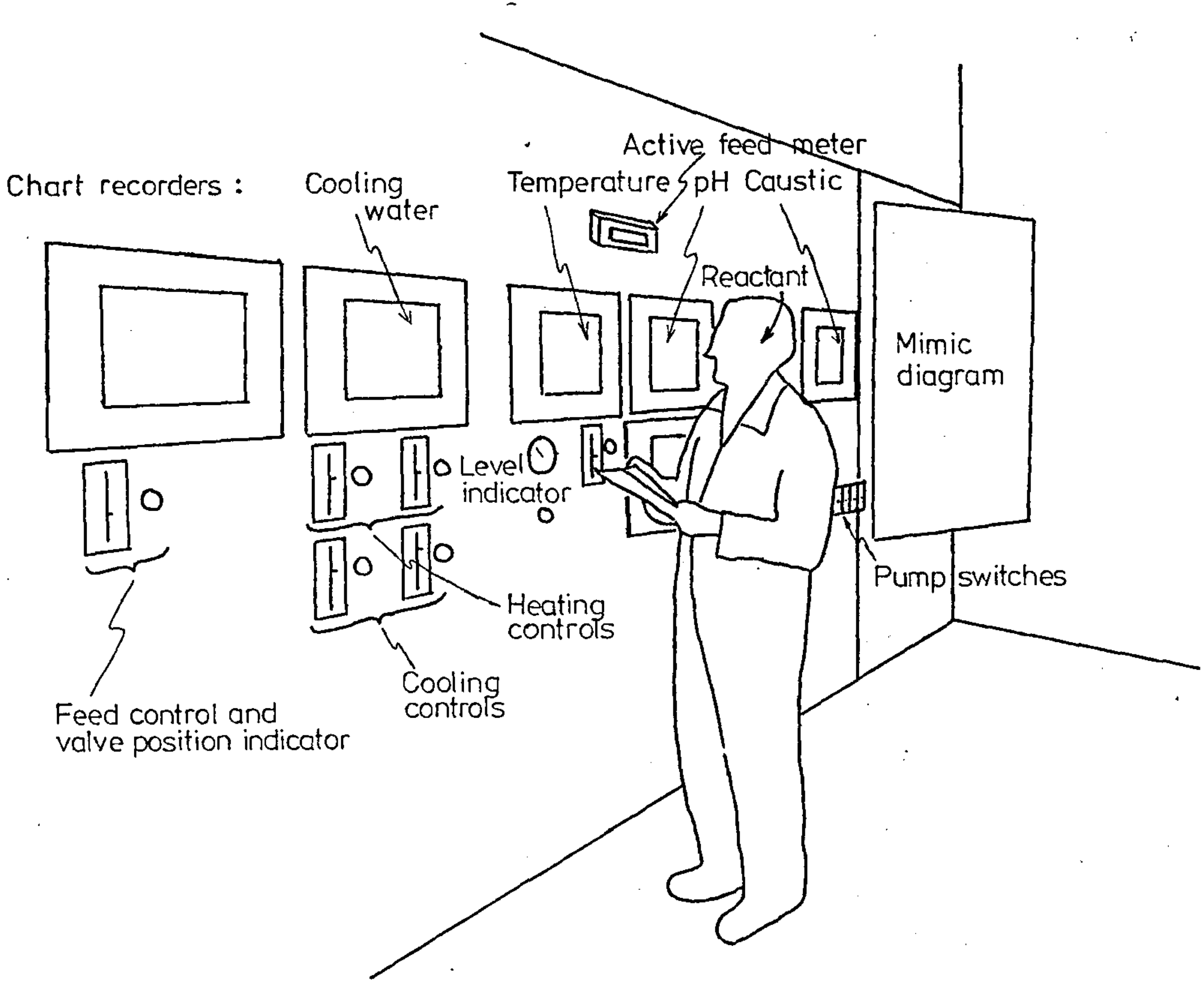


Figure 29: Batch kettle control panel

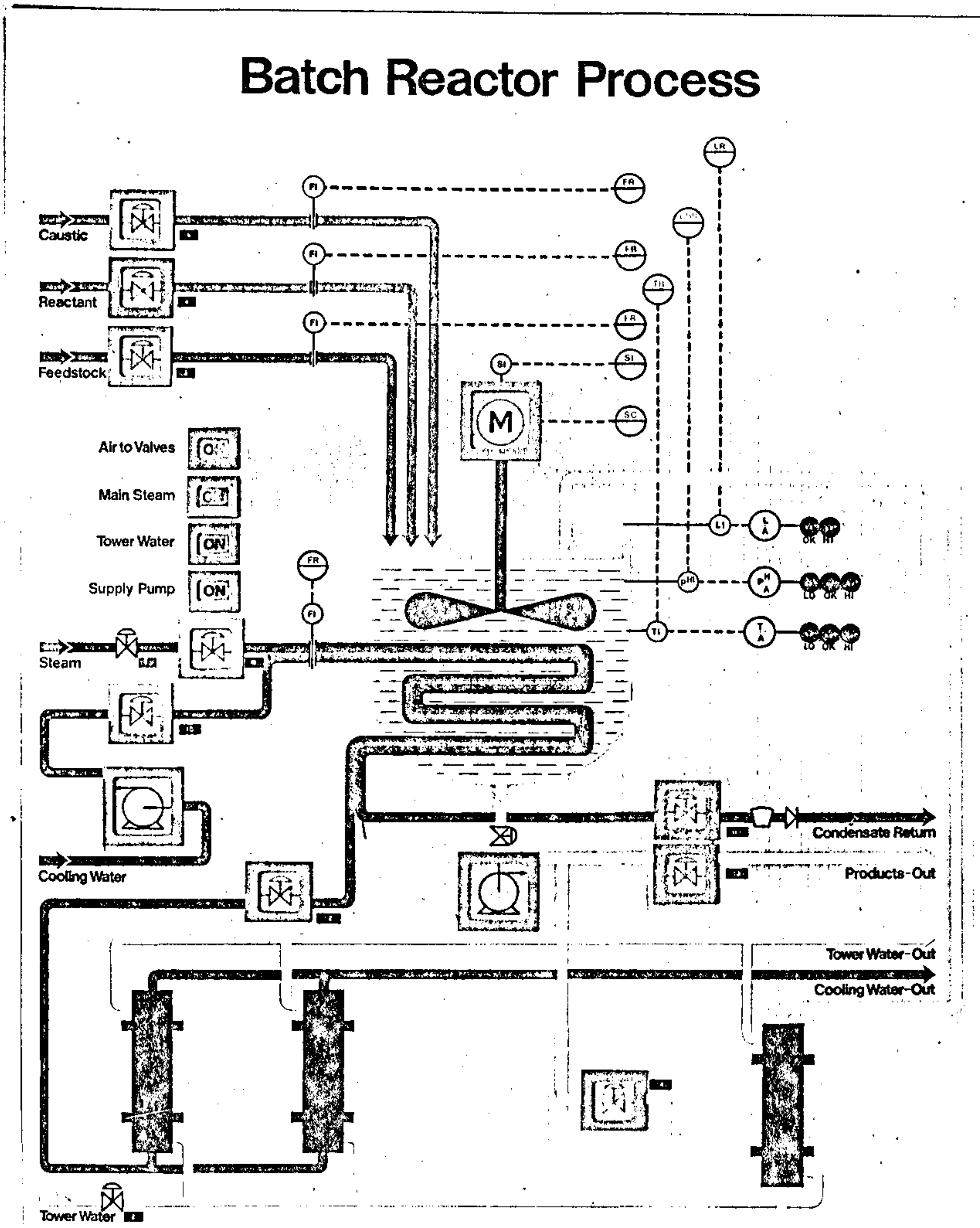


Figure 30: Mimic diagram of batch reactor process displayed in the control room (Photograph by courtesy of Warren Spring Laboratory).

The process was operated as follows:-

1) An initial charge of 'feed' was run into the reactor and the contents heated to 55°C.
2) 'Reactant' was then added to the vessel under the process operator's control, causing temperature and pH to vary. The recirculation and cooling controls were used to maintain the temperature at 60°C $\pm$ 5°C. 'Caustic' flow was used to neutralise the acid formed so that the pH remained in the range 7 $\pm$ 1 unit. The reactant flow could also be used to control pH.
3) When the reaction was complete the vessel was heated to 75°C for a finishing period of 5 minutes. The contents of the reactor were then cooled to 50°C and the vessel was emptied.

All process measurements and the plant operator's control actions were logged by the computer at 10 second intervals for subsequent analysis.

2.2 Process operator research at Warren Spring

A considerable body of knowledge had already been accumulated concerning the operation of the kettle, in terms of such standard measures as percentage conversion of reagents to finished product and the more sensitive measure of profit per batch (Cininas, 1975b), manual control strategies (King and Cininas, 1976) and verbal protocol evidence (Cininas, 1976). It had become clear that the batch kettle was of sufficient complexity to be representative of the problems encountered in chemical process control. Cininas (1975b) for example found that whereas most operators could achieve good percentage conversion figures, there was a large variability between operators in their profit scores, although he did not attempt any detailed statistical analysis. The operators aimed only to optimise percentage conversion and profit figures. Whilst a good percentage conversion implied good temperature and pH control, several interacting factors contributed to profit (e.g. cost of feed, length of run, amount of product produced). It was thus difficult for an operator to understand what contributed to profit and therefore how to control it. Warren Spring had already been experimenting with a rudimentary form of decision aid, in the form of a computer-driven volume display ('Active feed meter') to replace the original tank level indicator. The latter was unreliable due to agitation of the kettle contents by the stirrer. A computer-driven volume display not only provided a reliable level indication, but also gave an indication of when the reaction phase was over, and according to Cininas (op.cit.) led to higher profit figures.

King and Cininas (1976) further report that although manual control policy was vitally important during the reaction phase, there was little evidence amongst the batch kettle operators of 'feed-forward' (predictive) control, due to the difficulties they experienced in understanding the interaction effects of the process. The operators thus worked on non-optimal rules-of-thumb for their control. This suggested to the present author that the provision of a simple predictive facility might improve operator control policy and thus lead to higher production figures. Cininas (1976) also comments on the need for some external aid, particularly for pH control, based on verbal protocol evidence.

2.3 Multipen Predictive Recorder (m.p.r.)

An existing pen recorder (Watanabe type MC611-S4H) with provision for multiple pens was modified to serve as the m.p.r. by fitting extension arms to the existing pens (Shackel, Goillau and Laios, 1976). The final arrangement is detailed in Figure 31. The lowest pen provided a standard current time trace, and the upper 3 pens gave indications of computer predicted values at 10, 20 and 30 seconds respectively into the future. This arrangement gave discontinuous predicted paths in essence similar to those on the CRT based displays of Chapters 5 and 6, but displayed in a cheaper way and in accordance with current practice in process control using pen recorders.

The predicted values were derived from a Taylor series subroutine running within the suite of control programs on the Argus 500. The subroutine stored past values of smoothed temperature or pH and calculated derivatives, smoothing being achieved using the moving average technique. A listing of the Taylor series predictive subroutine is given in Appendix 7.3. Though the subroutine was capable of extrapolating on the basis of any number of derivative terms, due to the problems encountered with noise amplification a single term (i.e. straight line predictor) was used in practice. Predicted values were updated every 10 seconds and integrated over 1 second intervals to give a smoother response. In its operational form, the m.p.r. was trolley mounted directly below the temperature or pH panel recorder which it replaced (see Figure 29).

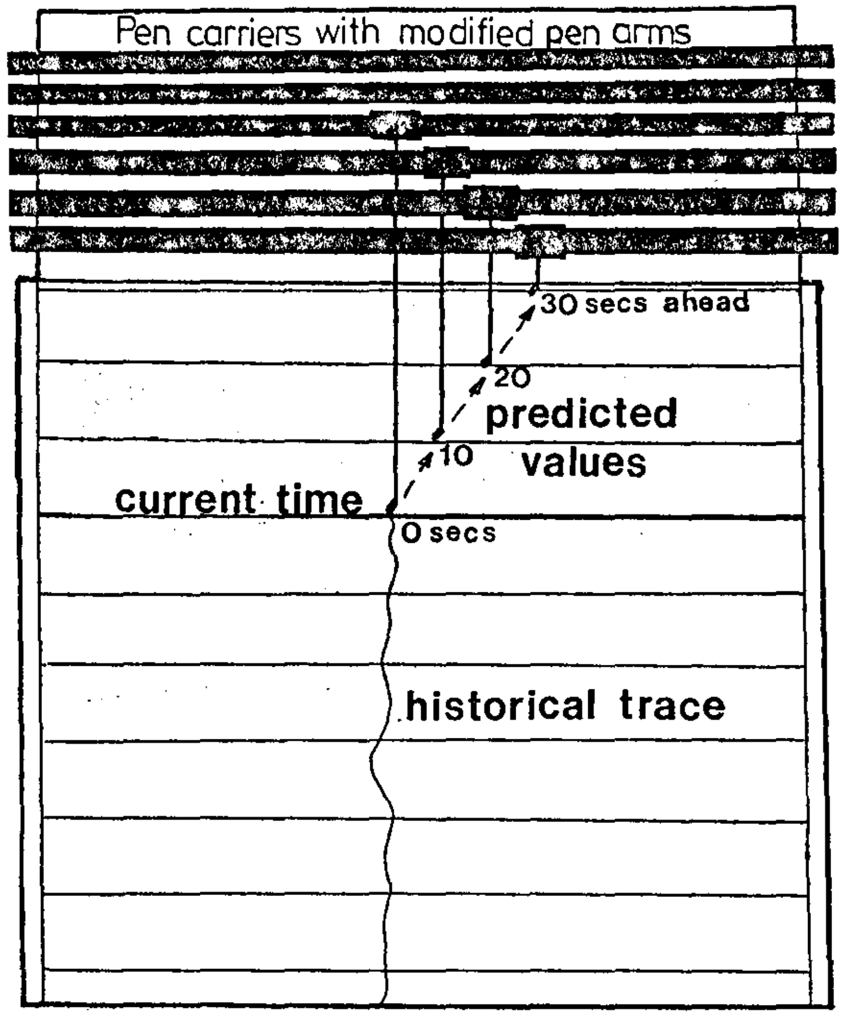


Figure 31: The multipen predictive recorder (m.p.r.)

In the example shown, the predictive pens show a trend away from the steady state towards an extreme of, say, temperature. This trend is not shown by the conventional historical trace, but as amplified by the predictive pens the operator can quickly decide whether the excursion is outside his prescribed limits of control and take the necessary corrective action. A steady state on this recorder is reached when the pens are vertically in line one above the other.

2.4 Experimental Design

Owing to the limited number of operators available and the need to use their time optimally, a balanced repeated measures design was employed. A preliminary analysis had already shown temperature and pH to be the critical variables in the control of the reactor, and these factors were treated separately. Four operators were assigned to control the reactor with prediction on temperature, and four with prediction on pH. The latter condition was included at the operators' suggestion. For both studies (temperature and pH), two subjects encountered the 'with predictor' condition first and two the 'without predictor' condition first. The design may be represented as follows for both temperature and pH:

	no predictor	m.p.r.
no predictor first	G1	G1
m.p.r. first	G2	G2

where G1 and G2 are independent groups of 2 operators, differing only in whether they received the m.p.r. condition as the first or second block of experimental trials.

2.5 Procedure

An initial period was necessary to 'bed in' the recorder and Taylor series subroutine. Experimental trials with the new recorder were slotted into the casual shift operator system employed at WSL for control of the kettle and in connection with other projects.

For each new operator, the nature of the m.p.r. and the purpose of the predictive pens was explained verbally. The experimental trials were then run. Operator comments were gleaned informally at the end of the experimental trials. The length of each run was such that a maximum of four runs could usually be obtained on a given day, two runs in the morning and two in the afternoon. It was customary for operators to take a break between runs.

It had been hoped that each operator would yield results from approximately five with- and five without- predictor runs, but with equipment problems and constraints of the shift system this proved to be impractical. However, in all cases averages over two to six valid runs were obtained.

2.6 Subjects

The batch reactor plant was operated by WSL personnel on a shift basis. All operators had had considerable experience in the control of the kettle, some since its commissioning in June 1974.

2.7 Data Collection

The Argus 500 computer maintained a process log of all variables and process states at 10 second intervals. A computer printout summarising the operator's performance was generated from each run. Figure 32 shows a typical example. An abbreviated version of this printout was output on a teletype in the operator's control room as feedback to him, and a keen sense of competition resulted amongst the operators. In addition a fuller breakdown of each trial by 10 second intervals was available on paper tape.

```
BATCH REACTOR RUN 10. 3      29/ 7/75

BATCH TIME   1H30M
PRODUCT VOL GALS          5.386      CONVERSION PERCENT          99.1
EXCESS FEED GALS       0.00     EXCESS REACTANT GALS      0.11
CAUSTIC VOL GALS       5.91
ABS ERROR ,TEMP=          0.700   ,PH=          4.140
PERFORMANCE INDICES

TEMPERATURE

SUM OF ERROR  - SOD=          0.290  - ABS=          0.700
MEAN ERROR    - SOD=          0.001  - ABS=          0.002
VALUE         - MEAN=        61.830  - S D=          0.885
NO OF RESULTS=    545


PH

SUM OF ERROR  - SOD=          2.606  - ABS=          4.140
MEAN ERROR    - SOD=          0.008  - ABS=          0.013
VALUE         - MEAN=         6.602  - S D=          0.411

PROCESS COSTS POUNDS


    FEED     REACT    CAUST    STEAM    WATER    PLANT    REPROC   PRODUCT
    COSTS----------------------------------------------------------VALUE
    29.7     59.5     11.8     14.0     11.2     15.0      0.0    323.7

RUN COST=        141.3REPROCESSING COST=          0.0

PRODUCT VALUE=       323.  PROFIT=       182.4
```

<u>Figure 32:</u> <u>Typical results printout at the end of a trial summarising operator performance.</u>

3. RESULTS AND STATISTICS

From the available performance measures listed in the printout of Figure 32, three representative indices were chosen on the advice of Warren Spring personnel. Mean performance scores were thus obtained in respect of:

a) Percentage conversion of reagents to product.

b) Standard deviation about the predicted variable (temperature or pH).

c) Calculated profit in £.

These measures are plotted in Figure 33 (percentage conversion), Figure 34 (Standard deviation), and Figure 35 (Calculated profit) for both temperature and pH prediction. An average over the four operators is also given. Original data scores can be found in Appendices 7.4-7.5.

By taking averages to equalise the sample sizes, it was possible to perform a statistical analysis of the data. Summary ANOVA tables are given in Tables 9, 10 and 11, showing the effects of the m.p.r. and presentation order.

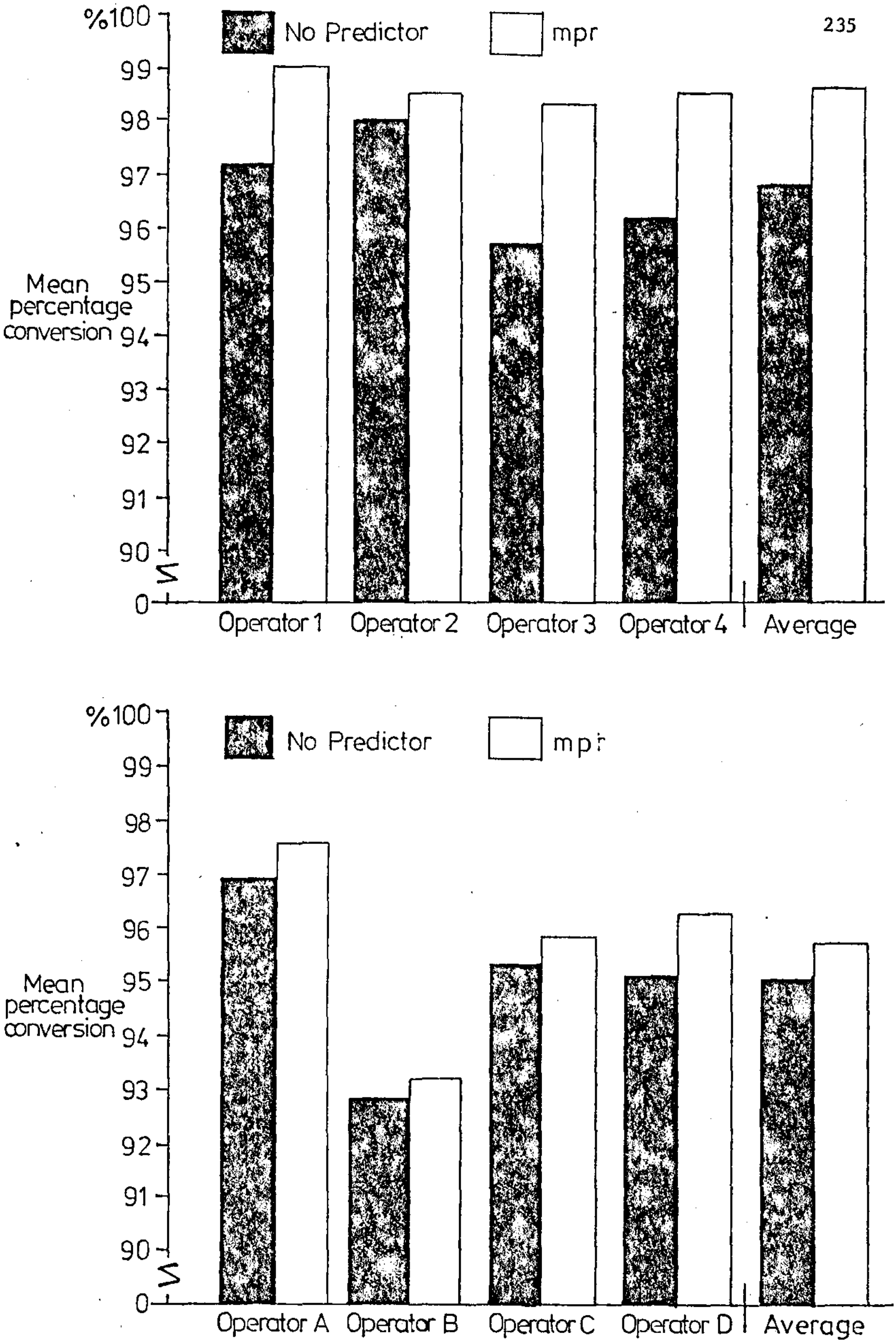


Figure 33: Effect of m.p.r. on percentage conversion of reagents to product: a) prediction on temperature (top)
b) prediction on pH (bottom)

TABLE 9: Summary ANOVA for percentage conversion scores

a) Prediction on temperature

Source	Sum of Squares	df	Variance estimate	'F'	Significance level
Between Subjects					
Order	2.030	1	2.030	30.136	5% (df 1,2)
Subjects within groups	0.135	2	0.067		
Within Subjects					
Prediction (m.p.r.)	6.534	1	6.534	29.652	5% (df 1,2)
Prediction x Order	0.865	1	0.865	3.924	- (df 1,2)
Prediction x S.w.g.	0.441	2	0.220		

b) Prediction on pH

Source	Sum of Squares	df	Variance estimate	'F'	Significance level
Between Subjects					
Order	0.505	1	0.505	.056	- (df 1,2)
Subjects within groups	17.993	2	8.997		
Within Subjects					
Prediction (m.p.r.)	0.865	1	0.865	12.855	- (df 1,2)
Prediction x Order	0.056	1	0.056	0.834	- (df 1,2)
Prediction x S.w.g.	0.135	2	0.067		

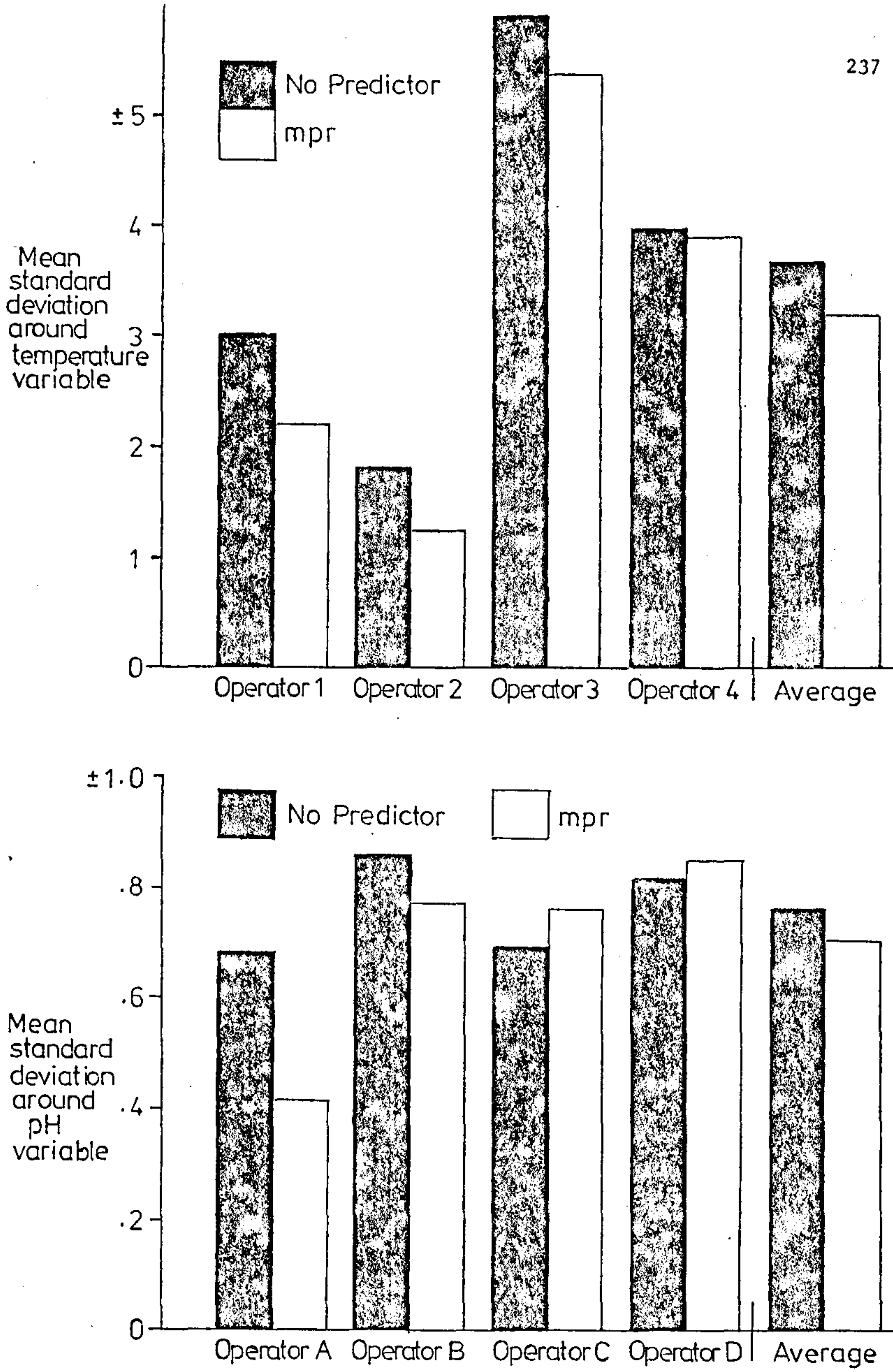


Figure 34: Effect of m.p.r. on standard deviation around predicted variable: a) prediction on temperature (top)
b) prediction on pH (bottom)

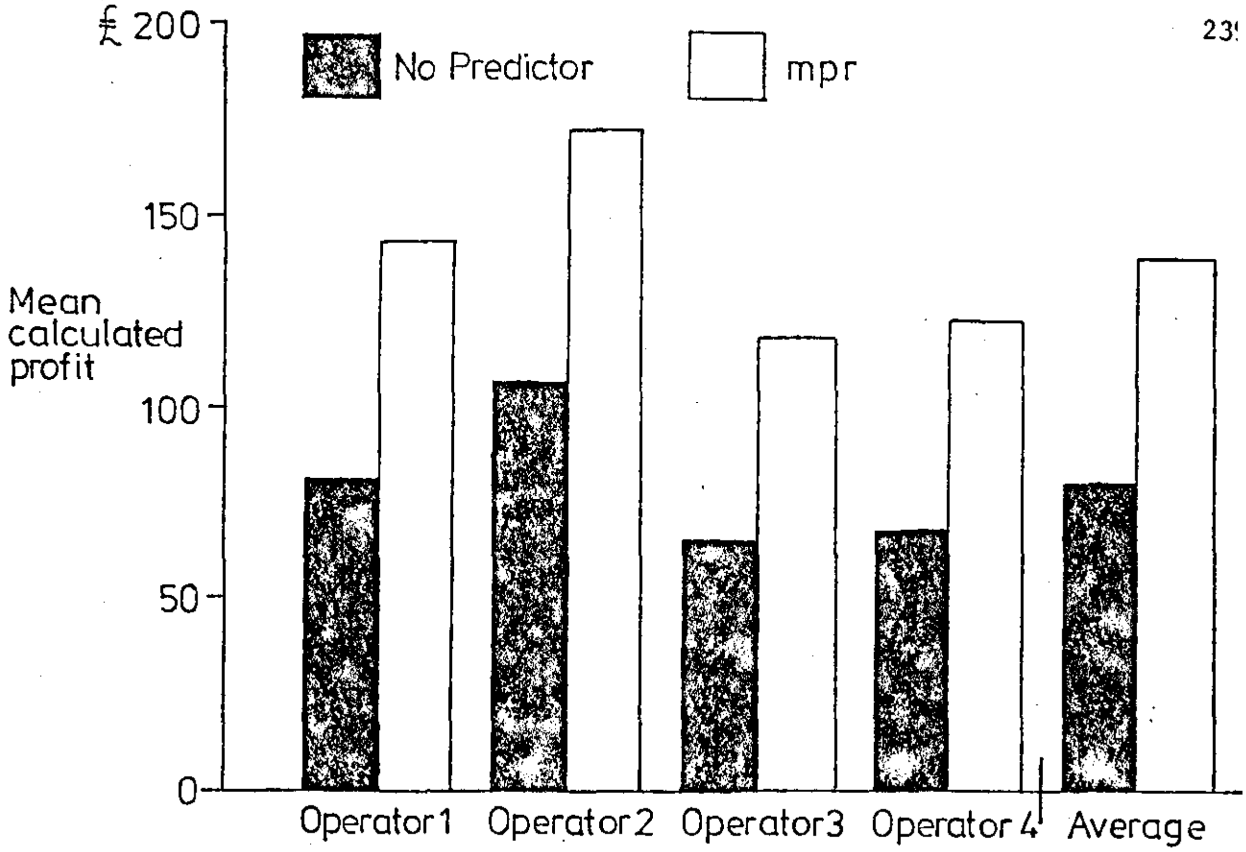
TABLE 10: Summary ANOVA for standard deviation around the predicted variable

a) Prediction on temperature

Source	Sum of Squares	df	Variance estimate	'F'	Significance level
Between Subjects					
Order	14.751	1	14.751	7.162	- (df 1,2)
Subjects within groups	4.119	2	2.059		
Within subjects					
Prediction (m.p.r.)	0.514	1	0.514	15.168	- (df 1,2)
Prediction x Order	0.069	1	0.069	2.06	- (df 1,2)
Prediction x S.w.g.	0.068	2	0.034		

b) Prediction on pH

Source	Sum of Squares	df	Variance estimate	'F'	Significance level
Between Subjects					
Order	0.0183	1	0.0183	0.45	- (df 1,2)
Subjects within groups	0.0816	2	0.0408		
Within Subjects					
Prediction (m.p.r.)	0.0083	1	0.0083	1.923	- (df 1,2)
Prediction x Order	0.0273	1	0.0273	6.35	- (df 1,2)
Prediction x S.w.g.	0.0086	2	0.0043		



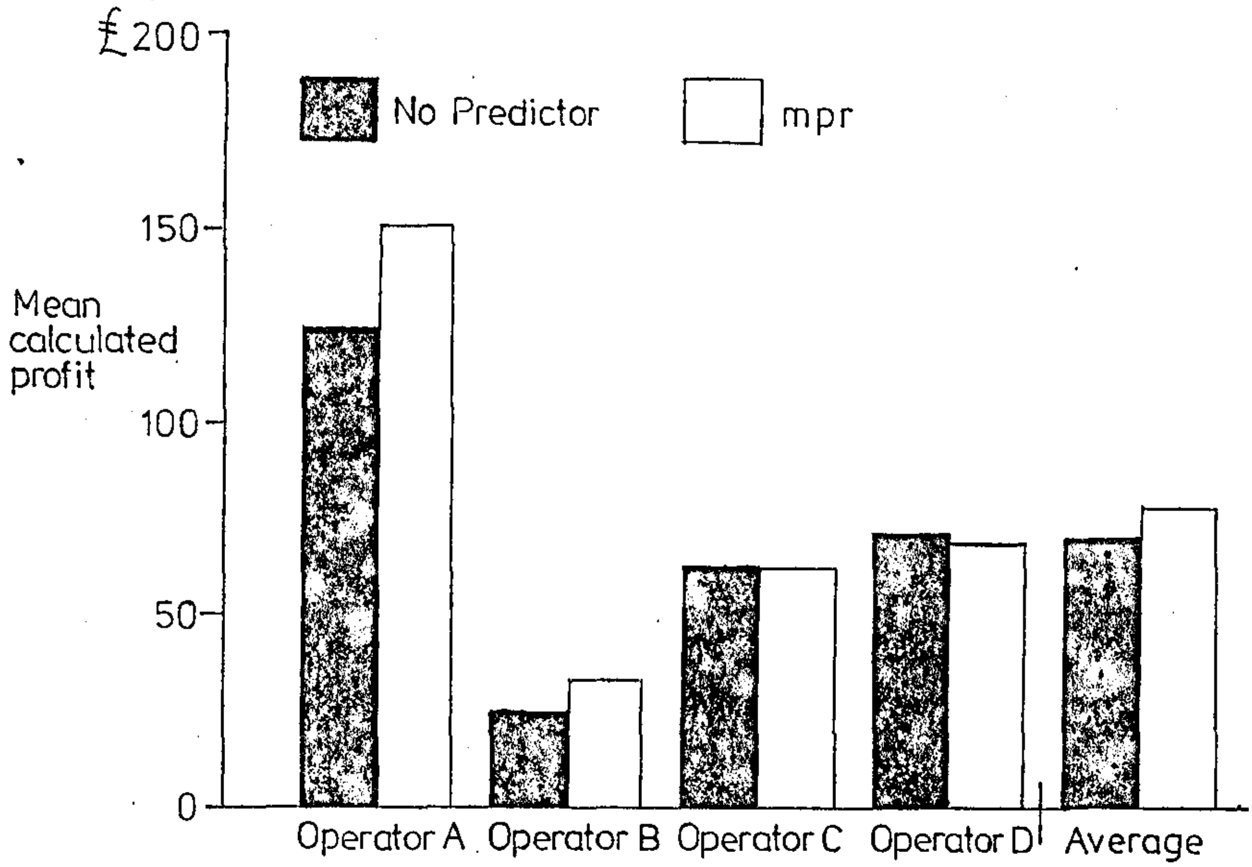


Figure 35: Effect of m.p.r. on calculated profit:
a) prediction on temperature (top)
b) prediction on pH (bottom)

TABLE 11: Summary ANOVA for calculated profit in £

a) Prediction on temperature

Source	Sum of Squares	df	Variance estimate	'F'	Significance level
Between Subjects					
Order	2,067.25	1	2,067.25	5.17	- (df 1,2)
Subjects within groups	799.21	2	399.61		
Within Subjects					
Prediction (m.p.r.)	6,809.45	1	6,809.45	3,902.26	0.1% (df 1,2)
Prediction x Order	51.01	1	51.01	29.23	5% (df 1,2)
Prediction x S.w.g.	3.49	2	1.745		

b) Prediction on pH

Source	Sum of Squares	df	Variance estimate	'F'	Significance level
Between Subjects					
Order	568.688	1	568.688	.095	- (df 1,2)
Subjects within groups	11,989.845	2	5,994.9225		
Within Subjects					
Prediction (m.p.r.)	137.365	1	137.365	2.894	- (df 1,2)
Prediction x Order	173.446	1	173.446	3.654	- (df 1,2)
Prediction x S.w.g.	94.946	2	47.473		

4. DISCUSSION

4.1 Percentage Conversion

Figure 33 (a,b) shows that in both temperature and pH cases control of the kettle using the predictive pens gave an average increase in percentage conversion. The improvement, though small, is consistent with what one would expect from operators with considerable experience of controlling the process and already achieving good percentage conversion figures.

The difference was apparently more marked in the case of temperature prediction, and this was borne out by the ANOVA (Table 9a). Here the effect of temperature prediction was just significant at the 5% level, though coupled with a similarly significant effect due to the order of testing (the improvement due to the predictive facility being greater when this condition was encountered first). In contrast, for pH prediction the ANOVA indicated that none of the experimental effects were statistically significant (Table 9b).

4.2 Standard deviation around the predicted variable

Figure 34a shows a consistently smaller standard deviation around the temperature mean with the m.p.r. set to predict on temperature. However, the effect of temperature prediction just missed statistical significance in the ANOVA (Table 10a), again probably due to the low numerator and denominator of the F ratio resulting from the small sample size.

Figure 34b, however, is more equivocal. In only two cases (Operators A and B) was the standard deviation around the mean pH smaller using the predictive pens, and in both these cases the predictive facility was encountered second. Perhaps not too much should be read into this finding as the standard deviation around pH mean was already quite small when compared with the temperature equivalent. As might be expected, no statistically significant difference was indicated by the ANOVA (Table 10b).

4.3 Calculated Profit

Profit is perhaps the most critical index, being a system measure and the criterion which any aid must demonstrably satisfy if it is to be adopted by a sometimes sceptical management.

Figure 35a indicates a large and consistent improvement in profit figures using the predictive facility for temperature prediction. The average profit per batch increased from £80 to £138 using the m.p.r., and by this reckoning the aid would pay for itself within months. The ANOVA (Table 11a) further showed the effect of temperature prediction on profit figures to be significant at beyond the 0.1% level, though coupled with a slight interaction term due to presentation order.

Prediction on pH (Figure 35b) showed a less consistent improvement, though the average profit per batch increased from £70 to £78 with pH prediction. However, this effect did not achieve statistical significance, again perhaps due to the small sample size.

In all it would seem that using the m.p.r.'s predictive facility was indeed beneficial to control, particularly in the case of temperature prediction. An explanation of the discrepancy between the temperature and pH cases is that the temperature response over time tended to be inherently smoother, and so the relatively crude Taylor series predictor model could cope better with temperature compared to the faster moving pH (the latter may have warranted a more sophisticated, e.g. stochastic, predictor).

It is interesting to note that for pH prediction, operators C and D obtained higher percentage conversion scores but higher standard deviation and lower profit figures when using the m.p.r. This highlights the fact that the profit figure was a function of several interacting factors, and depended not only on percentage conversion but on costs for process time, reagents used, steam and water consumption, and a value of product which depended on the quantity produced and its saleable value. (Details of how the profit figure was arrived at are given in Appendix 7.6.) This example emphasises the importance of using realistic criteria in system evaluation.

4.4 Operator Comments

The operators' opinions of the m.p.r. were elicited at the end of the trials. One operator mentioned that the new-style recorder was an improvement on the previous panel recorder, though he felt that the scale was really too large to be meaningful. It is convenient to consider operator comments separately for the cases of temperature and pH prediction.

In the former case, a consensus of opinion existed amongst the operators that, although it was "nice" to see how the pens altered with changes in the controls, the predictive facility was not of much practical use as temperature did not vary sufficiently to warrant a computer prediction. One is now left with trying to reconcile performance data which indicate a consistent improvement, with subjective reports of operators not consciously finding the predictive facility useful. It is evident from the trial logs and from operator comments that the predictive pens may have been used at least part of the time for control. If one discounts a Rosenthal (1966) 'trying to please the experimenter' effect as being incapable of explaining such a remarkably consistent improvement in performance, one is left with the possibility that the predictive information, when it was used, was taken in subconsciously. This hypothesis needs further testing, though it does fit with the notion of conscious and subconscious components to a process controller's hierarchical predictive 'internal model' of his plant, as recently suggested by Rasmussen (1974, 1976) and others. Information may be used routinely by an operator without his being aware of its input and hence being able to verbalise it.

Moving now to pH prediction, it is interesting to record that several operators actually requested a predictive facility be available on pH, and the plan of the original experiment was enlarged to accommodate this request. Unfortunately, as it has already been noted, a relatively unsophisticated predictor model is less well able to cope with fast-moving variables such as pH. One operator who reported using the pH predictor as a control aid became disillusioned with pen drift on the 30 second pen, and eventually gave up. In all, when predicting on pH the m.p.r. was judged to be more desirable, but less accurate, than on temperature.

5. CONCLUSIONS

The main conclusion from this chapter has been to confirm, in a realistic field setting, the findings of the previous chapter regarding the use of simple prediction models for slow response systems. The performance data indicate that the introduction of predictive information, in this case in the form of a multipen predictive recorder (m.p.r.) using a relatively unsophisticated smoothed Taylor series prediction model, can improve the control of a complex part-simulated semi-batch chemical reactor. The improvement was more marked when future values of slow-moving temperature were predicted than when predicting faster-moving pH.

A conflict between the objective performance measures and operators' subjective comments suggests that the predictive information may have been taken in at a subconscious level.

PART III

RESULTS SUMMARY, DISCUSSION AND CONCLUSIONS

CHAPTER 8

SUMMARY OF CONCLUSIONS FROM EXPERIMENTAL PROGRAMME

This chapter will, for convenience, present a brief summary of the major conclusions to date from the experimental programme.

1. EXPERIMENT 1 (Chapter 3)

1.1 Process

Discrete, soaking pit scheduling problem.

1.2 Situation

Laboratory simulation study.

1.3 Subjects

Students, varied academic backgrounds.

1.4 Factors Investigated

Input uncertainty (3 levels), prediction span (3 levels), academic background of subjects (maths/non-maths).

1.5 Performance measures

Scheduling errors, predictive activity, decision horizon.

1.6 Major conclusions

Decision aids designed to cope with uncertain environments are still adversely affected by the introduction of uncertainty. Scheduling performance deteriorated as information uncertainty was initially introduced, though performance actually improved when the level of uncertainty was further increased, as subjects gave up using the aid in the manner intended. A similar pattern to scheduling error variations with uncertainty was observed for the predictive activity data. A consistent reduction in the scheduler's decision horizon was also found when uncertainty was increased beyond a critical point.

Prediction span had a statistically reliable effect on scheduling performance only in deterministic (certain) conditions, where longer spans led to lower scheduling error scores. An optimum prediction span emerged in the average scores under uncertainty, but this was not confirmed by the statistical analysis. Reducing prediction span beyond a critical point did, however, lead to a reduction in both decision horizons and the amount of predictive activity.

Non-specialist, non-mathematical users achieved scheduling error scores which were not statistically different from their counterparts with a mathematical background.

2. EXPERIMENT 2 (Chapter 4)

2.1 Process

Discrete, job-shop scheduling problem.

2.2 Situation

Laboratory simulation study, test data from an operational job-shop.

2.3 Subjects

Students, mathematical background.

2.4 Factors investigated

Prediction span (3 levels).

2.5 Performance measures

Percentage machine utilisation, percentage of jobs unfinished, scheduling time.

2.6 Major conclusions

Under deterministic conditions, longer prediction spans lead to improved scheduling performance - a confirmation of the result from Experiment 1 (Chapter 3) under deterministic conditions. A critical value of prediction span was found below which machine utilisation performance deteriorated.

3. EXPERIMENT 3 (Chapter 5)

3.1 Process

Continuous, chemical process control, continuous stirred-tank reactor (CSTR).

3.2 Situation

Laboratory simulation study, hypothetical chemical industry task.

3.3 Subjects

Students/research workers, mathematical background.

3.4 Factors investigated

Train with/without predictor trace (2 levels), experimental trials with/without uncertainty (2 levels), presence and type of predictor model (3 levels).

3.5 Performance measures

Total product manufactured.

3.6 Major conclusions

This investigation should be regarded as an initial pilot study of the predictive display concept in continuous chemical process control. Due to inadequate smoothing, a simple Taylor series prediction model yielded reliably lower total product manufactured scores than did both a Perfect predictor and a No predictor condition. When uncertainty was present, the Perfect predictor trace gave the highest average figures. This was particularly noticeable for those subjects trained with the predictor. In the absence of uncertainty, the No predictor condition was to be preferred. It is suggested that predictive assistance may only be of benefit where the task is sufficiently demanding to warrant such assistance, as when uncertainty is introduced.

Predictive displays seem to facilitate training, in that those subjects trained with a predictor trace achieved higher average production scores than did those trained without. It may well be that the value of predictive displays lies in helping controllers to form an accurate internal model of the process during training.

Various problems are encountered in this pilot study, largely as a result of attempting to simulate a representative chemical industry task with naive student subjects in the laboratory. It was suggested that future work should concentrate on a much-simplified laboratory problem on which an in-depth study of predictive display characteristics could be conducted (Chapter 6). A second phase would comprise a full-scale field study to test the best overall design in a realistic industrial setting (Chapter 7).

4. EXPERIMENT 4 (Chapter 6)

4.1 Process

Continuous dual-meter monitoring and control task.

4.2 Situation

Laboratory simulation study.

4.3 Subjects

Students, mathematical background.

4.4 Factors investigated

Uncertainty (3 levels), prediction span (4 levels), prediction model fidelity (2 levels), plant gain (3 levels).

4.5 Performance measures

Time outside limits scores, control histograms, display switching.

4.6 Major conclusions

The introduction of a predictive display was found to lead to a consistent improvement in control performance. In general, time outside limits error scores were found to increase with faster plant response and increasing level of uncertainty. However, a complex interaction between plant gain, uncertainty, prediction span and prediction model fidelity emerged.

For systems with a slow speed of response, there was little to choose between a highly accurate (Perfect predictor) and a relatively unsophisticated (Taylor series) prediction model, given that adequate performance with the latter was achieved at the expense of greater control effort. For systems having moderate to fast response times, the more sophisticated prediction model was justified.

Recommendations can be made regarding the choice of an appropriate prediction span for simple and sophisticated prediction models under various levels of plant responsiveness and input uncertainty. The simple prediction model seemed relatively immune to uncertainty variations, but was affected by the plant gain. When applied to systems with low to medium gains, longer prediction spans led to lower error scores. A slight reduction in usable prediction span might be advisable with high gain systems. Using the hypothetical Perfect predictor model, the optimum prediction span was progressively reduced with the combined effect of increasing uncertainty level and increasing plant gain. The conflicting findings of previous workers can be explained in terms of the differing gains, levels of uncertainty, and prediction models of the systems with which they were concerned.

5. EXPERIMENT 5 (Chapter 7)

5.1 Process

Continuous, semi-batch chemical reactor.

5.2 Situation

Field study with real plant and instrumentation, reaction itself computer-modelled.

5.3 Subjects

Experienced process operators.

5.4 Factors investigated

With/without multipen predictive recorder (2 levels), temperature/pH prediction (2 levels).

5.5 Performance measures

Percentage conversion, standard deviation about predicted variable, calculated batch profit.

5.6 Major conclusions

The recommendations of Chapter 6 regarding the adequacy of simple prediction models for slow-response systems have been confirmed in a realistic industrial setting. Performance data indicated that the introduction of a relatively crude predictive facility in the form of a multipen predictive recorder brought about an improvement in the control of the complex semi-batch chemical process studied.

Mean percentage conversion of reagents to finished product and mean batch profit figures were both enhanced. The improvement was most marked when future values of slow-moving temperature were predicted than when predicting the faster-moving pH.

CHAPTER 9

DISCUSSION

1. OVERVIEW

It may be helpful at this stage to recap how the philosophy underlying the experimental work has developed. The introductory chapters established the need for a programme of research to demonstrate the usefulness of predictive displays across a variety of discrete and continuous tasks, and to examine how variations in parameters both internal and external to the predictive display affect operator performance. The factors most critical to predictive display operation were established as input uncertainty, prediction span, and complexity of the prediction model.

Following directly from Laios' (1975) work on Predictive Computer Displays, a first experiment (Chapter 3) examined the effects of a wide range of uncertainty and prediction span values on operator performance in a simulated steelworks soaking pit scheduling task. In the absence of a real-world scheduling environment in which to verify the findings of Chapter 3, it was deemed methodologically appropriate to employ a further laboratory simulation of another discrete scheduling environment, this time a manual job-shop scheduling problem, but using test data from an operational job-shop (Chapter 4). Only the effect of prediction span was tested, since the result of uncertainty in a real-world scheduling environment had been demonstrated elsewhere (Bibby, 1974). This concluded the experimental work carried out on aids for discrete systems.

Chapter 5 branched into continuous process control applications with a pilot study to investigate the potential benefits of predictive displays therein. It became clear that trying to simulate a representative chemical process in the laboratory was not feasible in the time scale available. A two pronged approach was therefore adopted. A simplified laboratory simulation of a predictor-assisted dual-meter monitoring and control task with student subjects (Chapter 6) was employed to examine such parameters as input uncertainty, prediction span, prediction model fidelity and system gain. Chapter 7 took the predictive display concept, formulated as a multipen predictive recorder (m.p.r.), into an industrial field situation (or as near to one as any systems engineer testing a prototype device is likely to get). Real plant and experienced operators were used to evaluate the m.p.r. in the control of slow-moving temperature and faster-moving pH.

The experiments have relied for the most part on objective performance measures, largely because these give a good impression of predictive display effectiveness, and because they were readily available from the computer-based systems emloyed. Where appropriate, however, objective measures were supplemented by formal or informal questionnaires and the collection of limited verbal protocols (Chapters 3, 4 and 6).

In the introduction to the experimental programme, six aims were outlined broadly relating to:-

1) What are the effects on performance of variations internal and external to the predictive display in discrete and continuous tasks?
2) Are the results of laboratory studies borne out in the real-world? and
3) What are the implications of predictive display research for the modelling of human control and decision-making behaviour?

Having reviewed the philosophy behind the experimental programme, the following sections will attempt to answer these questions and to discuss, expand and inter-relate the experimental findings within the existing body of predictive display knowledge.

2. BENEFITS OF PREDICTIVE DISPLAYS

2.1 Objective Performance Measures

It is encouraging to note that, in common with the previous predictive display studies reviewed in Chapter 2, the experiments uncovered distinct improvements in performance through the use of predictive displays to assist in the control of industrial processes. In the later experimental chapters (Chapters 5, 6 and 7) this was reflected as a straightforward improvement in the 'with predictor' conditions compared to the 'no predictor' conditions. In the earlier experimental chapters (Chapters 3 and 4) the benefits of predictive displays over conventional equipment was not tested directly, but rather were reflected indirectly through a gradual reduction in prediction span, i.e. by cutting down in stages the amount of predictive information displayed to the operator.

It can be stated with confidence that for the direct comparisons of operator performance with and without prediction, the predictive display is to be preferred on all counts to traditional non-predictive alternatives. In the dual-meter monitoring and control task of Chapter 6, for example, an improvement in the average time outside limits scores of up to 40% using the Taylor series extrapolation model and up to 70% using the hypothetical Perfect predictor resulted (Figure 23), coupled with a smoother pattern of control (Figure 27). Using the multipen predictive recorder (m.p.r.) of Chapter 7, the average batch profit increased from £80 to £138 (73%) when predicting future values of temperature, and from £70 to £78 (12%) when predicting the faster-moving pH variable. In addition, fluctuations around the temperature and pH target values of 65^{o} and 7 respectively were reduced. This study has particular significance since it was conducted in a pseudo real-world environment with actual plant and experienced operators.

The pilot study of Chapter 5 must be considered as an exception to the general rule, since as was noted in that Chapter, the unsmoothed Taylor series prediction model and the simulated task itself were open to criticism. However, even in this task a small improvement of up to 4.5% in the total amount of product manufactured was observed where student subjects had been specifically trained to use the predictor, and where the task was sufficiently demanding. It should be noted that even a marginal improvement in amount produced or similar scores (witness the seemingly negligible 1.8% increase in percentage conversion from Chapter 7) can mean a substantial financial gain, depending on the pricing structure. The general findings from the pilot study seem to fit in with Rouse's (1970) observation that predictive displays may only be useful in tasks of medium difficulty - in very simple tasks they are not strictly necessary and may only serve to distract. Similarly, in highly complex tasks the operator may be overloaded and ignores the aid, instead responding at an intuitive level. The pilot study was also worthwhile in that it tended to confirm Smith and Kennedy's (1975) suggestion that the true value of predictive displays may be in the training environment. In Kelley's (1968) language, predictive displays assist the operator to build up an accurate internal model of the process. In the pilot study, subjects trained with the predictor subsequently performed better in all conditions (though not reliably different in statistical terms) than their colleagues who had been trained conventionally.

The alternative method of testing predictive display effectiveness was indirectly via a gradual reduction in the prediction span. In the first of the experimental chapters (Chapter 3) to adopt this indirect approach, conflicting results were obtained depending on whether or not uncertainty was present. Under deterministic conditions (i.e. no uncertainty element) the mean number of scheduling errors rose by 40% as prediction span was reduced from full to quarter screen. It is fair to assume that the quarter screen condition was close to a 'no predictor' condition since insufficient predictive information was in fact visible to the operator for him to make effective use of it. The reduction confirms the work of Laios (1975) who obtained a reduction of 75% in scheduling errors with the introduction of a predictive facility. (The percentage difference between the two studies is probably due to variability in the subject pools, as well as the slightly different nature of the comparisons). Under input uncertainty, however, the reduction in prediction span had a somewhat different effect, with the emergence (though not statistically significant) of an optimum prediction span in the average scores (Figure 7). This point is of particular interest and will be elaborated in section 3.

Also under deterministic conditions, the results of Chapter 4 using test data from an operational job shop confirm that the facility to predict ahead with an analogue representation of the problem environment can be beneficial, even if the 'test loads' are done manually using wooden blocks rather than electronically on a c.r.t. display. Scheduling performance as reflected by the machine utilisation scores was observed to fall off by 6% as prediction span was reduced. The previously made point about marginal performance improvements and eventual financial benefits is again relevant here. Gibson and Laios (1978) had already shown the superiority of the scheduler's abacus arrangement over a non-predictive but otherwise equivalent, numerical representation.

2.2 User Acceptance

Subjective acceptance of the predictive display devices was on the whole also favourable. The only exceptions were when predictor credibility had been reduced by inadequate smoothing (as with the unsmoothed Taylor series of Chapter 5), or by the choice of an inappropriate prediction model (as when the m.p.r. of Chapter 7 was set to predict on pH); or when high levels of uncertainty or a very short prediction span reduced the effectiveness of the display (Chapters 3 and 6).

In the scheduling studies of Chapters 3 and 4, subjects made extensive use of the planning facility and commented to this effect. When part of the planning facility was blocked off, subjects complained of the increased task difficulty and several reported continuing the 'test load' procedure mentally. The predictive facility was flexible enough to cope with individual subjects' differing search patterns, i.e. depth of search, degree of organisation when testing alternatives, etc. Non-mathematical users appeared to be as capable of using the predictive aid as were users with a mathematical background.

In the pilot study of Chapter 5 the Perfect predictor model was reported to be useful in initially filling the kettle without incurring overheat errors, and also in the later phase of controlling the reaction. It is particularly interesting to record that without being told of this peculiar program idiosyncrasy, several subjects made use of the predicted temperature trace's irregularity pending kettle overflow to prevent the latter occurring. This evidently agrees with the common observation that operators can detect and employ informal communication channels (display quirks, noises, smells, etc.) to assist with their control.

Indeed, Agnew and Pyke (1969) have commented that left to his own devices the human will use anything that pre-packages information for him.

Subjects also stated their preference for the predictive display option in the dual-meter monitoring and control task of Chapter 6, though opinions were divided for preference between the accurate and the crude prediction models. The experienced operators of Chapter 7, however, commented that, whilst the predictive pens of the m.p.r. were "nice" to look at, the temperature control problem was not sufficiently complex to warrant predictive assistance. Performance was improved, however, and as noted in that chapter it may well have been that the predictive information was taken in subconsciously. In addition, the operators had been used as guinea-pigs to test various control aids prior to the present study, and had developed a somewhat jaundiced attitude towards any form of assistance. This could well explain the apparent discrepancy between the opinions of the naive student subjects and the experienced Warren Spring operators. The latter's comments may not therefore be representative of the display's reception in practice.

3. FACTORS AFFECTING PREDICTIVE DISPLAY PERFORMANCE

3.1 Input Uncertainty

The general impression from the programme of experimental studies is that input uncertainty, whether in the form of unreliable arrivals information or the contamination of signals by noise, degrades predictive display effectiveness. This finding is in accordance with the limited previous work in this area reviewed in section 3.1 of Chapter 2 (e.g. Bibby, 1974; Laios, 1975). The reasons for the degradation in performance are not difficult to comprehend. In discrete cases (e.g. Chapter 3) schedules made under uncertainty will be based on unreliable information and once planned may require frequent updating as arrivals do not occur at the expected times. If insufficient time is available to revise the schedule an off-the-cuff choice must be made, and with successive choices there is an increasing likelihood that one of these will be in error. Once an inappropriate choice has been made, it is a characteristic of multistage decision-making tasks that performance is usually affected for the whole task. In continuous applications (e.g. Chapter 6) with noise contamination present, the predicted path displayed to the operator will be in error by an amount proportional to the magnitude of the input disturbance. As in the discrete case, the operator will receive an erroneous impression of how the plant is or will be behaving, and his control actions will be inappropriate.

It should be noted that performance still deteriorates if information about the uncertainty is displayed to the operator, either by showing the likely interval within which the arrivals will occur (as in Chapter 3), or by incorporating diagnostic information about the uncertainty as part of the predicted trace (as in Chapter 6, Perfect predictor model). A true 'adaptive display' (Kelley and Prosin, 1972) would doubtless be necessary if it were desired to make an effective diagnosis about the uncertainty element. In fact, unacceptably high levels of uncertainty can occur in real life when an operator takes over from an automatic controller, either in an emergency or as part of routine maintenance procedures. It will take a certain amount of time for the operator to build up an accurate mental picture of the current process state.

Several qualifying remarks must, however, be made to the earlier statement that uncertainty degrades performance. Firstly, the level of uncertainty is important. Chapter 3 showed that although performance deteriorated at moderate levels of uncertainty, if the uncertainty level increases to such a degree that rational choices cannot be made on the basis of the displayed information, then operators may give up using the aid as intended and may revert to their previous unaided method of dealing with the task, perhaps using an internal model of the problem environment formed with the assistance of the predictive display. Rouse (1970) has also commented to this effect. In this case then, depending on the operator and the task, performance may deteriorate or may even improve.

Secondly, the effect of uncertainty is confounded by interactions with parameters of the predictive display and with characteristics of the task itself. This will be discussed further under subsequent headings. In Chapters 3 and 6 for example, with a perfectly accurate prediction model, one of the main effects of uncertainty was to reduce the usable prediction span (in Chapter 3 this was evident in the averages of Figure 7, though not statistically significant). Crude system models on the other hand seem relatively immune to uncertainty, as witnessed by the Taylor series expansion of Chapter 6. There is thus an apparent dichotomy between Class I and Class II predictive displays regarding their susceptibility to input uncertainty.

The effect of uncertainty on predictive display effectiveness is also dependent to some degree on the difficulty of the original task. In the pilot study of Chapter 5, the task without uncertainty was sufficiently easy for the operator to manage quite well without the predictive display on temperature, predictive information serving only to distract. When uncertainty was introduced however, the task became sufficiently demanding to warrant an accurate predictor and so performance improved. As noted earlier in Chapter 5, the introduction of uncertainty also served to start the reaction phase slightly earlier, which may go part-way to explaining the improved production figures under uncertainty in this task. In the dual-meter task of Chapter 6 however, the basic problem was in itself sufficiently taxing for the additional load caused by the introduction of uncertainty to result in a worsened performance.

3.2 Prediction Span

A main finding from the experimental programme is that the choice of prediction span is a function of other predictive display parameters and task characteristics. Such an interaction was hinted at, though not proven, by some previous workers, e.g. Kelley (1960a), Bernotat (1972), who commented that different systems would undoubtedly need different prediction spans, probably related to the 'responsiveness' of the system (plant gain and control order) and to the magnitude and frequency of unpredictable disturbances (input uncertainty). To this list must be added the fidelity of the prediction model. Considering that a four-way interaction is effectively present between these parameters, the present author believes that it is hardly surprising that in the past different authors have come up with conflicting findings regarding choice of prediction span.

Firstly, with a sufficiently demanding task and an entirely accurate prediction model (as in Chapters 3 and 6), uncertainty reduces the useful range of prediction spans. As prediction extends further into the future it becomes progressively more inaccurate and likely to mislead the operator. As Rouse (1970), Dey (1971) and Bernotat (1972) also found, an optimum prediction span thus emerges with spans either side leading to reduced performance (see Figures 7, 25 and 26). Under deterministic conditions (Chapters 3, 4 and 6) however, or when the plant is sufficiently slow moving (Chapter 6, low gain - Figure 24), there is little restriction on the length of prediction span. The 'optimum span' phenomenon thus seems to be most obvious in complex systems.

Secondly, with crude prediction models, the effective range of usable prediction spans is enhanced. This seems to confirm Bernotat and Widlok's (1966) finding that a first-order, extrapolative predictor was useful over a wide range of prediction spans, though as the order of the prediction model was increased, i.e. as it became more accurate and nearer to a hypothetical 'perfect predictor', the useful range of prediction spans was reduced. The findings of Chapter 6 suggest that crude prediction models of the type used by Bernotat and his colleagues are relatively immune to uncertainty, so that no shortening in prediction time was necessary even with quite high noise levels. This may well be because the slow response rate of such a model acts as a partial filter to the input noise.

Lastly, prediction span is related to the dynamics/response characteristics of the process. Chapter 6 demonstrated that, with a completely accurate prediction model, the effect that uncertainty had of reducing the usable prediction span was most marked for the higher gain conditions (Figures 25 and 26). In the low gain condition (Figure 24) it had far less effect. With the Taylor series prediction model, on the other hand, the effect of different prediction spans was most marked in the low gain condition, and its effect lessened as plant gain increased until in the high gain condition no significant difference could be detected between different spans.

In a sense, high uncertainty levels and fast plant dynamics can be thought of as having the same effect on accurate prediction models, since both effectively increase task difficulty. It may be that there is an optimum distance for looking ahead in all control and decision-making applications, related to the task complexity and any lags present. Tomizuka and Whitney (1975) have suggested that an optimum preview distance exists for automatic control systems, and have developed a rule of thumb relating practical preview distances to controlled plant eigenvalues. A similar concept may well apply to operator-inclusive predictive display systems. Performance will attain its maximum potential when the prediction span is equal to the required decision horizon for the task. For example, in the scheduling task of Chapter 3, it was calculated that adequate performance could be achieved by planning 2-3 stages ahead, and in continuous control applications by extrapolating up to the next required control reversal or 'turning point' of the track. In fact an optimum prediction span of 20 minutes (or roughly 1-2 stages ahead) did occur under uncertainty for the average scheduling error scores of Chapter 3, lending support to this hypothesis. The situation for continuous control tasks is more complex.

Hollister (1967) has reworked much of Bernotat and Widlok's (1966) early experimental data in an attempt to explain and mathematically model their results. (It has already been noted in section 5.1 of Chapter 1 that such control theoretic models may adequately describe a human operator's output, but do not necessarily represent his internal processes.) Hollister's estimates of optimum prediction times agreed to within 8% of Bernotat's experimental results. Hollister considers a third order controlled process similar to that studied in Chapter 6 as part of his analysis, and deduces the optimum prediction time to be inversely proportional to the plant gain, as well as inversely proportional to the difficulty the operator experiences in controlling the process. The results from Chapter 6 seem to bear out this relationship, particularly with the hypothetical perfect predictor model. Optimum prediction spans were reduced as plant gain increased and, if one considers the introduction of uncertainty as an increase in task difficulty, then also as task difficulty rose. With the Taylor series extrapolation model however (the same model used by Bernotat), system gain had a far lesser effect on optimum prediction times, and task difficulty as reflected by increased uncertainty apparently had no effect at all.

In practical terms, Bernotat (1972) appears to recommend that the choice of prediction time for minimum control error should be of the same order of magnitude as the response lag of the controlled process - as an example he suggests a prediction time of 0.7 seconds for a third order system where the lag is equal to 0.4 seconds. This seems a useful rule-of-thumb to follow, and is supported by analysis of Chapters 6 and 7: in Chapter 6 the measured response lag to a control change varied from approximately 15 seconds (low gain) to only a few seconds (high gain). The batch kettle process of Chapter 7 had a lag of between 40 seconds and 1 minute (King, 1975).

In general it would seem that the choice of prediction span should be roughly proportional to the response time of the system, where response time is a function of plant dynamics and the control task. In tasks of variable difficulty, an operator-adjustable prediction span as suggested by Kelley (1960b) may be the best solution.

3.3 Prediction Model Fidelity

This factor was examined directly in Chapters 5 and 6. A main finding of Chapter 6 was that for slow response systems there is little to choose between a completely accurate prediction model of the process and a far less sophisticated alternative, given that adequate performance with the latter is achieved at the expense of greater control effort. Even a crude mathematical prediction will often produce far better extrapolations than the human operator can manage. Smoothing of the data on which predictions are based is a vital factor, however. As Chapter 5 showed, inadequate smoothing can seriously reduce the credibility of a crude prediction model, giving worse performance with the predictive display than with no prediction at all.

In practical terms, the over-riding advantage of a crude prediction model such as the Taylor series expansion is that the computational requirements are greatly reduced. It would be technically feasible to implement such an extrapolative device on a micro-processor chip, the resulting predictions being displayed on a c.r.t. display (in line with current trends in process control) or on a modified pen recorder (in line with current practice). The latter approach was adopted in the pseudo-field test of Chapter 7. With the m.p.r.'s predictive pens set to predict on temperature, the practical usefulness of a predictive display based on a crude extrapolation model was confirmed. However, prediction of the faster-moving pH variable was not so successful, and only a small increase in average batch profit scores was obtained. This result was in a sense pre-empted by the findings of Chapter 6 : with faster-moving systems, a more sophisticated prediction model with an appropriately chosen prediction span is to be preferred (see Figures 25 and 26). (A practical drawback to using an accurate fast-time model is, of course, that it is considerably more difficult to implement.) Hence it is not surprising that when set to predict on the faster-moving pH variable, the crude Taylor series expansion was inadequate.

As previously noted, the choice of an appropriate prediction model is inter-related with other predictive display and task parameters. The design engineer has a wide range of models to choose from, ranging from the relatively unsophisticated Taylor series expansion, through Kalman filters and statistical predictors, to fully fledged dynamic system models. In practical terms, the choice of an appropriate prediction model must depend on the particular task.

Reports from the subjects of Chapter 6 suggest that operators did not form precise internal models of plant dynamics but rather used some form of crude process representation, e.g. a yo-yo, simple harmonic motion, inverted pendulum etc., on which to base their predictions. Process operators are also known to form and make use of crude rules-of-thumb (Ketteringham et al., 1970), category classifications of process output (Bainbridge,1975a), and simplistic control models which ignore process interaction effects (King and Cininas, 1976). It is interesting to note that the crude internal models which experienced operators are known to adopt have been shown, in the context of automated predictive displays and when adequately smoothed, to be almost immune to uncertainty and perhaps also to other variations in the environment. The implications of this statement for a process operator's choice amongst his assumed repertoire of possible internal models will be discussed at length in a later section. Umbers (1976) has also shown that grid controllers in uncertain environments either made rough estimates or did not bother at all when they realised that their predictions did not lead to accurate results.

3.4 Process dynamics/response characteristics

Chapter 6 was the only chapter specifically to investigate process dynamics, in the form of plant gain, on predictive display effectiveness. In absolute terms, performance fell off with increasing plant gain, though in relative terms the improvement with predictive assistance over the no prediction condition was greatest for the higher plant gains (Figure 25 and 26). This agrees with Bernotat and Widlok's (1966) findings. These authors had shown that, in a pilot study varying the gain of a third order system over the full range of possible values, the relative improvement due to the predictor was also greatest for high gain values. In absolute terms, performance similarly fell off at high gain values, and interestingly at very low values as well - in Chapter 6, there was some evidence that performance also deteriorated at very low gains, since the pointers moved so slowly that changes in pointer position were virtually impossible to detect. Warner's (1969) finding that exploratory prediction was independent of system gain over the ranges investigated is probably a reflection on the performance measures he used.

It must be stressed that the effect of gain went hand-in-hand with the various predictive display parameters mentioned earlier. At low gains, performance was virtually identical using sophisticated or crude prediction models, though the latter was at the expense of greater control effort. Although performance deteriorated for both prediction models as gain increased, a gain x prediction span interaction was present with the crude prediction model. This meant that for the Taylor series extrapolation model the effect of different prediction spans lessened as plant gain increased. The reverse effect was found for the perfect predictor, both prediction span and plant gain interacting with the level of uncertainty.

Though the gain, or process 'speed' factor, was not investigated specifically in the scheduling application of Chapter 3, there is no reason to suppose that a faster process would not have reduced predictive display effectiveness, since less time would have been available to try out alternative schedules. Indeed, when designing the simulation a simulated process speed of twice real time was chosen, since a faster process was difficult to control and running in real time induced problems of fatigue and boredom.

3.5 Repetition rate/frequency of updating

These parameters were not investigated experimentally, being largely pre-determined by the computer systems on which the simulations were run. Updating of the prediction model occurred every second for the PDP-12 laboratory simulations, and once every 10 seconds on the Argus 500 computer at Warren Spring. The impression gained was that the inaccuracies of a crude prediction model could to some extent be tolerated if the interval between successive displayed predictions were sufficiently small.

3.6 Mode of control

It is difficult to talk about modes of control without again raising the distinction between discrete and continuous predictive display applications. As was mentioned earlier, the predictive display concept was found to benefit both discrete scheduling and continuous control applications. The main differences between the two areas seems to be in the appropriate choice of useful prediction spans, and in the mode of predictive display operation. In discrete applications, the effective time scale seems to be longer, so prediction spans in the order of minutes or hours are appropriate. Since time is available to test alternative options, exploratory prediction is the best choice - indeed, this mode of control was adopted in Chapters 3 and 4. For continuous applications, little time is available to test alternative actions, and therefore on-line prediction is most appropriate. This form was used in Chapters 5, 6 and 7. Here prediction spans were in the order of seconds.

However, it should be stressed that in either case the use of the predictive display is the same - to provide future-oriented information about the state of the system. In exploratory prediction, the operator must search out the prediction for himself, whereas in on-line prediction the consequences of his actions are provided automatically for each control change that he makes. This lends further support to the hypothesis that a single model may be appropriate to explain human operator behaviour in both continuous control and decision-making tasks. As was noted in the introductory chapter, Gregory (1970) has pointed out that, whether the process be discrete or continuous, the human operator samples information and so his processing of it must be through a discrete mechanism. Subsequent sections will attempt to construct a model of the human operator combining the best features of existing models in accordance with the present experimental evidence.

4. ROLE OF PREDICTION IN HUMAN CONTROL AND DECISION-MAKING

It would seem, from the improvement in performance when predictive assistance was introduced, and from subjects' reported mental procedures when the predictive displays were removed (or became unusable due to high uncertainty levels or short prediction spans), that prediction is an important part of human control and decision-making skills. This conclusion is in accord with the existing body of research reviewed in Chapter 1. In addition, anecdotal evidence suggests that operators build up over time an internal or mental model of the process, and that the predictive display is of considerable benefit in helping them to do this. As noted in Chapter 5, subjects trained with the predictive display achieved higher production scores in all subsequent trials. Hence it may well be that the practical significance of predictive displays is in the training environment, as suggested by Smith and Kennedy (1975).

A separate study on the Batch Kettle plant at Warren Spring had also shown that experienced operators did not employ open-loop, 'feedforward' control based on prediction since they experienced difficulty in understanding process interactions (King and Cininas, 1976). Instead they reduced their control model to single loops wherever possible, and treated interaction effects as disturbances. This 'rule of thumb' approach was sub-optimal. Hence it is not surprising that providing a job aid to assist the operators in making feedforward control actions (the m.p.r.) brought about an improvement in performance.

Since prediction is such an important component of human control and decision-making skills, it is evident that any model of the human operator in the experimental tasks of this thesis must centre around an internal predictive model of the process. The performance measures discussed so far have been mainly objective, coupled with subjective comments, and whilst useful in determining the effectiveness of the various predictive displays they have been inadequate for modelling purposes. To gain a better understanding of human thought processes in industrial tasks, the verbal protocol approach was used and is discussed in section 5.

5. VERBAL PROTOCOLS

It was mentioned earlier that in addition to the limited subjective reports whilst performing the various experimental tasks, detailed verbal protocols were collected for the experiments of Chapters 3, 4 and 6. As has become the tradition in protocol studies, one subject was used as the source of the protocol data. The subject in question was experienced in all the experimental tasks and had achieved consistently good performance scores. He was chosen not only for his ability, but for his willingness to verbalise his thought processes.

A separate study of operator strategy using verbal protocols had also been carried out on the Warren Spring Batch Kettle plant (Cininas, 1976) before the introduction of the m.p.r., and reference will be made to the findings of this study for comparison purposes. The pilot study of Chapter 5 was not considered a suitable task for protocol analysis in view of the problems encountered with the smoothing of the prediction model and with the task itself.

Analysis of the protocol data followed the now standard format of:

1) Transcription of protocols.
2) Classification of phrases into an exclusive set of activities.
3) Grouping of activities to form recognisable and recurrent patterns.
4) Development of a flow-diagram representation of the behaviour.

A detailed description of the technique is given in Bainbridge (1972) or Umbers (1976). The next sections present the flow diagrams developed for the experimental tasks of Chapters 3, 4, 6 and 7. Although they are based on the thoughts of only one subject, the basic protocol material was supplemented by discussing the controller's strategy with him after each protocol had been recorded. The flow diagrams should therefore be treated as a general description of how a typically competent operator might set about controlling the various processes. This approach is deemed sufficient to permit comparisons of control and decision-making behaviours between the tasks.

5.1 Discrete tasks

The routines used in the scheduling applications of Chapters 3 and 4 are described respectively in Figures 36-38, and in Figure 39. It can be seen that distinct similarities are present between the two cases. In both, much time was spent in trying out the effect of schedule adjustments in relation to the target objectives, by evaluating the consequence of a particular allocation of ingot to pit, or of job to machine. Evidence from the two tasks tends to lend support to Bainbridge's (1974) concept of a loose, hierarchical goal-directed model in decision-making tasks. As has been previously noted, subjects appeared to be setting up goals and sub-goals within the main task objectives of achieving a constant output flow from the soaking pit complex, or maximising machine utilisation whilst minimising jobs unfinished in the production scheduling problem. The sub-goals related to the allocation of particular ingot arrivals (or groups of these), and particular jobs (or groups of jobs), which the subject scheduled to his satisfaction before moving on. The sequence of operations did not appear to be particularly rigid. For example, when re-scheduling in the soaking pit problem the subject might clear the entire screen of all test loads, clear only the pits he was interested in, or superimpose a new value onto an old test load. This suggests a flexibility of approach, confirmed in the job shop scheduling problem.

Figure 36 shows that a central feature of the soaking pit scheduling problem with virtually no predictive assistance was an internal model of the pit soaking times, with which schedule clashes could be predicted before making a test load. Ingots were scheduled as they became visible from beneath the card, and the main use of the PCD was to determine pit status before test loading. The effect of the predictive aid was that the consequences of a decision could be worked through on an external representation of the problem environment, whereas in the unaided task this had to be done mentally. The PCD also gave an explicit representation of how the schedule 'pattern', or distribution of test loads, was developing. Though the job shop scheduling problem was considerably more complex, the planning facility was similarly employed not only to check for schedule conflicts but as an aid to the detection of spaces and odd gaps where a job or part-job might fit. In both the soaking pit and job shop scheduling problems, the subject seemed capable of talking through the consequences

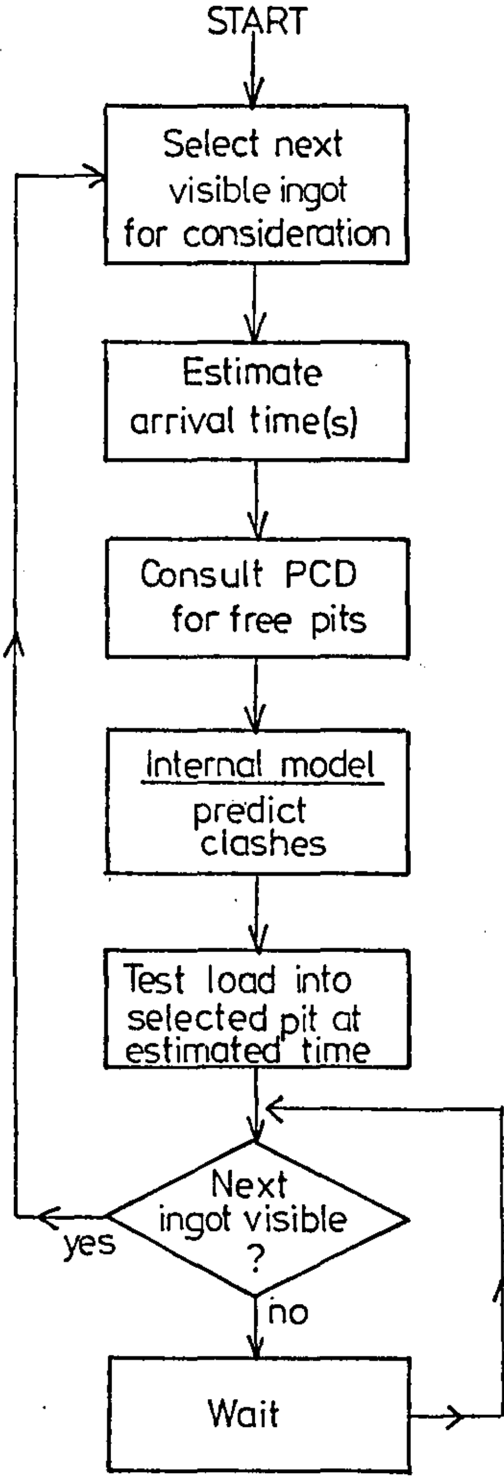


Figure 36: Flow diagram of behaviour in soaking pit scheduling problem with minimal assistance (quarter screen prediction span, low uncertainty) based on verbal protocols.

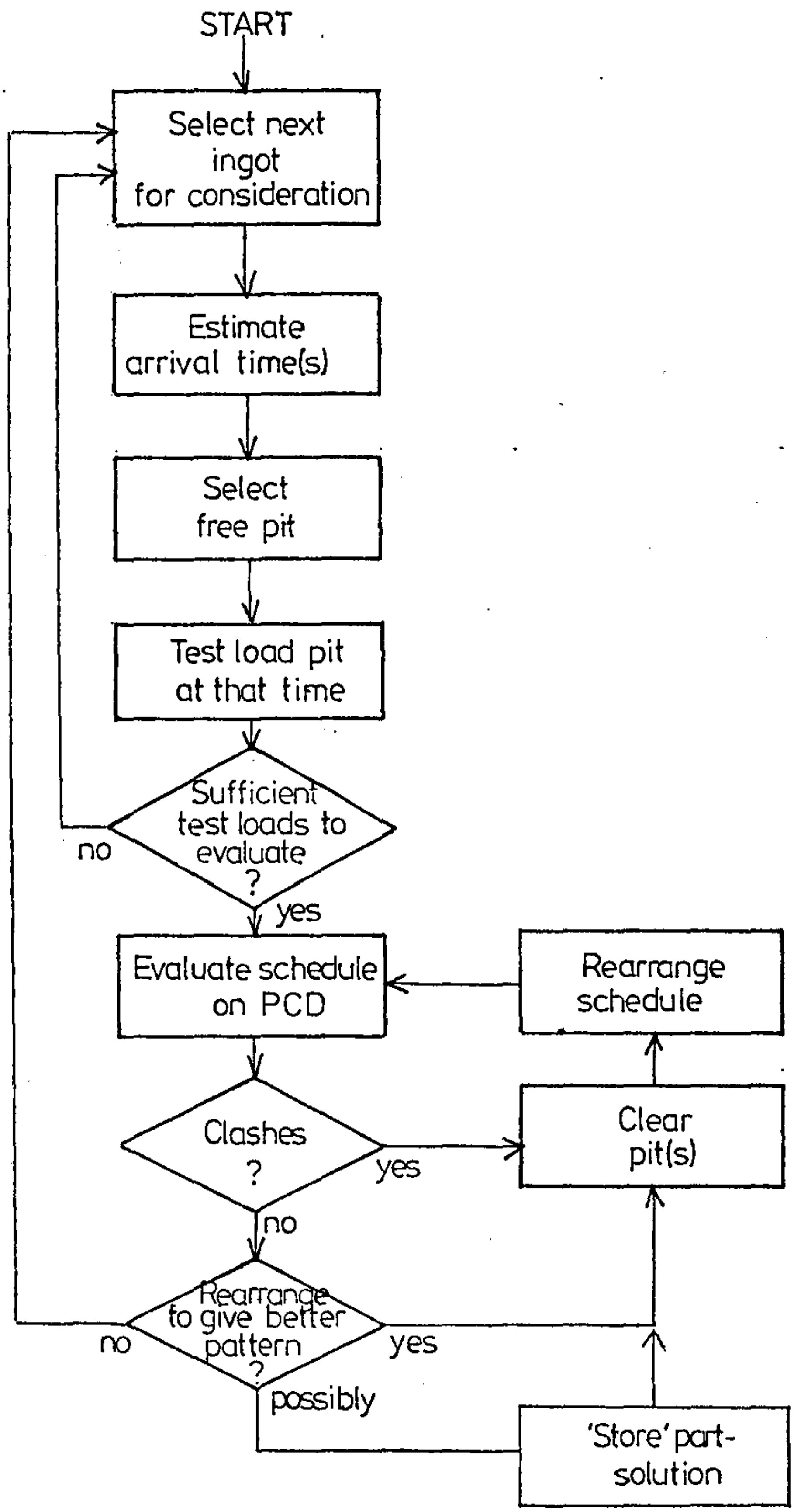


Figure 37: Flow diagram of behaviour in soaking pit scheduling problem with predictive assistance (full screen prediction span, low uncertainty) based on verbal protocols.

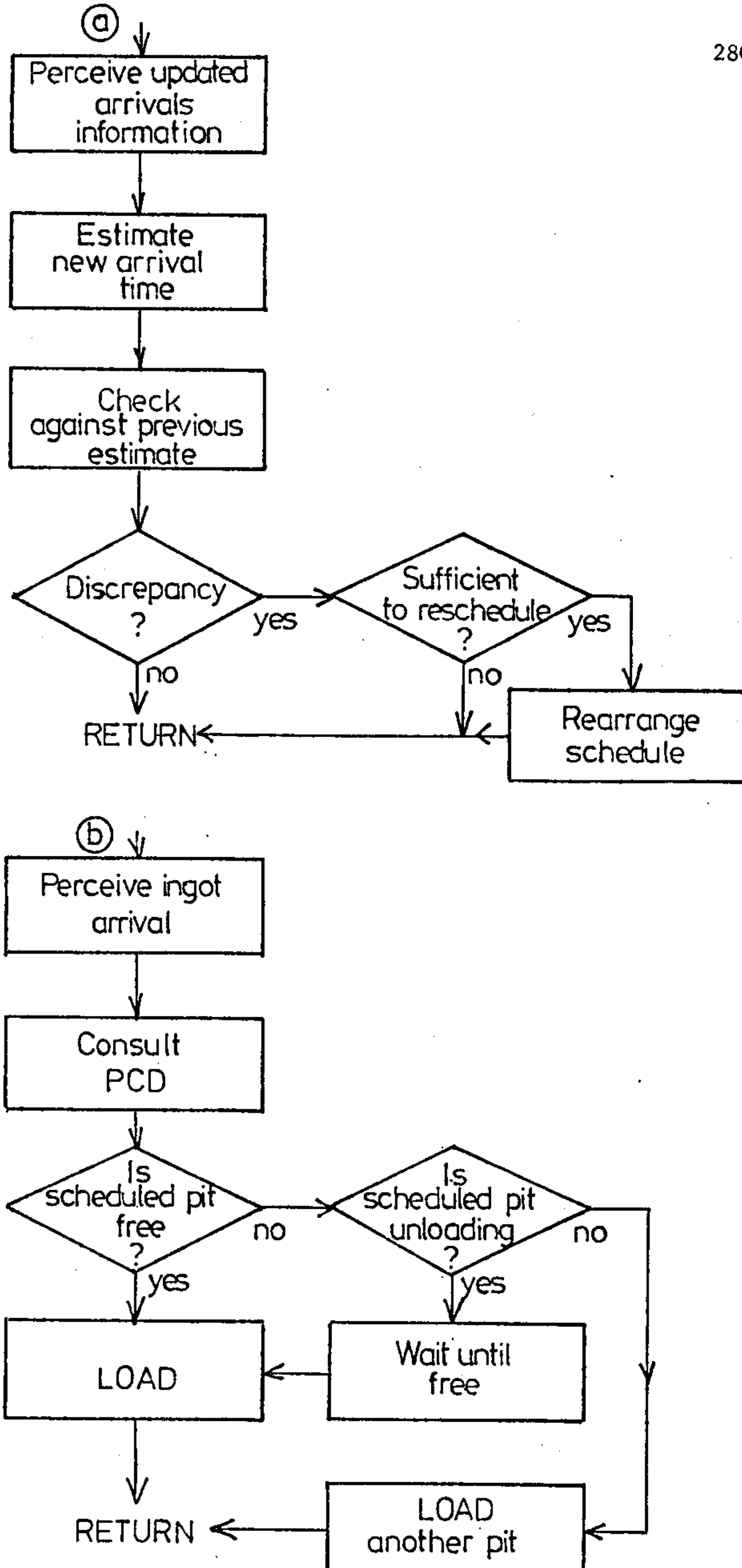


Figure 38: Subroutines common to Figures 36 and 37 to account for behaviour in soaking pit scheduling problem when
(a) the arrivals information was updated, and
(b) a cast arrived at current time

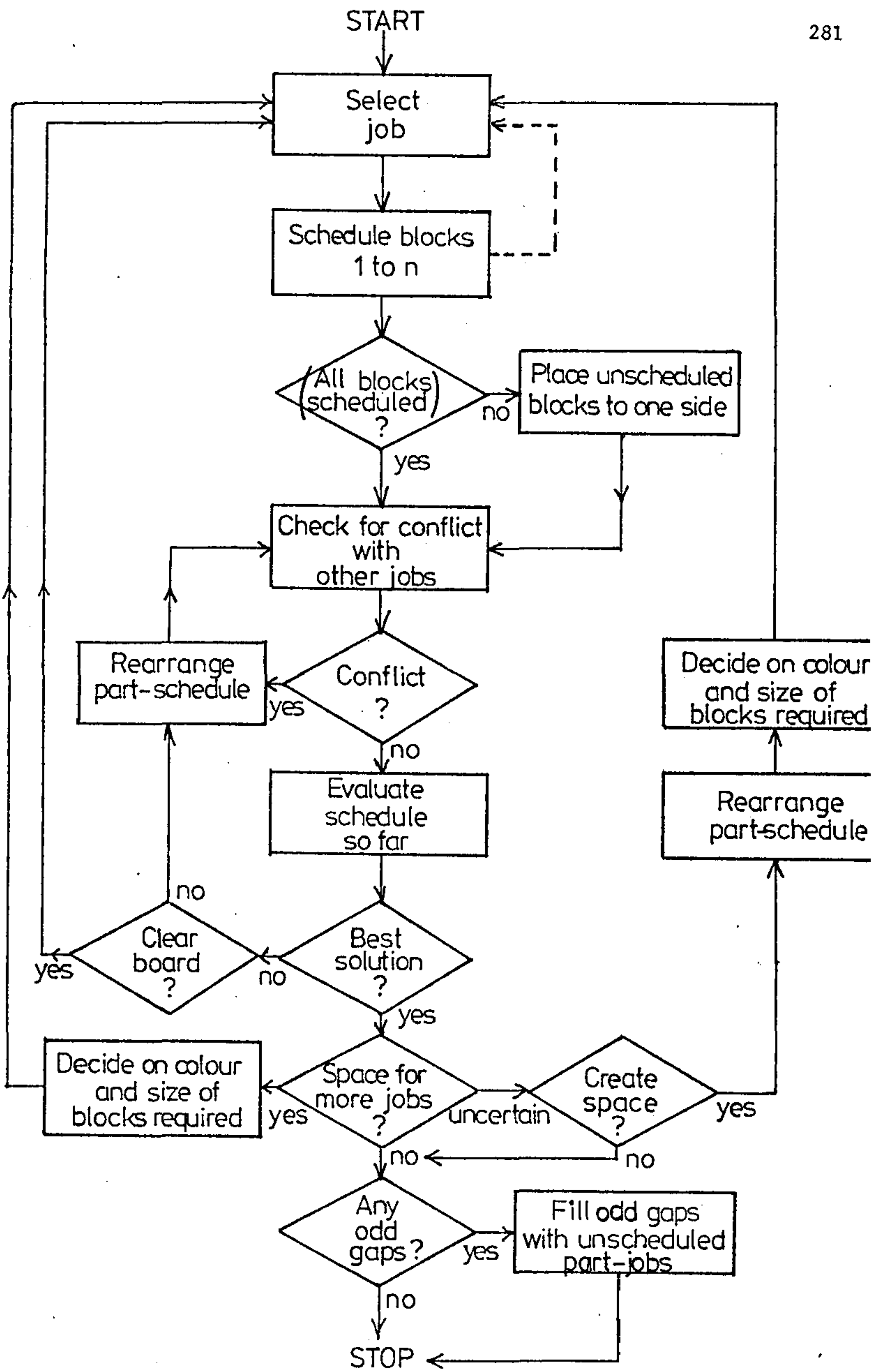


Figure 39: Decision-making strategy in job-shop scheduling problem (full board visible) based on verbal protocols.

of a particular decision. The ability to verbalise implies conscious processing.

Note that separate 'panic' subroutines were present in the soaking pit scheduling task when the ingot arrivals information was updated, and when an ingot actually arrived (Figure 38). In the latter case the standard Select, Test, Evaluate procedure became temporarily abandoned whilst the operator judged the best pit into which the ingot could be loaded, either from the options on the display or from a previously remembered allocation. In some cases the operator was caught unawares, and with no schedule planned on the display had to make an off-the-cuff choice.

5.2 Continuous Tasks

Flow diagrams for the continuous control applications of Chapters 6 and 7 are given in Figures 40-42 and Figures 43-44 respectively. (Figures 43 and 44 are modified from Cininas, 1976, who recorded protocols from an experienced operator on the Batch Kettle.) Again there is a distinct similarity between the protocols for the two tasks. In both cases the subjects follow a standard pattern of Examine value, Judge whether acceptable, Control action if necessary, Switch displays. Separate subroutines (Figure 44) were delineated from the Batch Kettle control task to account for control patterns under temperature and pH alarm conditions, whereas in Figures 40-42 these were incorporated into the main diagrams.

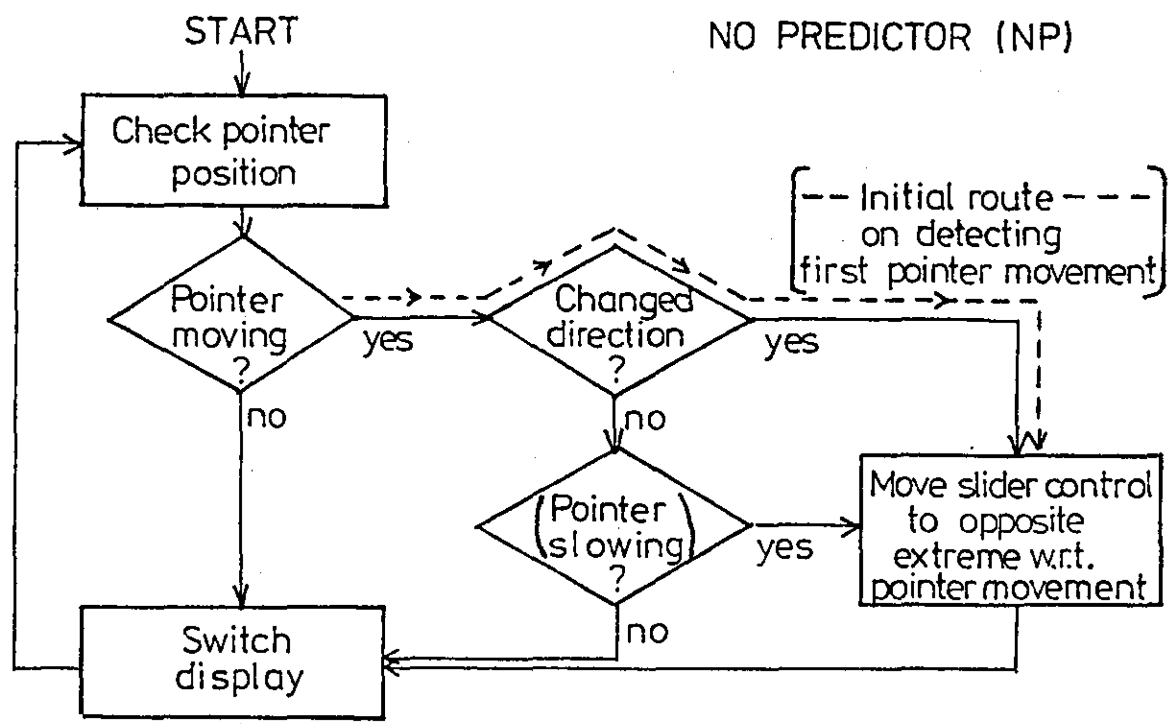


Figure 40: Flow diagram of behaviour in dual-meter monitoring and control task (medium gain and uncertainty) based on verbal protocols - unaided task

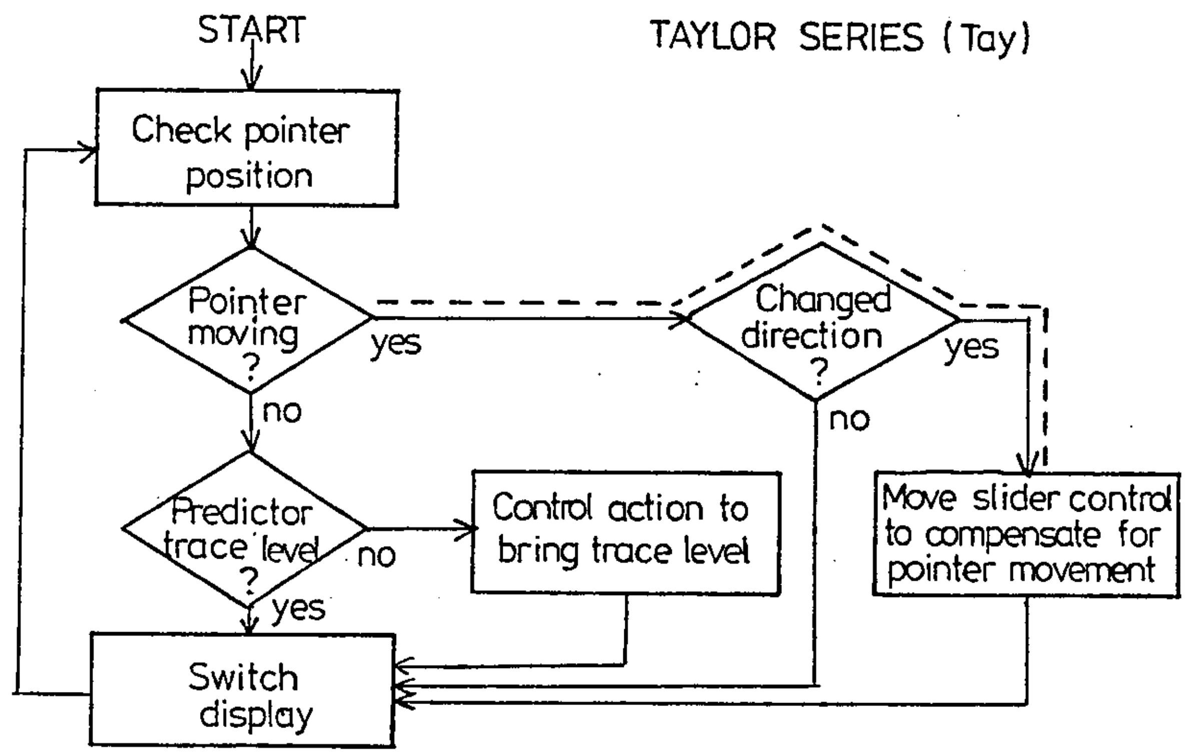


Figure 41: Flow diagram of behaviour in dual-meter monitoring and control task (medium gain and uncertainty) based on verbal protocols - Taylor series extrapolation model (30 seconds span).

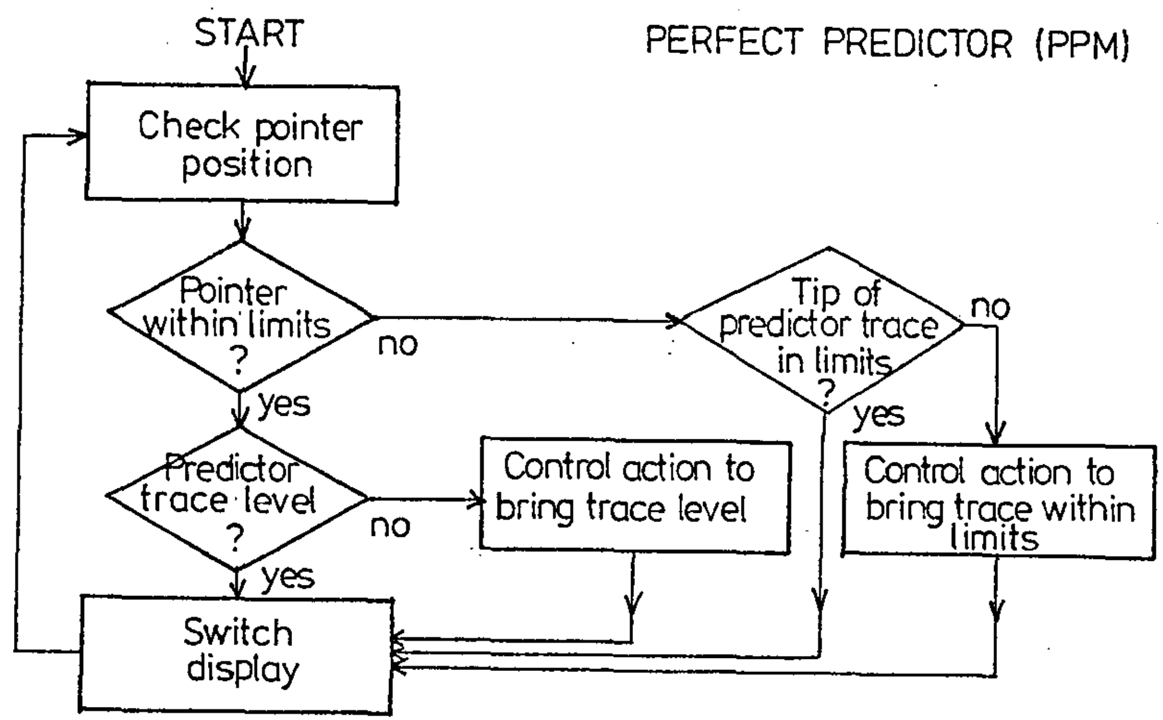
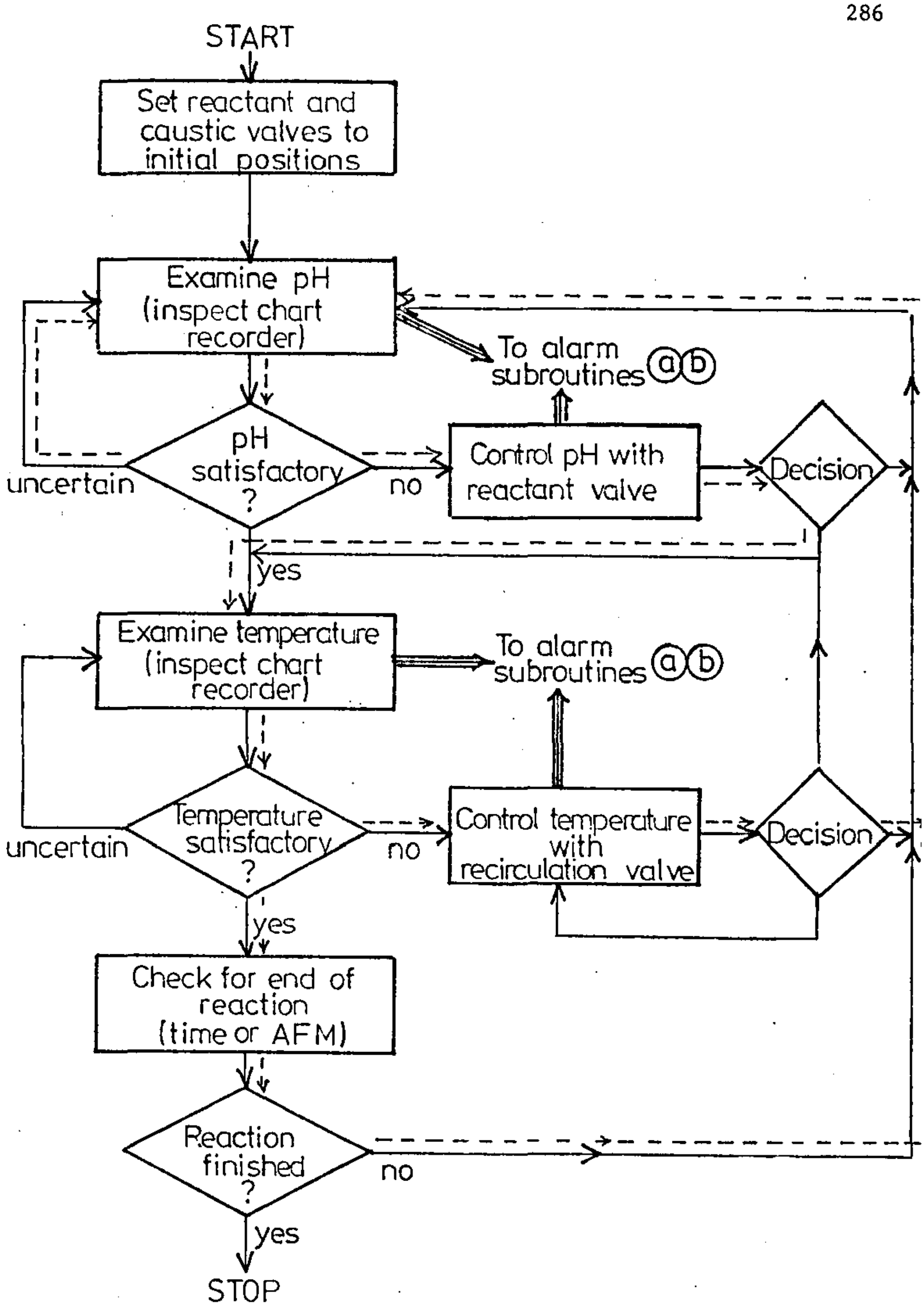


Figure 42: Flow diagram of behaviour in dual-meter monitoring and control task (medium gain and uncertainty) based on verbal protocols - Perfect predictor model (30 seconds span)



Key: ——→—— indicates route most frequently used by the operator during the reaction phase

– –→– – indicates route most frequently used by the operator during the first 800 seconds of the reaction

⟹ indicates interrupt subroutine followed when alarm occurs

<u>Figure 43:</u> <u>Batch kettle control strategy without m.p.r. based on verbal protocols of an experienced operator</u> (from Cininas, 1976).

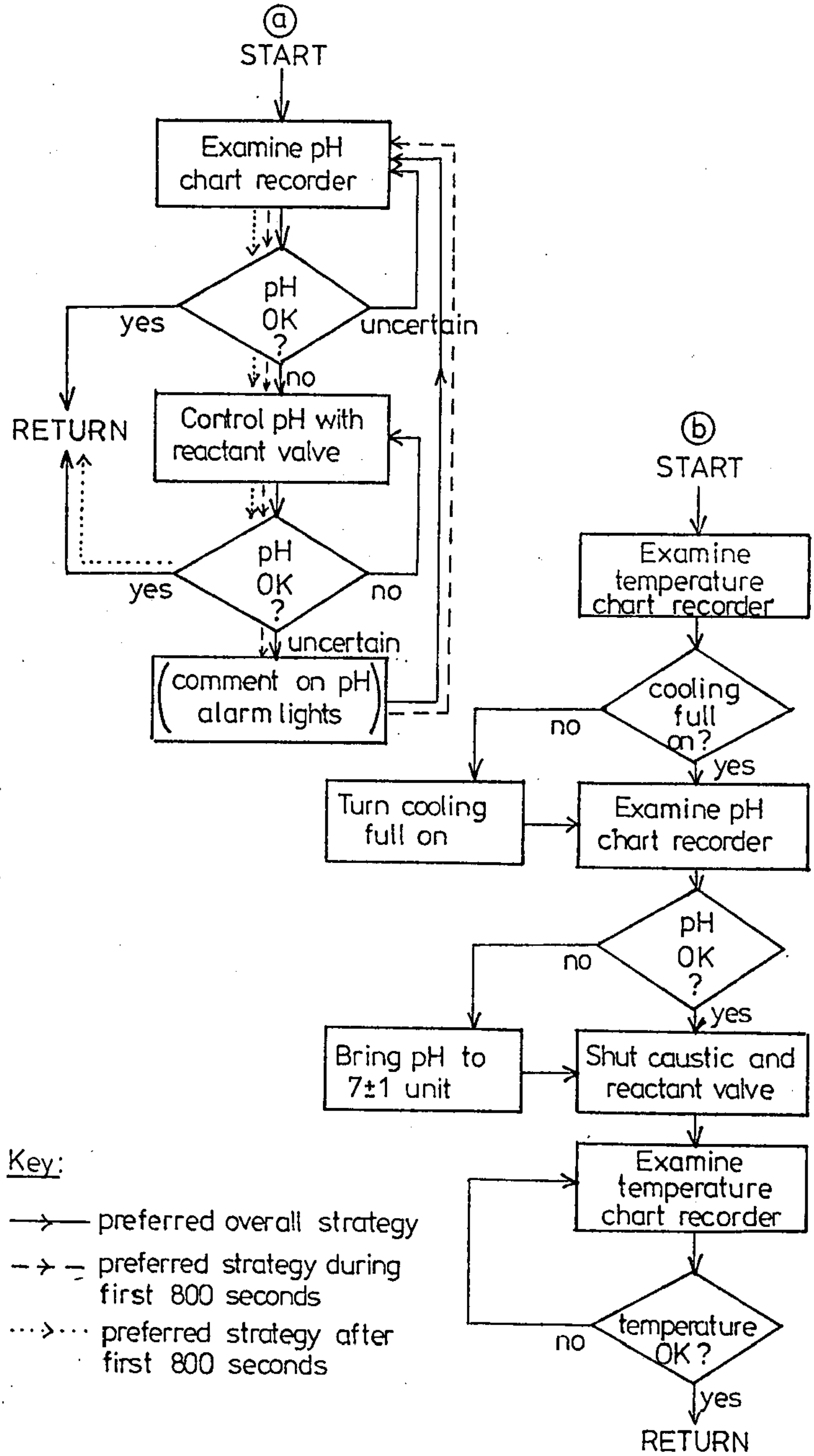


Figure 44: Alarm subroutines from Figure 43 for control of Batch kettle
a) pH alarm conditions, and
b) temperature alarm conditions (from Cininas, 1976)

Perhaps the most interesting feature of Figure 40 is that when controlling without the predictor the subject made control adjustments not only when the pointer was moving or had changed direction, but when it was perceived as slowing. In other words, control actions were being made based on the pointer's predicted point of turn. This suggests that some form of internal model was being used to extrapolate the pointer's path into the future. However, since the subject did not make any verbal reference to the pointer's future position but merely commented whether it was slowing or not, this implies a subconscious extrapolation process.

When the Perfect predictor trace (Figure 42) was introduced, the prediction mechanism was externalised and so the extrapolative judgement "Is the pointer slowing?" could be replaced by the direct perceptual judgements of "Pointer and tip of predictor trace within limits?" and the finer question "Predictor trace horizontal?" This statement was verbalised, implying that its locus of processing was conscious. When the Taylor series prediction model (Figure 41) was introduced, the resultant flow diagram was a compromise between the unaided and Perfect predictor situations. The question "Pointer moving?" coupled with "Changed direction?"was still a predominant part of the behaviour, but rather than judging whether the pointer was slowing, a direct perceptual judgement "Predictor trace level?" could again be made. This statement was also verbalised, again implying that its locus of processing was conscious.

Some mention will be made of the findings from the Batch Kettle protocol study (Cininas, 1976). Protocol evidence from this study suggested that the operator, even though experienced, had considerable difficulty in controlling pH during the first 800 seconds of the reaction since he was overloaded and the pH changed very quickly during this period. This phase is indicated by the dashed line in Figure 43. pH alarms were also a problem throughout the reaction. As noted previously (King and Cininas, 1976), the batch kettle operators did not practice feedforward control since they could not comprehend the process interaction effects. Coupled with the difficulty in controlling pH, this suggested the need for some form of semi-automated external aid, such as the m.p.r. It is therefore surprising that the kettle operators did not report making conscious use of the m.p.r. - there is an obvious discrepancy with the naive subject's protocols from Chapter 6. However, it should be remembered that the Warren Spring operators were highly experienced in controlling the Kettle. It may well have been that since they were already very familiar with the task, then the predictive facility was employed in a subconscious 'checking' mode, which was nonetheless still of sufficient benefit to bring about an improvement in performance. Smith and Crabtree (1975) also noted that their operators with experience used their predictive facility as an error-checking device only.

6. DEVELOPMENT OF A HIERARCHICAL, PREDICTIVE MODEL OF THE HUMAN OPERATOR

6.1 Relation between internal and external predictive models

In order to link the findings of parameter changes in external predictive models to human internal predictive models, it is necessary to assume that the human's internal model is of the same form as an external model which can successfully replace it. Considered without qualification, the implications of such an assumption seem plausible. However, it is a tenuous link, and the evidence for such an assumption must be considered carefully.

When the information on which human judgements are made becomes uncertain, predictions cannot be made with such a degree of confidence and performance deteriorates, as it is not possible to plan so far ahead. This was confirmed by Laios (1975) in his work on unaided decision-making. Similarly, performance would deteriorate if the prediction span of a human internal model were reduced beyond a critical point. For human internal predictive models, prediction span becomes equivalent to decision horizon. In fact, Kelley (1968) has proposed a model in which the operator may adjust his own prediction span. If accurate internal models were adversely affected by uncertainty, one might expect the human operator to adopt a simpler internal model less sensitive to such uncertainty. In fact, this seems to be the case. As was previously noted, operators are known to adopt crude decision heuristics or 'rules of thumb', category classifications of process output, and simplistic control models (Ketteringham et al., 1970; Bainbridge, 1975; King and Cininas, 1976), thus saving on valuable processing capacity. Simpler rules can be delegated to lower cognitive levels. Protocol evidence from the present studies suggests the adoption of a crude internal model both in unaided scheduling tasks (e.g. "if an ingot arrives now, it would fit into pit A; but if it doesn't arrive for 5 minutes it would fit more easily into C") and in unaided continuous tasks (e.g. yo-yo model, inverted pendulum, etc.).

6.2 Necessary components of a human operator model

From previous attempts to model the human operator, and evidence from the present experimental programme, it seems that any adequate representation of the human operator must have the following features:

1) It must be an internal model representation, as in the work of Kelley (1968). Control and decision theory models have already been shown as inadequate to represent the workings of the human mind, though capable of successfully mimicking its output, whereas the internal model has been shown to be a central feature of unaided control and decision-making.
2) There must be a repertoire of internal models for the operator to choose from, depending on the requirements of the situation. Kelley (1968, page 212) has also proposed that the operator should be able to adjust adaptively the parameters of his internal model, e.g. prediction span, to match particular situational requirements and so adapt his behaviour to suit.
3) It must be a hierarchical representation, as suggested by Miller, Galanter and Pribram (1960), Bainbridge (1975a), and Broadbent (1977). Behaviour is known to be organised as a series of global objectives, with goals, sub-goals and routines linked in a flexible and loosely-structured way for achieving these objectives.
4) It must have conscious and sub-conscious components, as suggested by Rasmussen (1974); the higher, conscious processes delegating responsibilities wherever possible to lower level, subconscious processes. The conscious-subconscious distinction was evident in the protocols from the experimental programme.

Taking all these factors together, the model of Figure 45 is proposed. Like so much in psychology, it is not completely original but is rather a composite of previous models by Kelley (1968), Laios (1975), and Rasmussen (1974). In this hybrid model, monitoring is defined as a largely sub-conscious process, involving a crude internal model used to extrapolate future behaviour in the short-term, from sampled values of the process output state(s). The model is accessed from a central store or repertoire of such models. Only when the process is judged to be going outside specification are conscious processes called down to reason out which action to implement in order to rectify the situation. The action is selected from a set or repertoire of potential actions. The internal model is again involved, but this time working to a longer time scale and concerned with the consequences of actions. The choice of internal model is flexible to meet the particular situation, as is the choice of prediction span and other model parameters. When an appropriate control action or decision has been found it is implemented, so closing the loop between man and process.

It is important to stress that this monitor/control loop represents only one level (level 'i', say) in a hierarchy of cognitive activities, ranging from writing poetry at one end of the scale, to reflex arcs at the other (Figure 46). Whenever one level finds itself unable to cope with the processing adequately, it can call on the resources of a higher processing centre: that is, the operator is obliged to 'think' rather more about the situation. However, through the use of standard internal models or 'templates' many of the simpler tasks, such as extrapolation of future values, can be delegated to lower

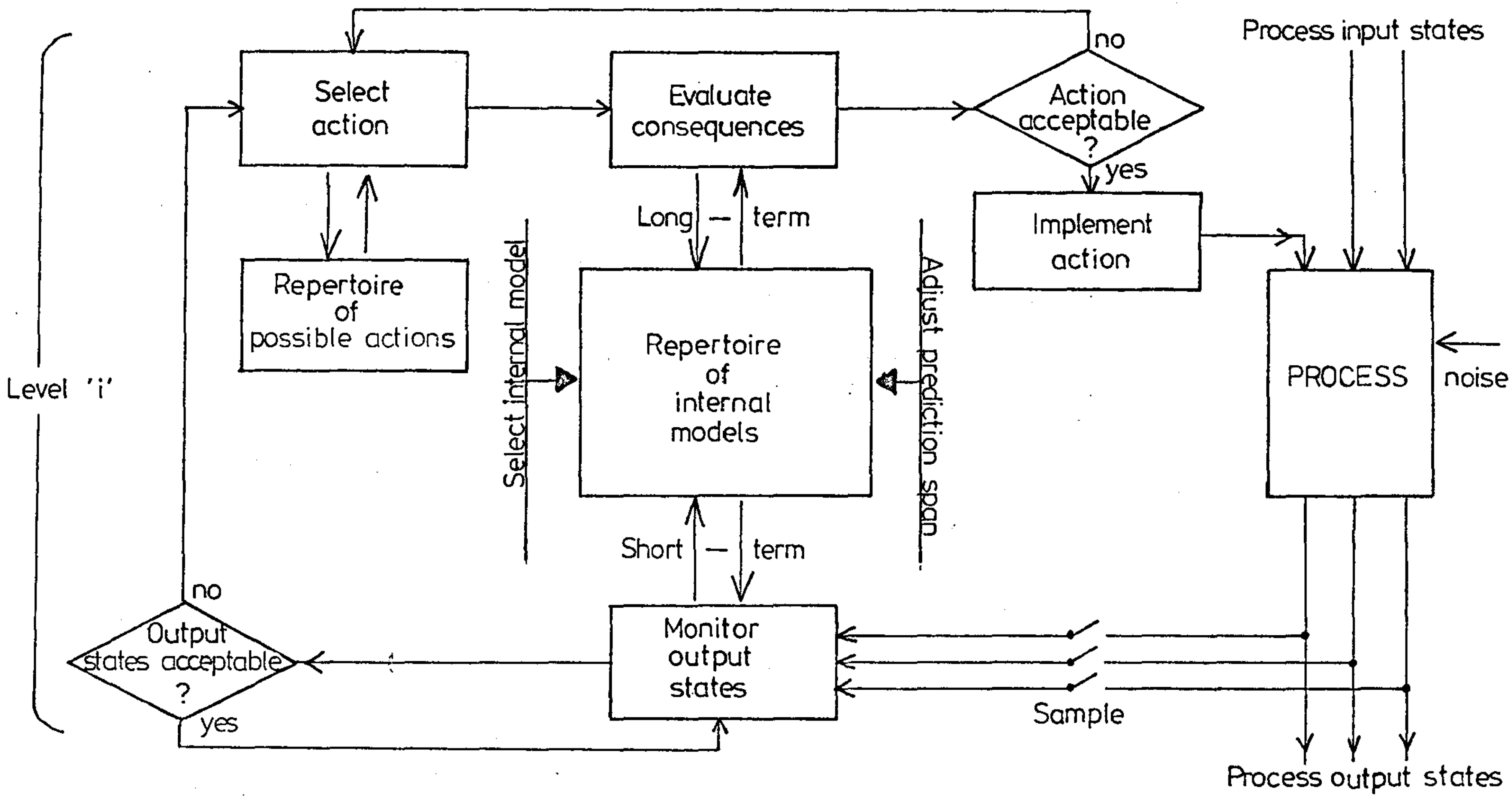


Figure 45: Proposed hybrid model of human operator control and decision-making behaviour

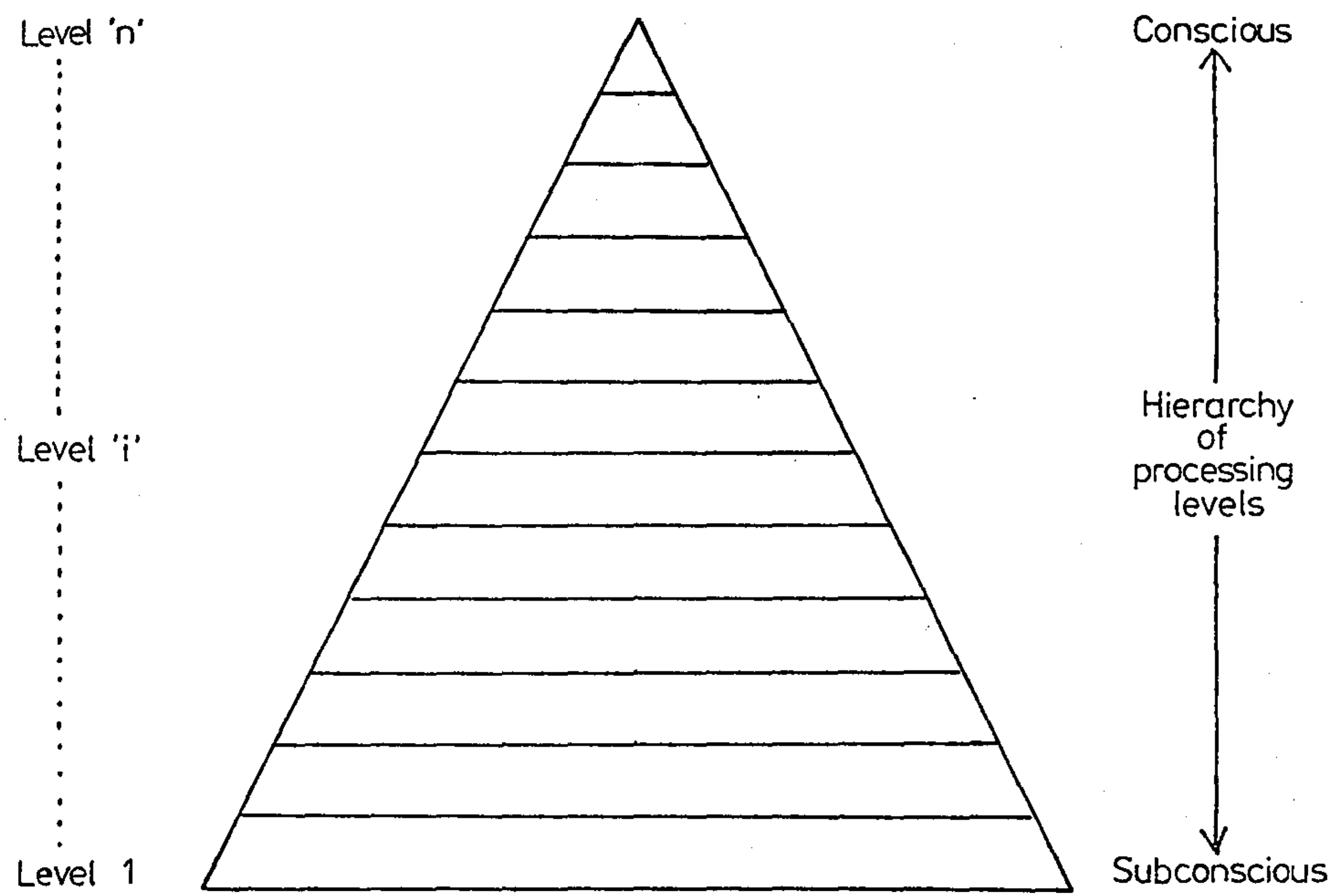


Figure 46: Hierarchy of processing levels (see text)

subconscious levels, leaving the higher centres to be concerned with more intellectually worthwhile pursuits. At the same time the higher levels have access to the information from lower cognitive levels on which to base their decisions. Control thus passes down the hierarchy with experience as the appropriate responses become over-learned, but can revert to higher levels if dictated by the situation. It is also important to stress that in the model there is no rigid conscious/subconscious boundary, but a rather more blurred threshold which can shift depending on the operator's processing load.

6.3 Implications of the proposed model

It is acknowledged that the preceding model can be criticised on the grounds that it is perhaps too simplistic and flexible even to merit the title of a 'model'. Taken no further, the present author believes that it would still be useful in itself as a working hypothesis, as a framework from which to develop future theoretic models. However, if one does accept the foregoing model, what then are its attendant implications for human control and decision-making behaviour? These are best considered taking each feature of the model in turn.

The central feature of the model, its repertoire of internal models, is necessary in order to cope with different situations, or the same situation in different contexts. (Compare approaching an examination from the viewpoint of a student on the one hand, and an examiner on the other. Although the situation is the same, the contexts and therefore the models are different.) Considering the vast number of possible situations in which a person may find himself, it would be impossible to have a specific model or 'template' for each occasion. A single general model may thus have to be adapted to fit different specific situations. Kelley (1968) for example distinguishes between a full internal model incorporating all aspects of a situation, and a simpler derivative model. A danger exists in that the wrong model may be chosen for a particular situation, or at a lesser extreme the inappropriate trimmings may be specified for the correct general model. In this instance a perceptual illusion may occur (Gregory, 1970) with the person's perception of his environment failing to match reality. Witness the commonest indictment of human error: "I thought that xyz was the case ...". The consequences for practical situations such as driving a car or controlling a chemical process could be disastrous, and there seems to be considerable scope for future work in the area of implanting inaccurate models by instruction - how inaccurate a model can the operator tolerate, before his behaviour becomes overtly inappropriate, or before he notices the model mismatch? Clearly, if predictive display research is any indication, there is considerable tolerance of model inaccuracies.

A second feature of the hybrid model as presented is that the store of internal models can be accessed from many different levels. A particular model can be used to extrapolate future process states in the short term, or to evaluate the outcomes of decisions in the longer term. This raises the question of whether several levels can have access to the model store simultaneously (as in the analogy of lines into a telephone switchboard) or whether they must queue for access to a single model (as in the case of two subscribers using a party line). There is, of course, a time-worn and unresolved argument in psychology as to whether mental processing is conducted in parallel or serial mode. If one accepts the logic of the single 'limited channel school, one would expect that a given internal model could only be accessed by one processing level at a time. It is known as a by-product of verbal protocol research that the act of verbalising depresses task performance (Henderson, 1975). A problem exists with task interference experiments, however, in that interference could occur equally well at the processing stage through competition for a central store of internal models or at the sensory stage. A pilot study by the present author suggested that subjects were unable to perform an extrapolative tracking task (short-term use of the internal model store) at the same time as an evaluation-of-outcomes decision-task (long-term use of the store) without mutual interference. This may well be a fruitful area for future research, and could shed some light on the die-hard serial vs. parallel processing argument.

A third feature of the model, its flexible, hierarchical nature, is necessary in that responsibility can be delegated with experience to lower levels wherever possible in order to relieve the processing load at higher levels. This has the potential disadvantage that lower levels are less aware of (and so able to cope with) potential danger signs in the process output, and as such corresponds to the well-known vigilance decrement. If the operator has delegated plant failure recognition to lower levels on the grounds that it never happens, he will be unprepared for it when it does.

Lastly, where do predictive displays fit into this model? Evidently they provide an external model of the problem environment through which future values can be predicted and the consequences of actions evaluated. It is always assumed that predictive displays provide some lightening of the operator's processing load. However, protocol evidence from the experimental programme suggested that the operator's monitoring and (subconscious) extrapolation process was replaced by a (conscious) perceptual judgement relating to the predicted trace when predictive assistance was introduced. According to the model, this implies a higher cognitive level and therefore an increase in processing load. In this sense predictive displays actually increase processing load rather than decrease it. This statement squares with the fact that there is more for the operator to look at, since the task has been brought 'into the open'. It also explains why in particularly complex tasks where the operator is overloaded, the predictive information is often totally ignored. The increased perceptual load imparted by predictive displays is evidently worthwhile in terms of the improved performance scores obtained; and with practice the operator becomes completely familiar with the predictor and the task, so that its use can once more be delegated to lower, subconscious levels. This appeared to be the case with the Warren Spring operators.

In summary, the flexibility of the proposed model is both its strength and its weakness. Human behaviour is flexible in order to cope with a changing world, and any model must itself incorporate this flexibility. At the same time, experimental testing of such a model does become difficult.

CHAPTER 10

CONCLUSIONS

The conclusions from the present programme of research can best be presented as a set of recommendations, together with some suggested areas for future work.

1. RECOMMENDATIONS

Taking the quantitative and qualitative findings of the experimental programme together with previous research, the following recommendations can be made:

* Predictive display systems should be used in preference to conventional alternatives in those control and decision-making applications where process complexity prevents adequate anticipation of events and prediction of consequences by the human operator, and where full automation is not technically feasible or would be undesirable for a variety of reasons.

* Typical applications include aircraft, spacecraft, ship and industrial process control.

* Likely benefits are improved task performance, reductions in training time, and a high degree of user acceptance.

* Prediction should, if at all possible, be directly related to system objectives.

* The distinction between discrete and continuous predictive aids is mainly one of time scales.

* Performance improvements will depend on predictive display parameters and task characteristics. These are detailed as follows:

* Uncertainty in the form of unreliable input information or signal contamination by noise will generally reduce predictive display effectiveness, even where uncertainty has been designed into the display.

* The choice of prediction span is influenced by the prediction model fidelity, the process complexity (gain) and the level of input uncertainty. Interactions are present between these parameters.

* In general, a crude prediction model has longer useful prediction spans than an accurate model, though for high gain systems an upper limit to the prediction span is recommended on practical grounds.

* A complex prediction model will have its useful range of prediction spans reduced by input uncertainty, particularly with high gain systems, as operators make full use of the display they are given. The reduction should in practice be proportional to the level of uncertainty and the responsiveness of the system.

* The prediction model should be chosen to match the process concerned. Even a crude prediction model may be superior to an operator's own forecasts.

* For slow response (low gain) systems, there is little to choose between a complex and a crude process model. The latter is simple to implement in practice, and has been shown to be cost-effective in bringing about financially worthwhile improvements in a pseudo real-world operational process.

* Further development work is now needed to exploit the predictive display concept in practical situations.

2. SUGGESTED AREAS OF FUTURE WORK

Two basic areas of future work are seen as necessary. Firstly, there is now a need to develop and commercially exploit the predictive display concept in a wide range of applications and operational settings. The pseudo-industrial study at Warren Spring has given some idea of the potential benefits, but there is a long way to go before predictor displays make an impact on systems design. In particular the widespread use of microprocessors seems an ideal basis for compact prediction models, perhaps using Kalman filters or similar techniques (Page, 1978 - in forthcoming IERE Conference proceedings). It is to be hoped that some enterprising manufacturer will take up the idea and market it as part of future generations of information display systems.

Secondly, a tentative theoretical hybrid model has been proposed of the human operator's control and decision-making behaviour, but this obviously needs further refinement and testing. It will be of no mean satisfaction to the author if the present thesis can be used as a working hypothesis to stimulate future work in this area. In particular, the question of how inaccurate an operator's internal process model can be before control and decision-making errors become evident seems especially worth investigating.

3. EPILOGUE

It has been demonstrated that predictor displays have much to recommend them in the instrumentation of industrial process plant. Their application is well suited to any area where process characteristics combine with man's inherent limitations to render unaided control difficult. A classic man-computer symbiosis thus emerges, with computer extrapolation compensating for man's weakness in predicting complex responses. Although predictor displays have been around for 20 years they have not been applied for two reasons. Firstly, like so many sound but unproven ideas, they have come to be regarded as a kind of sophisticated toy. This tag is surely undeserved. Secondly, they have not been implemented because their use contradicts Birmingham and Taylor's (1954) design philosophy, a philosophy which has for too long dominated human factors thinking. As Poulton (1974) has noted: "The fashions of design engineers are perhaps not very different from the fashions of designers in other fields. However, they may cost lives when they lead to the wrong design decisions". It is to be hoped that today's designers will seriously consider the adoption of predictive techniques in the design of future operator-instrument interfaces.

REFERENCES

* Indicates reprinted in Edwards and Lees (1974).

ADAMS, P.G. and SCHOOLEY, A.T. (1968)

Adaptive predictive control of a batch reaction. In Proceedings of 23rd Instrument Society of America Annual Conference, New York. Advances in Instrumentation, 23, 818, 1-8.

AGNEW, N.M. and PYNE, S.W. (1969)

The Science Game: an introduction to research in the behavioral sciences. New Jersey: Prentice Hall.

ANDERSON, N.H. (1974)

Information integration theory: a brief survey. In Krantz, D.H. (ed.) Contemporary developments in mathematical psychology - Measurement, psychophysics and neural information processing. San Francisco: Freema

ANGEL, E.S. and BEKEY, H.A. (1968)

Adaptive finite state models of manual control systems. IEEE Transactions on Man-Machine Systems, MMS-9, 9, 15-20.

ANYAKORA, S.N. and LEES, F.P. (1972)

Detection of instrument malfunction by the process operator. The Chemical Engineer, 264, 304-309. *

ARNOLD, J.E. and BRAISTED, P.W. (1963)

Design and evaluation of a predictor for remote control systems operating with signal transmission delays. NASA Technical Note TN D-2229.

ATKINSON, P. (1968)

Feedback control theory for engineers. London: Heineman Educational Books.

ATTWOOD, D. (1970)

The interaction between human and automatic control. In Bolam, F. (ed.) Paper-making systems and their control. London: British Paper and Board Makers Association. *

AUSTIN, R.W. and RYKEN, J.M. (1963)

Trajectory control and energy management of lifting re-entry vehicles Advances in the Astronautical Sciences, 16, 7, 847-866.

BAKER, F.G. (1967)

The internal organisation of computer models of cognitive behaviour. Behavioural Science, 12, 156.

BAKER, J.D. and GOLDSTEIN, I. (1966)

Batch vs. Sequential displays: effects on human problem-solving. Human Factors, 8, 3, 225-236.

BAINBRIDGE, L. (1967)

A discrete predicting controller. British Journal of Mathematical and Statistical Psychology, 20, 1, 31-50.

BAINBRIDGE, L. (1971)

The influence of display-type on decision-making strategy. In Displays, IEE Conference Record No. 80, p.209-215. London: Institution of Electrical Engineers.

BAINBRIDGE, L. (1972)

An analysis of a verbal protocol from a process control task. Unpublished Ph.D. Thesis, University of Bristol.

BAINBRIDGE, L. (1974)

Analysis of verbal protocols from a process control task. In Edwards, E. and Lees, F.P. (eds.) The human operator in process control. London: Taylor and Francis.

BAINBRIDGE, L. (1975a)

The process controller. In Singleton, W.T. (ed.) The study of real skills. London: Academic Press.

BAINBRIDGE, L. (1975b)

The representation of working storage and its use in the organisation of behaviour. In Singleton, W.T. and Spurgeon, P. (eds.) Measurement of human resources. London: Taylor and Francis.

BAINBRIDGE, L. (1978)

Forgotten alternatives in skill and work-load. Ergonomics, 21, 3, 169-185.

BAINBRIDGE, L., BEISHON, J., HEMMING, J.H. and SPLAINE, M. (1968)

A study of real-time human decision-making using a plant simulator. Operational Research Quarterly, 19, 91-106. *

BALL, R.G., LLOYD, R.B. and ORD, G. (1975)

Interactive conflict resolution ATC. AGARD: 20th Guidance and Control Panel Meeting/Symposium; Cambridge, Massachusetts.

BARTLETT, F.C. (1932)

Remembering.

Cambridge: Cambridge University Press.

BATY, D.L. (1976)

Evaluating a CRT map predictor for airborne use. IEEE Transactions on Systems, Man and Cybernetics, SMC-6, 3, 209-215.

BAUM, A.S. and DRURY, C.G. (1976)

Modelling the human process controller. International Journal of Man-Machine Studies, 8, 1-11.

BEACH, B.H. (1975)

Expert judgements under uncertainty: Bayesian decision-making in realistic settings. Organisational Behavior and Human Performance, 14, 10-59.

BECKER, G.M. and McCLINTOCK, C.G. (1967)

Value: behavioral decision theory. Annual Review of Psychology, 18, 239-286.

BEISHON, R.J. (1966)

A study of some aspects of mental skill in the performance of laboratory and industrial tasks. Unpublished D.Phil. Thesis, University of Oxford.

BEISHON, R.J. (1969)

An analysis and simulation of an operator's behaviour in controlling continuous baking ovens. In Bresson, F. and de Montmollin, M. (eds.) The simulation of human behaviour. Paris: Dunod. *

BELLMAN, R. (1957)

Dynamic programming. Princeton: Princeton University Press.

BERBERT, A.G. and KELLEY, C.R. (1962)

Piloting nuclear submarines with controls that look into the future. Electronics, 35, 35-39.

BERNOTAT, R.K. (1972)

Prediction displays based on the extrapolation method. In Bernotat, R.K. and Gärtner, K.P. (eds.) Displays and Controls. Proceedings of NATO Advanced Study Institute, Berchtesgaden, F.R. Germany. Amsterdam: Swets and Zeitlinger, N.V.

BERNOTAT, R.K. and WIDLOK, H. (1966)

Principles and applications of prediction display. Journal of the Institute of Navigation, 19, 3, 361-370.

BESCO, R.O. (1964)

Manual attitude control systems. Vol II: Display format considerations. Culver City, California: Hughes Aircraft Corporation Technical Report to NASA.

BIBBY, K.S. (1974)

Human and organisational problems of production control. In Proceedings of IFAC Workshop on Productivity and Man, Bad Boll, Germany, 35-38.

BIRD, J.M. and WHITFIELD, D. (1975)

Some evaluation studies of predictive aids for human controllers. Applied Psychology Department, University of Aston in Birmingham, A.P. Note 46.

BIRMINGHAM, H.P. and TAYLOR, F.V. (1954)

A design philosophy for man-machine control systems. Proceedings of American Institute of Radio Engineers, 42, 1748-1758.

BLAKE, B.M. (1968)

Selecting the right technique. Control Engineering, 15, 6, 88-93.

BODE, H.W. and SHANNON, C.E. (1950)

A simplified derivation of linear least square smoothing and prediction theory. Proceedings of American Institute of Radio Engineers, 38, 417-425.

BOERSMA, F.J., DE JONGE, J.T. AND STELLWAGEN, W.R. (1964)

A power comparison of the F and L tests. Psychological Review, 71, 505-513.

BOLLENBACHER, R. (1970)

Computer eases small parts scheduling. Industrial Engineering, 2, 10, 48-50.

BREWER, A.C. (1971)

Interactive scheduling system. IBM Systems Journal, 10, 63-79.

BRIGHAM, F.R. (1972)

Ergonomics problems in ship control. Applied Ergonomics, 3, 1, 14-19.

BRIGHAM, F.R. (1974)

A study of decision-making problems in six control tasks. Department of Human Sciences, Loughborough University of Technology. HUSAT Memo No. 61.

BRIGHAM, F.R. and LAIOS, L. (1975)

Operator performance in the control of a laboratory process plant. Ergonomics, 18, 1, 53-66.

BROADBENT, D.E. (1977)

Levels, hierarchies and the locus of control. Quarterly Journal of Experimental Psychology, 29, 181-201.

BRUNER, J. (1957)

On perceptual readiness. Psychological Review, 64, 123-152.

BURGHART, J.H. and LEFKOWITZ, I. (1969)

A technique for on-line, steady-state optimizing control. IEEE Transactions on Systems, Science and Cybernetics, SSC-5, 2, 125-132.

BURTON, K. (1976)

An investigation of possible selection tests for process controllers. Unpublished undergraduate thesis, Department of Human Sciences, Loughborough University of Technology.

CHAPANIS, A. (1967)

The relevance of laboratory studies to practical situations. Ergonomics, 10, 557-578.

CHESTNUT, H., SOLLECITO, W.E. and TROUTMAN, P.H. (1961)

Predictive-control system application. Transactions of American Institute of Electrical Engineers: Applications and Industry, 80, 128-139.

CININAS, A. (1975a)

The batch reactor plant and process. Warren Spring Laboratory, Stevenage. Internal Report IR 270 (CON).

CININAS, A. (1975b)

Operator performance assessment on the batch kettle plant. Warren Spring Laboratory, Stevenage. Internal Report IR 304 (CON).

CININAS, A. (1976)

An analysis of a verbal protocol from an operator on the WSL batch kettle plant. Warren Spring Laboratory, Stevenage. Internal Report IR 314 (CON).

COCHRAN, W.G. (1947)

Some consequences when assumptions for the analysis of variance are not satisfied. Biometrics, 3, 22-38.

COHEN, D. (1962)

Prediction-marker computers and display systems for controlling a lunar robot vehicle from earth. Long Island, New York. Airborne Instruments Laboratory Technical Memorandum.

COHEN, H.S. and FERRELL, W.R. (1969)

Human operator decision-making in manual control. IEEE Transactions on Man-Machine Systems, MMS-10, 2, 41-53.

COOKE, J.E. (1965)

Human decisions in the control of a slow-response system. Unpublished D.Phil. Thesis, University of Oxford.

CRAIK, K.J.W. (1943)

The nature of explanation. Cambridge: Cambridge University Press.

CRAIK, K.J.W. (1947)

Theory of the human operator in control systems I - The operator as an engineering system. British Journal of Psychology, 38, 56-61.

CRAIK, K.J.W. (1948)

Theory of the human operator in control systems II - Man as an element in a control system. British Journal of Psychology, 38, 142-148.

CRAWLEY, J.E. (1968)

The present and future contribution of the human operator to the control of LD steelmaking. In Proceedings of British Iron and Steel Research Association Symposium on Steelmaking. Battersea: British Steel Corporation. *

CROSSMAN, E.R.F.W. and COOKE, J.E. (1962)

Manual control of slow-response systems. In International Congress on Human Factors in Electronics, Long Beach, California. *

CROSSMAN, E.R.F.W., COOKE, J.E. and BEISHON, R.J. (1964)

Visual attention and the sampling of displayed information in process control. University of California, Berkeley, California, Human Factors in Technology Research Group Report HFT-64-11-7. *

CURRY, R.E. and GAI, E.G. (1976)

Detection of random process failures by human monitors. In Sheridan, T.B. and Johannsen, G. (eds.) Monitoring behavior and supervisory control. London: Plenum Publishing Company.

DALE, H.G.A. (1968)

Weighing evidence: an attempt to assess the efficiency of the human operator. Ergonomics, 11, 215-230.

DAVIS, R.H. and BEHAN, R.A. (1960)

A study of visual prediction behavior. Human Factors, 2, 228-234.

DEY, D. (1971)

The influence of a prediction display on the quasilinear describing function and remnant measured with an adaptive analog-pilot in a closed loop. In Proceedings of Seventh Annual NASA-University Conference on Manual Control, University of Southern California, Los Angeles. NASA SP-281.

DEY, D. (1972)

Results of the investigation of different extrapolation displays. In Bernotat, R.K. and Gärtner, K.P. (eds.) Displays and Controls. Proceedings of NATO Advanced Study Institute, Berchtesgaden, F.R. Germany. Amsterdam: Swets and Zeitlinger, N.V.

DEY, D. and JOHANNSEN, G. (1969)

Stabilisierung und lenkung von fahrzeugen mit hilfe der voranzeige. Institut für Flugführung und Luftverkehr, Technische Universitat Berlin. Report No. 50.

DIXON, W.J. (ed.) (1971)

BMD biomedical computer programs. Berkeley, California: University of California Press.

DRURY, C.G. and BAUM, A.S. (1976)

Manual process control: a case study and a challenge. Applied Ergonomics, 7, 1, 3-9.

DUPUIS, L.J. (1975)

Computer operator interface for steel plant application. In Preprints of Sixth International Federation of Automatic Control World Congress, Massachusetts. Volume IV A, Session 5.1, 1-8.

EBERT, R.J. (1972)

Human control of a two-variable decision system. Organisational Behavior and Human Performance, 7, 237-264.

EDWARDS, E. and LEES, F.P. (1973)

Man and computer in process control. London: Institution of Chemical Engineers.

EDWARDS, E. and LEES, F.P. (eds.) (1974)

The human operator in process control. London: Taylor and Francis.

EDWARDS, W. (1961)

Behavioral decision theory. Annual Review of Psychology, 12, 473-98.

EDWARDS, W. (1968)

Conservatism in human information processing. In Kleinmuntz, B. (ed.) Formal representation of human judgement. New York: Wiley.

EDWARDS, W. (1972)

Dynamic aspects of decision-making. In Proceedings of Twentieth International Congress of Psychology, Tokyo.

EDWARDS, W., PHILLIPS, L.D., HAYS, W.L. and GOODMAN, B.C. (1968)

Probabilistic information processing systems: design and evaluation. IEEE Transactions on Systems Science and Cybernetics, SSC-4, 3, 248-265.

ELKIND, J.I. and MILLER, D.C. (1967)

The adaptive response of the human controller to sudden changes in controlled process dynamics. IEEE Transactions on Human Factors in Electronics, HFE-8, 3, 218-223.

FARGEL, L.C. and ULBRICH, E.A. (1963)

Aerospace simulators prove predictor displays extend manual operation. Control Engineering, 10, 57-60.

FERGUSON, R.L. and JONES, C.H. (1969)

A computer aided decision system. Management Science, 15, 10 B, 550-561.

FLOWERS, K.A. (1978)

The predictive control of behaviour: appropriate and inappropriate actions beyond the input in a tracking task. Ergonomics, 21, 2, 109-122.

FOGEL, L.J., OWENS, A.J. and WALSH, M.J. (1966)

Artificial intelligence through simulated evolution. New York: Wiley.

FOOT, H.C. (1969)

Visual prediction of the point of coincidence of two moving targets. Ergonomics, 12, 5, 723-734.

FROST, G. (1972)

Man-machine dynamics. In van Cott, H.P. and Kinkade, R.G. (eds.) Human engineering guide to equipment design. Washington, D.C.: United States Government Printing Office, Joint Army-Navy-Air Force Steering Committee.

GAINES, B.R. (1969)

Linear and non-linear models of the human controller. International Journal of Man-Machine Studies, 1, 333-360.

GAINES, B.R. (1976)

Foundations of fuzzy reasoning. International Journal of Man-Machine Studies, 8, 6 (Special Issue), 623-668.

GALLAHER, P.D., HUNT, R.A. and WILLIGES, R.C. (1976)

A regression approach to generate aircraft predictor information. In Proceedings of Twelfth Annual NASA-University Conference on Manual Control, University of Illinois. NASA TMX-73, 170.

GARNER, W.R. (1962)

Uncertainty and structure as psychological concepts. New York: Wiley.

GIBSON, R. (1978)

Personal communication.

GIBSON, R. and LAIOS, L. (1978)

The presentation of information to the job-shop scheduler. Human Factors, in press.

GIFFORD, R.N. and LYMAN, J. (1967)

Tracking performance with advanced and delayed visual displays. Human Factors, 9, 2, 127-132.

GILCHRIST, J.D. and SOLAND, D.E. (1967)

A manual optimal guidance scheme using a predictive model. Paper presented at American Institute of Aeronautics and Astronautics Guidance, Control and Flight Dynamics Conference, Huntsville, Alabama.

GILCHRIST, W. (1976)

Statistical forecasting. London: Wiley.

GODIN, V. and JONES, C. (1969)

The interactive shop supervisor.

Industrial Engineering, 1, 11, 16-22.

GOLDBERG, L.R. (1968)

Simple models or simple processes? Some research on clinical judgements. American Psychologist, 23, 483-496.

GOTTSDANKER, R.M. (1952)

The accuracy of prediction motion. Journal of Experimental Psychology, 43, 26-36.

GOTTSDANKER, R.M. (1955)

A further study of prediction motion. American Journal of Psychology, 68, 432-437.

GOTTSDANKER, R.M. (1956)

The ability of human operators to detect acceleration of target motion. Psychological Bulletin, 53, 477-487.

GOTTSDANKER, R., FRICK, J.W. and LOCKARD, R.B. (1961)

Identifying the acceleration of visual targets. British Journal of Psychology, 52, 1, 31-42.

GREEN, B.F. (1968)

Descriptions and explanations: a comment on papers by Hoffman and Edwards. In Kleinmuntz, B. (ed.) Formal representation of human judgement. New York: Wiley.

GREENHOUSE, S.W. and GEISSER, S. (1959)

On methods in the analysis of profile data. Psychometrika, 24, 95-112.

GREGORY, R.L. (1970)

On how little information controls so much behaviour. Ergonomics, 13, 1, 25-35.

GREGORY, R.L. (1973)

Visual perception. Oxford: University Press, Oxford Biology Readers Series.

HAIDER, S.W., BUCK, F.R. and MOODIE, C.L. (1977)

Man-computer vs. automatic scheduling. In Proceedings of the Human Factors Society 21st Annual Meeting, Santa Monica, California.

HAINSWORTH, T. and OLINGER, M. (1958)

Safe take-off predictor (STOP). Lear Inc., Advanced Engineering Report No. 8.

HALL, D.J. (1970)

Man-machine projects at SRI. International Journal of Man-Machine Studies, 2, 363-394.

HAMMERTON, M. and TICKNER, A.H. (1970)

The effects of temporary obscuration of the target on a pursuit tracking task. Ergonomics, 13, 723-725.

HARTNETT, O.M. and MURRELL, K.F.H. (1973)

Some problems of field research. Applied Ergonomics, 4, 4, 219-221.

HEBB, D.O. (1949)

The organisation of behavior. New York: Wiley.

HENDERSON, A.I. (1975)

A note on the interference between verbal output and a cognitive task. Acta Psychologica, 39, 495-497.

HERMAN, L.M., ORNSTEIN, G.N. and BAHRICK, H.P. (1964)

Operator decision performance using probabilistic displays of object location. IEEE Transactions on Human Factors in Electronics, HFE-5, 13-19.

HOLLISTER, W.M. (1967)

An analytic measure for the difficulty of human control. Journal of the Institute of Navigation, 20, 2, 167-175.

HOWELL, W.C. and BRIGGS, G.E. (1959)

The effects of visual noise and locus of perturbation on tracking performance. Journal of Experimental Psychology, 58, 166-173.

JOHNSON, E.M. and BAKER, J.D. (1974)

Field testing: the delicate compromise. Human Factors, 16, 3 (Special Issue), 203-214.

JONES, C.H., HUGHES, J.L. and ENGOLD, K.J. (1970)

A comparative study of management decision-making from computer terminals. In Proceedings of the AFIPS Spring Joint Computer Conference, Atlantic City, New Jersey.

de JONG, J.J. and KÖSTER, E.P. (1971)

The human operator in the computer controlled refinery. In Proceedings of Eighth World Petroleum Congress, Moscow. London: Institute of Petroleum. *

JORDAN, N. (1963)

Allocation of function between men and machines in automated systems. Journal of Applied Psychology, 47, 3, 161-165.

KAHNEMAN, D. and TVERSKY, A. (1973)

On the psychology of prediction. Psychological Review, 80, 4, 237-2[illegible]

KALMAN, R.E. (1960)

A new approach to linear filtering and prediction problems. Journal of Basic Engineering, Transactions of American Society of Mechanical Engineers, Series D, 82, 35-45.

KAPLAN, M.F. and SCHWARTZ, S. (eds.) (1975)

Human judgement and decision processes. New York: Academic Press.

KELLEY, C.R. (1958)

A predictor instrument for manual control. Paper presented at Eighth Annual Office of Naval Research Human Engineering Conference, Ann Arbor, Michigan.

KELLEY, C.R. (1960a)

Developing and testing the effectiveness of the 'predictor' instrument. Dunlap and Associates Inc. Technical Report TR-252-60-1 to Office of Naval Research.

KELLEY, C.R. (1960b)

Further research on the 'predictor' instrument. Dunlap and Associates Inc. Technical Report TR-252-60-2 to Office of Naval Research.

KELLEY, C.R. (1962)

Predictor instruments look into the future. Control Engineering, 9, 3, 86-90.

KELLEY, C.R. (1963)

Manual remote control of a roving lunar surface vehicle. Santa Monica, California: Dunlap and Associates Inc., Technical Memorandum.

KELLEY, C.R. (1968)

Manual and automatic control. New York: Wiley.

KELLEY, C.R. (1969)

The measurement of tracking proficiency. Human Factors, 11, 1, 43-64.

KELLEY, C.R. (1972)

Manual vs. automatic control. In Bernotat, R.K. and Gärtner, K.P. (eds.) Displays and Controls. Proceedings of NATO Advanced Study Institute, Berchtesgaden, F.R. Germany. Amsterdam: Swets and Zeitlinger, N.V.

KELLEY, C.R., BEGGS, J.D. and PROSIN, D.J. (1973)

Automatic control plus manual equals automanual control. Control Engineering, 20, 5, 49-51.

KELLEY, C.R., MITCHELL, M.B. and STRUDWICK, P.H. (1964)
Application of predictor displays to the control of space vehicles. Santa Monica, California, Dunlap and Associates. Technical Report to NASA Office of Manned Space Flight.

KELLEY, C.R. and PROSIN, D.J. (1972)
Adaptive display using prediction. In Bernotat, R.K. and Gärtner, K.P. (eds.) Displays and Controls. Proceedings of NATO Advanced Study Institute, Berchtesgaden, F.R. Germany. Amsterdam: Swets and Zeitlinger, N.V.

KEMP, G.G. (1969)
A VTOL prediction display. Unpublished Masters thesis, Massachusetts Institute of Technology.

KENNEDY, R.S., SMITH, R.L., WULFECK, J.W., QUEEN, J.E. and BURGER, W.J. (1975)
Two studies of predictor displays for jet aircraft landings. 1. Optimisation of control stick dynamics and prediction span parameters. 2. A comparison between experienced and novice navy F-4 pilot simulated landing performances. Pacific Missile Test Center, Pt. Mugu, California. Report PMTC-TP-75-55 (AD-A021 925).

KETTERINGHAM, P.J.A. and O'BRIEN, D.D. (1974)
A simulation study of computer-aided soaking pit scheduling. In Edwards, E. and Lees, F.P. (eds.) The human operator in process control. London: Taylor and Francis.

KETTERINGHAM, P.J.A., O'BRIEN, D.D. and COLE, P.G. (1970)
A computer-based interactive display system to aid steel plant scheduling. In Man-Computer Interaction, IEE Conference Record No. 68 London: Institution of Electrical Engineers.

KING, P.J. (1975)

The dynamics of a partially simulated semi-batch reactor. Warren Spring Laboratory, Stevenage. Report CR 1026 (CON).

KING, P.J. and CININAS, A. (1976)

Manual control strategy determination using questionnaires. Warren Spring Laboratory, Stevenage. Report LR 236 (CON).

KING, P.J. and CININAS, A. (1978)

Performance evaluation and information display in process control. In Proceedings of Symposium on the Operator-instrument interface, Middlesbrough. Teeside: Institute of Measurement and Control.

KING, P.J. and MAMDANI, E.H. (1975)

The application of fuzzy control systems to industrial processes. Warren Spring Laboratory, Stevenage. WSL Paper 109/3/621. Presented to Sixth International Federation of Automatic Control World Congress, Massachusetts.

KING, P.J. and RAY, D.J. (1972)

Digital computer makes data collection easy. Process Engineering (July), 38-39.

KIRK, R.E. (1968)

Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks-Cole Publishing Company.

KLEINMAN, D.L., BARON, S. and LEVISON, W.H. (1971)

A control-theoretic approach to manned-vehicle systems analysis. IEEE Transactions on Automatic Control AC-16, 6, 824-832.

KLUBNIKIN, P.F. (1966)

A non-linear self-adjusting system with linear prediction. In Hammond, P.H. (ed.) Theory of self-adaptive control systems. New York: ISA/Plenum Press. (Proceedings of Second IFAC Symposium on Self-adaptive control systems, Teddington, Middlesex.)

KRAGT, H. and LANDEWEERD, J.A. (1974)

Mental skills in process control. In Edwards, E. and Lees, F.P. (eds. The human operator in process control. London: Taylor and Francis.

KREIFELDT, J.G. and WEMPE, T. (1973)

Pilot performance during a simulated standard instrument procedure turn with and without a predictor display. Paper presented at Ninth Annual NASA-University Conference on Manual Control, Massachusetts Institute of Technology, Cambridge.

KREIFELDT, J.G. and WEMPE, T.E. (1974)

Future terminal air-traffic management concepts. In Proceedings of the Tenth Annual NASA-University Conference on Manual Control, Wright Patterson Air Force Base, Dayton, Ohio.

LAIOS, L. (1975)

Human decision-making and predictive aids in discrete multi-stage decision tasks. Unpublished Ph.D. Thesis, Department of Human Science Loughborough University of Technology.

LAIOS, L. (1976)

The effects of uncertainty on human performance in a multi-stage decision task. International Journal of Man-Machine Studies, 8, 465-476.

LAIOS, L. (1978)

Predictive aids for discrete decision tasks with input uncertainty. IEEE Transactions on Systems, Man and Cybernetics, SMC-8, 1, 19-29.

LAIOS, L. and GIBSON, R. (1976)

A planning computer display for job-shop scheduling. Department of Human Sciences, Loughborough University of Technology. HUSAT Memo No. 120.

LEE, W. (1971)

Decision theory and human behaviour. New York: Wiley.

LEES, F.P. (1974)

Research on the process operator. In Edwards, E. and Lees, F.P. (eds The human operator in process control. London: Taylor and Francis.

LEVINE, J.M. and SAMET, M.G. (1973)

Information seeking with multiple sources of conflicting and unreliable information. Human Factors, 15, 4, 407-419.

LICKLIDER, J.C.R. (1960)

Man-computer symbiosis. IRE Transactions on Human Factors in Electronics, HFE-1, 4-10.

LUCE, R.D. and SUPPES, P. (1965)

Preference, utility and subjective probability. In Luce, R.D., Bush, R.R. and Galanter, E. (eds.) Handbook of mathematical psychology 3, 249-410. New York: Wiley.

LUYBEN, W.L. (1973)

Process modelling, simulation and control for chemical engineers. New York: McGraw-Hill.

MANO, N.M. and ULBRICH, E.A. (1965)

Co-planar adapredictive rendezvous simulation. North American Aviation Space and Information Systems Division, Technical Report No. SID-65-660.

MASON, G.W. (1976)

Computer and work scheduling for small firms. Machinery and Production Engineering, 130, 3352, 242-248.

McCORMICK, E.J. (1976)

Human factors in engineering and design, 4th edition. New York: McGraw-Hill.

McCOY, W.K. Jr. and FROST, G.G. (1965)

Investigation of 'predictor' displays for orbital rendezvous: program summary. Wright-Patterson Air Force Base, Dayton, Ohio. Aerospace Medical Research Laboratories Report AMRL-TR-65-138 (AD-624 467).

McCOY, W.K. Jr. and FROST, G.G. (1966)

Predictor display techniques for on-board trajectory optimisation of rendezvous maneuvers. Wright-Patterson Air Force Base, Dayton, Ohio. Aerospace Medical Research Laboratories, Technical Report AMRL-TR-66-60.

McEWING, R.W. (1977)

Man-computer systems simulations in production and process control. In Proceedings of Symposium on Human operators and simulation, Loughborough. London: Institute of Measurement and Control.

McLANE, R.C. and WOLF, J.D. (1966)

Symbolic and pictorial displays for submarine control. Paper presented at NASA-MIT Working Conference on Manual Control, Cambridge, Massachusetts.

McLANE, R.C. and WOLF, J.D. (1967)

Symbolic and pictorial displays for submarine control. IEEE Transactions on Human Factors in Electronics, HFE-8, 2, 148-158.

McLEOD, P.D. (1972)

Recovery strategy during temporary obscuration of a tracked target. Ergonomics, 15, 57-64.

McRUER, D.T. and JEX, H. (1967)

A review of quasi-linear pilot models. IEEE Transactions on Human Factors in Electronics, HFE-8, 3, 231-248.

MILLER, G.A. (1956)

The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 63, 81-97.

MILLER, G.A., GALANTER, E. and PRIBRAM, K.H. (1960)

Plans and the structure of behaviour. London: Holt, Rinehart and Winston.

MILNER, P.M. (1971)

Physiological psychology. New York: Holt, Rinehart and Winston.

MONTGOMERY, D.C. and JOHNSON, L.A. (1976)

Forecasting and time series analysis. New York: McGraw-Hill.

MURRELL, H. (1976)

Men and machines. London: Methuen Essential Psychology Series.

NEWELL, A. and SIMON, H.A. (1972)

Human problem solving. New Jersey: Prentice Hall.

OATLEY, K. (1972)

Brain mechanisms and mind. London: Thames and Hudson.

O'BRIEN, J.J. (1969)

Scheduling handbook. New York: McGraw-Hill.

OLIVERSON, R.L. (1971)

Planning and scheduling boards. Plant Engineering, 8, 45-48.

OLSHAVSKY, S. (1971)

Search limits as a function of tree size and storage requirements. Organisational Behaviour and Human Performance, 6, 336-344.

OPPENHEIM, A.N. (1966)

Questionnaire design and attitude measurement. London: Heinemann Educational Books.

ORD, G. and WHITFIELD, D. (1977)

Small scale simulations in the development of new concepts for air traffic control. In Proceedings of Symposium on Human operators and simulation, Loughborough. London: Institute of Measurement and Control.

PAGE, M. (1978)

Microprocessor implementation of the Kalman filter. In forthcoming IERE Conference on Microprocessors in Automation and Communications, University of Kent at Canterbury.

PATERNOTTE, P.H. (ed.) (1976)

Short survey of research activities. Eindhoven University of Technology, Eindhoven, Netherlands. Department of Industrial Engineering, Research Group on Man-Machine Systems, Report No. 14.

PETERSON, C.R. and BEACH, L.R. (1967)

Man as an intuitive statistician. Psychological Bulletin, 68, 1, 29-46.

PETIT, J. and FAVREL, J. (1976)

Conception and implementation of an information system for production control. Automatisme, 21, 11, 347-353. (In French.)

PEW, R.W., DUFFENDACK, J.C. and FENSCH, L.K. (1966)

On the relationship between discrete and continuous performance. In IEEE Symposium on Human Factors in Electronics, Minneapolis.

PHATAK, A.V. (1976)

Formulation and validation of optimal control theoretic models of the human operator. Man-Machine Systems Review, 2, 11-12.

PITRELLA, F. (1978)

Personal communication.

POULTON, E.C. (1952)

The basis of perceptual anticipation in tracking. British Journal of Psychology, 43, 295-302.

POULTON, E.C. (1957)

On prediction in skilled movements. Psychological Bulletin, 54, 6, 467-478.

POULTON, E.C. (1964)

Postview and preview in tracking with complex and simple inputs. Ergonomics, 7, 3, 257-266.

POULTON, E.C. (1969)

Bias in experimental comparisons between equipments due to the order of testing. Ergonomics, 12, 4, 332-344.

POULTON, E.C. (1972)

The environment at work. Applied Ergonomics, 3, 1, 24-29.

POULTON, E.C. (1973)

Bias in ergonomic experiments. Applied Ergonomics, 4, 1, 17-18.

POULTON, E.C. (1974)

Tracking skill and manual control. London: Academic Press.

POULTON, E.C. and GREGORY, R.L. (1952)

Blinking during visual tracking. Quarterly Journal of Experimental Psychology, 4, 57-65.

POUNDS, W.F. (1963)

The scheduling environment. In Industrial Scheduling. New Jersey: Prentice Hall.

PRICE, H.E., HONSBERGER, W.D. and ERENETA, W.J. (1966a)

A study of potential roles of supersonic transport crews and some implications for the flight deck. Volume I: Workload, crew roles, flight deck concepts and conclusions. NASA Report CR-561.

PRICE, H.E., HONSBERGER, W.D. and ERENETA, W.J. (1966b)

A study of potential roles of supersonic transport crews and some implications for the flight deck. Volume II: Feasible automated and manual implementation concepts for SST activities and functions. NASA Report CR-562.

PROSIN, D.J. and KELLEY, C.R. (1973)

Improving operator performance with predictive displays. Instruments and Control Systems, 46, 9, 93-95.

RAPOPORT, A. (1966)

A study of human control in a stochastic multi-stage decision task. Behavioral Science, 11, 18-32.

RAPOPORT, A. and WALLSTEN, T.S. (1972)

Individual decision behavior. Annual Review of Psychology, 23, 131-176.

RASMUSSEN, J. (1968)

On the communication between operators and instrumentation in automatic process plants. Danish Atomic Energy Research Establishment Risö, Denmark. Report Risö-M-686. *

RASMUSSEN, J. (1974)

The human data processor as a system component - bits and pieces of a model. Danish Atomic Energy Commission, Research Establishment, Risö, Denmark. Report Risö-M-1722. (ISBN 87 550 0268 4).

RASMUSSEN, J. (1976)

Outlines of a hybrid model of the process plant operator. In Sheridan, T.B. and Johannsen, G. (eds.) Monitoring behaviour and supervisory control. London: Plenum Publishing Company.

RASMUSSEN, J. and JENSEN, A. (1974)

Mental procedures in real-life tasks: a case study of electronic trouble-shooting. Ergonomics, 17, 3, 293-307.

RATCLIFFE, B.J. (1977)

Stock control of a continuous process using a conversational predictiv model and a VDU. In Displays for Man-Machine Systems, IEE Conference Record No. 150, 75-77. London: Institution of Electrical Engineers.

REITMAN, W.R. (1965)

Cognition and thought: an information processing approach. New York: Wiley.

RICE, M. (1969)

Man plus computer the best scheduling system. Industrial Engineering, 1, 7, 31-35.

ROSENTHAL, R. (1966)

Experimenter effects in behavioral research. New York: Appleton Century Crofts.

ROUSE, W.B. (1970)

An application of predictor displays to air traffic control problems. Masters Thesis, Dept. of Mechanical Engineering, Massachusetts Institute of Technology. NASA Report CR-111372 (N-71-11466).

ROUSE, W.B. (1973a)

Models of man as a sub-optimal predictor. In Proceedings of Ninth Annual NASA-University Conference on Manual Control, Cambridge, Massachusetts.

ROUSE, W.B. (1973b)

A model of the human in a cognitive prediction task. IEEE Transactions on Systems, Man and Cybernetics, SMC-3, 5, 473-477.

ROUSE, W.B. (1977)

Special issue preface - applications of control theory in human factors. Human Factors, 19, 4, 313-314.

SARKAR, G.A. (1972)

Scheduling system for an X-ray department. Industrial Engineering, 4, 12, 20-23.

SENDERS, J.W. (1964)

The human operator as a monitor and controller of multi-degree-of-freedom systems. IEEE Transactions on Human Factors in Electronics, HFE-5, 1-6.

SHACKEL, B. (1976)

Process control - simple and sophisticated display devices as decision aids. In Sheridan, T.B. and Johannsen, G. (eds.) Monitoring Behavior and Supervisory Control. London: Plenum Publishing Company.

SHACKEL, B., GOILLAU, P.J. and LAIOS, L. (1976)

Improvements in or relating to display apparatuses. U.K. Patent Application No. 14750/76.

SHACKEL, B. and KLEIN, L. (1976)

Esso London Airport refuelling control centre redesign - an ergonomics case study. Applied Ergonomics, 7, 1, 37-45.

SHANNON, G.F. (1976)

The use of predictive displays illustrated by a lunar docking. Australian Institution of Engineers, Electrical Engineering Transactions, EE-12, 1, 1-3.

SHEPHERD, A. (1977)

Fidelity of simulation for training control panel diagnosis. In Proceedings of Symposium on Human operators and simulation, Loughborough. London: Institute of Measurement and Control.

SHERIDAN, T.B. (1962)

The human operator in control instrumentation. In Macmillan, R.H., Higgins, T.J. and Naslin, P. (eds.) Progress in Control Engineering, 1, 141-188. London : Heywood and Company.

SHERIDAN, T.B. (1966)

Three models of preview control. IEEE Transactions on Human Factors in Electronics, HFE-7, 2, 91-102.

SHERIDAN, T.B. and FERRELL, W.R. (1974)

Man-machine systems. Cambridge, Massachusetts: MIT Press.

SHERIDAN, T.B. and JOHANNSEN, G. (eds.) (1976)

Monitoring behavior and supervisory control. Proceedings of NATO International Symposium, Berchtesgaden, F.R. Germany. London: Plenum Publishing Company.

SHERIDAN, T.B., JOHNSON, W.M., BELL, A.C. and KREIFELDT, J.G. (1964)

Control models of creatures which look ahead. In Proceedings of IEEE Fifth National Symposium on Human Factors in Electronics, San Diego, California.

SHERIDAN, T.B. and ROUSE, W.B. (1971)

Supervisory sampling and control: sources of suboptimality in a prediction task. In Proceedings of Seventh Annual NASA-University Conference on Manual Control, University of Southern California, Los Angeles, NASA SP-281.

SINGLETON, W.T. (1972)

General theory of presentation of information. In Bernotat, R.K. and Gärtner, K.P. (eds.) Displays and Controls. Proceedings of NATO Advanced Study Institute, Berchtesgaden, F.R. Germany. Amsterdam: Swets and Zeitlinger, N.V.

SLOVIC, P., FISCHHOFF, B. and LICHTENSTEIN, S. (1977)

Behavioral decision theory. Annual Review of Psychology, 28, 1-39.

SLOVIC, P. and LICHTENSTEIN, S. (1971)

Comparison of Bayesian and regression approaches to the study of information processing in judgement. Organisational Behavior and Human Performance, 6, 649-744.

SMALLWOOD, R.D. (1967)

Internal models and the human instrument monitor. IEEE Transactions on Human Factors in Electronics, HFE-8, 3, 181-187.

SMITH, H.T. (1976)

Perceptual organisation and the design of the man-computer interface in process control. In Sheridan, T.B. and Johannsen, G. (eds.) Monitoring Behavior and Supervisory Control. London: Plenum Publishing Company.

SMITH, H.T. and CRABTREE, R.G. (1975)

Interactive planning: a study of computer aiding in the execution of a simulated scheduling task. International Journal of Man-Machine Studies, 7, 213-231.

SMITH, R.L. and KENNEDY, R.S. (1975)

Predictor displays: history, research and applications. Inglewood, California, Dunlap and Associates Inc. Report N00123-74-C-0952, to Human Factors Engineering Branch, Pacific Missile Test Centre, Pt. Mugu, California.

SMITH, R.L., PENCE, G.G., QUEEN, J.E. and WULFECK, J.W. (1974)
Effect of a predictor instrument on learning to land a simulated jet trainer. Inglewood, California: Dunlap and Associates Inc., Report AD A000 586.

SMITH, R.L. and QUEEN, J.E. (1975)
Use of a predictor display in landing a simulated remote piloted vehicle (RPV). Inglewood, California: Dunlap and Associates Inc., Technical Report.

SPENCER, J. (1961)
Estimating averages. Ergonomics, 4, 4, 317-328.

STASSEN, H.G. (1975)
Man-machine systems: manual and supervisory control. In Michon, J.F. (ed.) Handbook for psychonomy. Zwolle, Netherlands: Kluwer Publishing Company.

STŘÍŽENEC, M. (1976)
Mental models of problems and their experimental investigation. Studia Psychologica, 18, 2, 125-130.

SWEENEY, J.S., TODD, N.C. and HEATON, E.C. (1965)
Studies in predictor display technique. North American Aviation Inc., Aeronautics Division Report C5-1819/311.

TAINSH, M.A. (1977)
The operator's estimation of a track's statistical parameters in a simulated process control task. Admiralty Research Laboratory, Teddington, Middlesex. Report APU/H23/1/71/77.

TELFORD, C.W. (1931)

Refractory phase of voluntary and associative responses. Journal of Experimental Psychology, 14, 1-35.

THOMAS, M.U. (1973)

A human response model of a combined manual and decision task. IEEE Transactions on Systems, Man and Cybernetics SMC-3, 5, 478-484.

TOLMAN, E.C. (1948)

Cognitive maps in rats and men. Psychological Review, 55, 4, 189-20{

TOMIZUKA, M. (1975)

On the design of adaptive servosystems. Transactions ASME Journal of Dynamic Systems - Measurement and Control, 6, 326-332.

TOMIZUKA, M. and WHITNEY, D.E. (1975)

Optimal discrete finite preview problems - why and how is future information important? Transactions of ASME, Journal of Dynamic Systems, Measurement and Control, 6, 319-325.

TONG, R.M. (1977)

A control engineering approach to fuzzy systems. British Steel Corporation, Battersea. BSC Report BL/CF/7/77.

TOUTENHOOFD, H. (1974)

An introduction to human prediction models. Unpublished thesis, Ergonomics Laboratory, Twente University of Technology, Netherlands. (In Dutch.)

TRIGG, D.W. and LEACH, A.G. (1967)

Exponential smoothing with an adaptive response rate. Operational Research Quarterly, 18, 53-59.

UMBERS, I.G. (1976)

A study of cognitive skills in complex systems. Unpublished Ph.D. Thesis, University of Aston in Birmingham.

van COTT, H.P. and KINKADE, R.G. (eds.) (1972)

Human engineering guide to equipment design. Revised edition. Washington, D.C.: United States Government Printing Office, Joint Army-Navy-Air Force Steering Committee.

van HEUSDEN, A.R. (1977)

Models for human prediction of process variables. In Proceedings of Symposium on Human operators and simulation, Loughborough. London: Institute of Measurement and Control.

VAUGHAN, W.S. and MAVOR, A.S. (1972)

Behavioural characteristics of men in the performance of some decision-making task components. Ergonomics, 15, 3, 267-277.

VELDHUYZEN, W. and STASSEN, H.G. (1977)

The internal model concept: an application to human control of large ships. Human Factors, 19, 4, 367-380.

VERHAGEN, L. (1976)

Computer programs for data management and experimental control of ergonomic experiments with a simulated process. Tutorial paper presented to NATO Advanced Study Institute on Man-Computer Interaction, Mati, near Athens, Greece.

WARNER, J.D. (1969)

A fundamental study of predictive display systems. University of Michigan. Ann Arbor, Michigan. NASA Report CR-1274.

WEMPE, T.E. (1974)

Flight management - pilot procedures and system interfaces for the 1980-1990's. In Proceedings of American Institute of Aeronautics and Astronautics, Life Sciences and Systems Conference, Arlington, Texas. AIAA Paper No. 74-1297.

WHITFIELD, C. and GOILLAU, P.J. (1978)

Human operators and simulation - a report on the Symposium. Measurement and Control, 11, 3, 99-102.

WHITFIELD, D. (1975)

Man-computer symbiosis: a 1975 review. Applied Psychology Department, University of Aston in Birmingham, A.P. Report 57.

WHITFIELD, D. (1977)

A review of recent developments in man-computer symbiosis. Ergonomics, 20, 5, 561. (Abstract of paper presented to Ergonomics Society Annual Conference, University of Aston in Birmingham.)

WIENER, E.L. (1962)

Motion prediction as a function of target speed and duration of presentation. Journal of Applied Psychology, 46, 6, 420-424.

WIENER, N. (1948)

Cybernetics. New York: MIT Press/Wiley.

WIENER, N. (1949)

Extrapolation, interpolation and smoothing of stationary time series. Cambridge, Massachusetts: MIT Press/New York: Wiley.

WIERWILLE, W. (1964)

Improvement of the human operator tracking performance by means of optimum filtering and prediction. IEEE Transactions on Human Factors in Electronics, HFE-5, 1, 20-24.

WILKINSON, W.L. (1972)

A brief exposition on man-computer dynamic scheduling systems. George Washington University, Institute for Management Science and Engineering, Program in Logistics, Serial T-272. (TRC Report T73-00982.)

WILLIAMS, W.L. (1969)

Predictor displays for low-altitude, high-speed flight. Wright-Patterson Air Force Base, Dayton, Ohio. Aerospace Medical Research Laboratory, Report No. AMRL-TR-68-118.

WINER, B.J. (1971)

Statistical principles in experimental design. 2nd edition. Kogakusha, Japan: McGraw-Hill.

WINGROVE, R.C. and COATE, R.E. (1961)

Piloted simulation tests of a guidance system which can continuously predict the landing point of a low L/D vehicle during atmospheric re-entry. Moffett Field, California: NASA Technical Note TN D-787.

WULFECK, J.W., PROSIN, D.J. and BURGER, W.J. (1973)

Effect of a predictor display on carrier landing performance. Part I: experimental evaluation. Inglewood, California: Dunlap and Associates Inc., Report prepared for Office of Naval Research, AD 767 982.

WUNDT, W. (1874)

Grundzüge der physiologischen psychologie. Leipzig: Engelmann.

YNTEMA, D.B. and MEUSER, G.E. (1960)

Remembering the present state of a number of variables. Journal of Experimental Psychology, 60, 18-22.

YNTEMA, D.B. and TORGENSON, W.S. (1961)

Man-computer co-operation in decisions requiring common sense. IRE Transactions on Human Factors in Electronics, HFE-2, 20-26.

YOUNG, L.R. (1969)

On adaptive manual control. Ergonomics, 12, 4, 635-675.

ZIEBOLZ, H. and PAYNTER, H.M. (1954)

Possibilities of a two-time scale computing system for control and simulation of dynamic systems. Proceedings of U.S. National Electronics Conference, 9, 215-223.

APPENDICES

STRUCTURE OF APPENDICES A1.1

Glossary of predictive display terminology A2.1-2.3

EXPERIMENT 1 (Chapter 3)

Instructions to Subjects A3.1-3.3

Scheduling Error Scores A3.4

Predictive Activity Data A3.5

Decision Horizon Results A3.6

EXPERIMENT 2 (Chapter 4)

Instructions to Subjects A4.1-4.3

Machine utilisation, jobs unfinished and scheduling time scores A4.4

List of jobs available to be scheduled A4.5-4.6

EXPERIMENT 3 (Chapter 5)

Instructions to Subjects A5.1

Detailed Instructions A5.2

Total amount of product manufactured in the various experimental conditions A5.3

EXPERIMENT 4 (Chapter 6)

Instructions to Subjects A6.1

Taylor Series Data A6.2

PPM Data A6.3

Questionnaire presented after each trial A6.4

EXPERIMENT 5 (Chapter 7)

The Warren Spring Batch Kettle Plant A7.1

The Batch Kettle Control Panel A7.2

Listing of Taylor series subroutine implemented on Argus 500 computer A7.3

Operator performance scores with prediction on temperature A7.4

Operator performance scores with prediction on pH A7.5

Details of profit figure calculation A7.6

Glossary of predictive display terminology

Continuous process - a system where the material flows are unbroken (e.g. continuous chemical plant, vehicle control, some steel production techniques).

Discrete process - a system where distinct items or stages are evident (e.g. production scheduling, soaking pit scheduling, air-traffic control).

Input uncertainty - a measure of the accuracy of the updating information fed to the prediction model. Determined by the degree of signal contamination due to noise, the unreliability of input information, or simply the normal variability inherent in system operation.

Internal model (mental model) - a process operator's internal representation of the system he is controlling, built up over time from experience. An operator's internal model may be supplemented, or even replaced, by an external predictive aid.

Mode of control - refers to whether the predictive display operates in an on-line or off-line configuration, exploratory control and supervisory control being special cases of the latter category.

Moving average - a mathematical technique for smoothing noisy data. Works by generating an average value over the past, say, ten data points as a basis for prediction of future values.

Perfect predictor model (PPM) - a hypothetical prediction model which matches the behaviour of the controlled system under all eventualities. Cannot be achieved with present technology, but is approached by Class 2 and 3 fast-time models.

Prediction model - a mathematical representation of the system to be controlled and through which future values of system output(s) are calculated. May be implemented by digital or analog means.

Prediction model fidelity - the accuracy with which the prediction model represents the system under control. Usually expressed in terms of Bernotat and Widlok's (1966) 3 stage classification.

Prediction span (extrapolation interval) - the real-time period over which the predicted plant response is displayed.

Prediction time - the real-time interval over which the predicted plant response is computed. Where a single predicted end-point is displayed, prediction span and prediction time are equivalent.

Predictive (predictor) display - a control or decision aid based on an external predictive model of the system under control, and displaying the consequences on system output(s) of an operator's control actions or decisions.

Process dynamics/response characteristics - for continuous systems, this is usually expressed in terms of the plant gain (K) and the order of the controlled process. In discrete applications, response characteristics are a function of the complexity and speed of the process.

Repetition (refresh) rate - the number of successive predictions displayed to the operator per unit of time.

Smoothing - the reduction of noise contamination of a signal, usually applied to plant variables before prediction can take place.

Taylor series extrapolator model (Tay) - a relatively simple mathematical extrapolation technique based on the Taylor series expansion. Depending on the number of derivative terms employed, can project a straight line or curve from a number of past data points. Often used to generate a single predicted end-point.

Updating frequency - the frequency at which the prediction model is updated with the current state of the plant.

Instructions to Subjects

Thank you for agreeing to be a subject in this experiment. I am interested in the usefulness of computer aids for human decision-making in discrete scheduling tasks such as occur in the steel industry. In this experiment your job will be to make a series of decisions to schedule the output from a simulated steel plant which has been represented on the computer.

I'll begin by explaining the simulated system (Figure 3). Imagine that a series of steel ingots (or 'casts') arrive randomly at a set of four soaking pits, where they are heated to the required temperature before passing onto the next stage of the process. Unfortunately the four soaking pits vary in efficiency, so that whereas the first two pits (we'll call them A and B) can each raise an ingot to the required temperature in 10 minutes (simulated time), the next pit (pit C) takes 15 minutes and the final pit (pit D) takes 20 minutes to do the same work. After an ingot has reached the required temperature it takes a further 3 minutes to unload it from the soaking pit, and you should allow for this by not reloading that pit during this emptying period.

As the controller of the soaking pits your job is to try and assign ingots to the soaking pits so that a regular output of heated casts is achieved - a rate of about one cast every 5 minutes should be your target, though one minute either side would still be reasonable. What you must try to avoid is the situation where two pits are ready to be emptied at the same time. I is also important, though less so, to load an ingot as soon as possible afte its arrival. Is there anything you would like to ask about the process at this stage?

Next I shall describe how the task is represented graphically on the compute screen. The computer display (Figure 4) is organised as follows: on the left hand side of the screen you can see the four pits (A, B, C and D), and time scale extending to the right as you look further into the future. The simulation starts at time 0 and the time scale moves along to the left one minute at a time as the simulation proceeds. Current time is shown in the bottom right of the screen. The display allows you to 'game play' with the computer and so work out different allocation strategies of ingots to soakin pits up to 35 minutes ahead of current time. By examining the computer predictions of when the soaking pits will empty you can see at a glance whether your target of one heated cast leaving every 5 minutes is likely to be reached.

Underneath the bottom time scale you will probably notice some rows of crosses. Each row of crosses represents the arrival of one ingot within the interval marked by the crosses against the time scale. As in real life it is often the case that the exact arrival time is not known, since it depends on a number of factors outside your control. The time interval within which an ingot will definitely arrive can be forecast, however, and it is this interval which is represented on the screen by a row of crosses. So an ingot will arrive at a time corresponding to one of the crosses in each row, and though there is a tendency for it to arrive towards the middle of an interval, it could equally well arrive at either end. Your job is to guess when an ingot will arrive and schedule it into one of the four soaking pits. The rows of crosses move to the left with the time scale, but at times 10, 20, 30 and 40 the intervals are updated and become more accurate i.e. narrower.

The soaking pits can be loaded by pressing the appropriate push-buttons on the 'black box' in front of you. There are two ways of loading a soaking pit: a 'test load' when you are trying out different combinations of pits ahead of current time and 'system loads' when an ingot has arrived at current time and requires to be loaded. Test loads can be cleared and a different combination of soaking pits tried, whereas system loads cannot be cleared.

When you have guessed on which of the crosses in a row an ingot will arrive, you can test load it into pit A, B, C or D at the corresponding arrival time by pressing one of the <u>orange</u> buttons. For example:

Test load $\boxed{A}$ at time $\boxed{2\ 0}$ $\boxed{\text{ENTER}}$

The number 20 appears in the bottom left of the screen as you type it. When the 'ENTER' button is pressed a single bar appears on the display in pit A, starting at time 20 and extending 10 minutes ahead. The square shape at the end of the bar (at time 30) tells you when soaking pit A will be ready for emptying. The length of the bar corresponds to the time each soaking pit needs to heat an ingot. If pit B had been chosen, for example, the bar would also have extended 10 minutes ahead, but in pit C it would have extended 15 minutes and in pit D 20 minutes. You should also remember to allow a further 3 minutes for a soaking pit to empty before attempting to reload it. If you attempt to test load more than two ingots into any pit, the computer will tell you "no space in pit".

Besides the 'ENTER' button, there are three other push buttons colour coded green. To clear all the test loads from any soaking pit, just press the test load button for that pit followed by the green 'CLEAR' button. If you want to clear all the test loads on the screen, press 'CLEAR ALL'. And if you make a mistake typing in an arrival time press 'CANCEL' and then key in the correct time. If at any time you lose the display, e.g. if the computer displays an error message, pressing any of the test load buttons will return the display to you.

When an ingot has actually arrived, corresponding to one of the crosses at current time, a bell will ring and the message 'CAST ARRIVED' is flashed onto the screen. You should then press one of the red buttons, e.g. load $\boxed{A}$ depending on where you had planned to load this particular ingot. A double bar appears on the screen to indicate that pit has been loaded. You are not obliged to load the same pit you had test loaded. The computer will tell you if the pit you have tried to load is already occupied, or is in the 3 minute unloading period. In the latter case pressing the appropriate load button a second time will over-ride the 3 minute rule and will load the ingot into that pit.

Finally, some practical points:
There is often a time gap between the arrival of consecutive ingots. You would be wise to use this time profitably to try out various schedules until you find one which satisfies the target of one heated cast leaving every 5 minutes or so. (It can be done!)

As a starting point you may find that it helps to begin with the soaking pits with the longer soaking times, so that the pits with shorter times may be 'nested' within them.

There will be three practice runs to let you get the feel of the system, followed by three experimental trials. There will be eight ingot arrivals for each practice run and nine during each experimental trial.

Please do not be afraid to ask if there is anything you do not understand.

Scheduling Error Scores (relative units)

		UNCERTAINTY		
		LOW	MEDIUM	HIGH
FULL SCREEN (Prediction Span = 35 minutes)	S1	0.25	2.25	0.25
	S2	0.12	1.25	1.25
	S3	0.25	0.37	1.75
	S4	0.87	1.25	0.37
	S5	0.25	0.37	0.37
	S6	0.25	0.37	1.12
	S7	0.75	2.37	0.12
	S8	0.0	0.5	0.5
	S9	1.75	2.0	1.62
	S10	0.75	2.25	0.25
	S11	0.12	0.5	0.12
	S12	0.75	0.62	0.87
HALF SCREEN (Prediction Span = 20 minutes)	S13	0.87	2.75	1.5
	S14	0.75	1.75	0.12
	S15	0.75	0.0	0.12
	S16	0.71	0.14	1.14
	S17	0.12	0.25	1.5
	S18	0.0	0.62	0.57
	S19	0.5	0.5	0.62
	S20	0.75	0.37	0.12
	S21	0.75	0.5	0.12
	S22	0.5	0.0	1.62
	S23	0.12	2.37	0.12
	S24	1.25	2.0	0.12
QUARTER SCREEN (Prediction span = 10 minutes)	S25	1.25	1.37	1.62
	S26	0.25	1.12	0.5
	S27	0.75	0.37	0.37
	S28	0.75	2.0	1.75
	S29	2.62	0.87	0.12
	S30	0.75	2.0	0.25
	S31	0.25	0.75	0.12
	S32	0.87	1.75	0.12
	S33	0.86	1.0	0.0
	S34	0.87	1.75	0.25
	S35	0.62	0.0	1.5
	S36	0.12	2.25	1.5

Predictive Activity Data (test loads)

		UNCERTAINTY		
		LOW	MEDIUM	HIGH
FULL SCREEN (Prediction span = 35 minutes)	S1	8.25	10.25	2.25
	S2	10.75	18.75	6.75
	S3	16.0	23.5	9.75
	S4	3.5	10.25	6.75
	S5	8.25	23.75	13.25
	S6	6.0	9.5	7.5
	S7	9.5	16.25	7.5
	S8	6.25	7.75	7.5
	S9	6.75	9.0	9.25
	S10	2.75	8.0	4.5
	S11	8.25	9.75	7.75
	S12	9.75	11.5	9.0
HALF SCREEN (Prediction span = 20 minutes)	S13	8.0	11.75	7.5
	S14	7.75	9.75	6.0
	S15	5.25	5.5	4.5
	S16	4.0	4.5	4.5
	S17	6.0	5.25	5.75
	S18	5.25	7.0	6.25
	S19	6.75	10.5	7.0
	S20	5.0	6.5	3.25
	S21	14.75	13.5	7.5
	S22	7.0	9.75	10.5
	S23	5.25	7.25	6.75
	S24	11.25	14.0	11.75
QUARTER SCREEN (Prediction span = 10 minutes)	S25	4.0	7.75	4.0
	S26	4.25	3.0	0.0
	S27	2.5	2.25	3.25
	S28	2.5	3.0	3.75
	S29	4.25	7.0	5.0
	S30	0.0	0.0	0.0
	S31	3.75	2.0	1.5
	S32	0.0	2.75	0.0
	S33	0.25	0.0	0.0
	S34	0.75	2.75	1.5
	S35	5.5	6.5	5.0
	S36	0.0	0.0	1.0

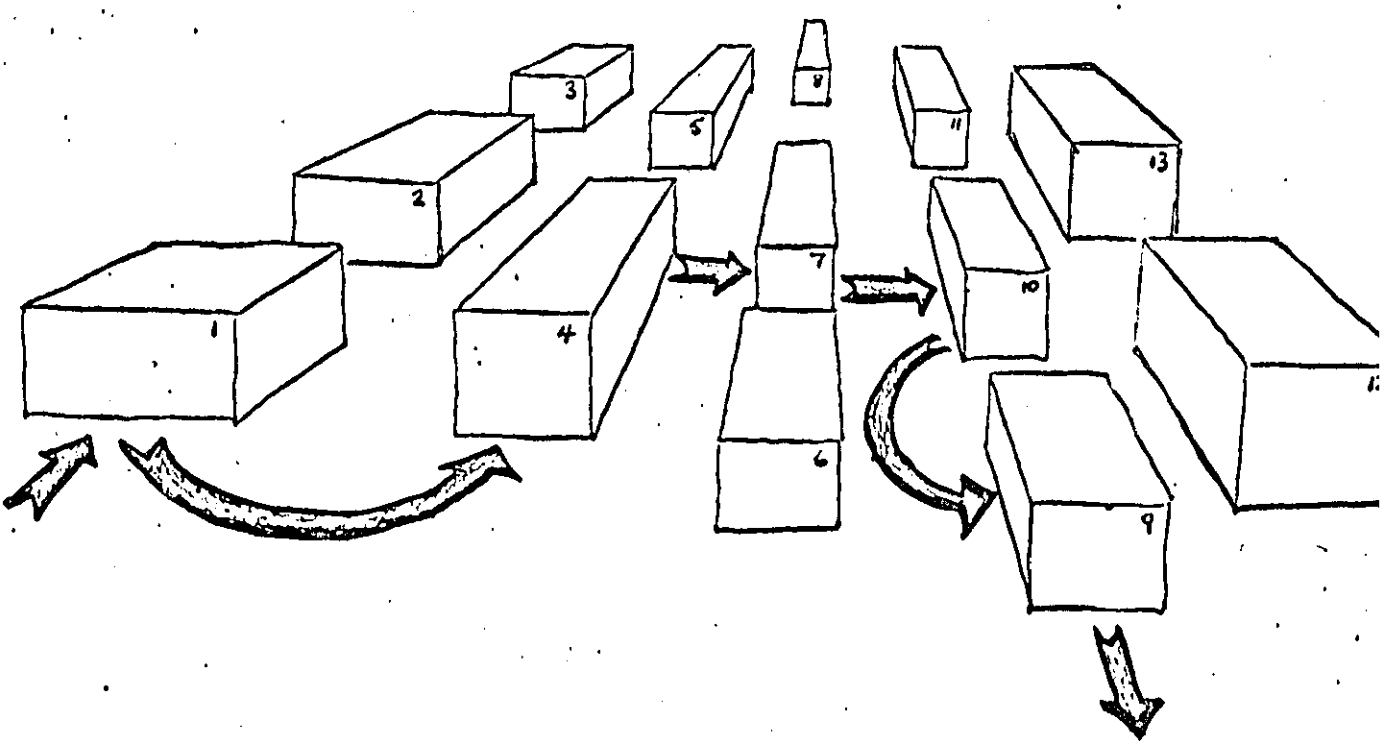
Decision Horizon Results (number of ingots scheduled ahead)

		UNCERTAINTY LOW	MEDIUM	HIGH
FULL SCREEN (Prediction span = 35 minutes	S1	3.55	3.42	1.65
	S2	3.72	2.8	1.45
	S3	2.95	2.5	0.8
	S4	4.87	3.52	4.35
	S5	4.27	2.82	2.12
	S6	3.37	5.4	4.75
	S7	3.82	2.47	4.25
	S8	5.45	4.05	1.67
	S9	2.62	4.8	2.9
	S10	3.87	2.9	2.0
	S11	4.62	4.62	3.6
	S12	4.85	3.97	3.45
HALF SCREEN (Prediction span = 20 minutes)	S13	3.75	2.32	1.85
	S14	3.72	2.42	2.57
	S15	3.55	5.77	3.7
	S16	3.65	3.07	3.25
	S17	3.72	4.87	2.57
	S18	2.7	1.97	1.67
	S19	3.75	3.57	2.25
	S20	2.85	4.4	2.3
	S21	4.62	3.47	3.12
	S22	4.87	5.6	2.6
	S23	3.42	3.52	3.3
	S24	2.35	1.8	1.77
QUARTER SCREEN (Prediction span = 10 minutes)	S25	2.02	2.22	0.7
	S26	4.6	4.2	0.0
	S27	3.42	3.6	3.07
	S28	3.72	3.4	2.67
	S29	3.87	3.32	3.0
	S30	0.0	0.0	0.0
	S31	2.9	2.17	1.62
	S32	0.0	3.27	0.0
	S33	0.0	0.0	0.0
	S34	0.27	1.45	1.62
	S35	3.25	1.62	1.35
	S36	0.0	0.0	2.43

Instructions to Subjects

MACHINE SHOP SCHEDULING

You are the production scheduler for a small machine shop. There are 13 machines in the shop and they are used to work on many different kinds of job. When a job arrives at your shop it comes with a specification of the machines needed to work on it, the time it will take on each machine and the order in which the machine operations must be done. For example, job number two might require 2 hours on machine 1, then 3 hours on machine 4, then 1 hour on machine 7, then 1 hour on machine 10, and finally, 3 hours on machine 9.



There are many jobs waiting to be processed, not all as complicated as this one. The problem is that several jobs must go through the same machine and machines can only process one job at a time. Your task is to decide which jobs are going to be on which machines at a particular time. This plan or 'schedule' must take into account the following rules:

(a) There are some jobs already on some of the machines and these must be finished before any of the new jobs can be put on.

(b) If an operation of a job is started on a particular machine it cannot be taken off until the operation is finished.

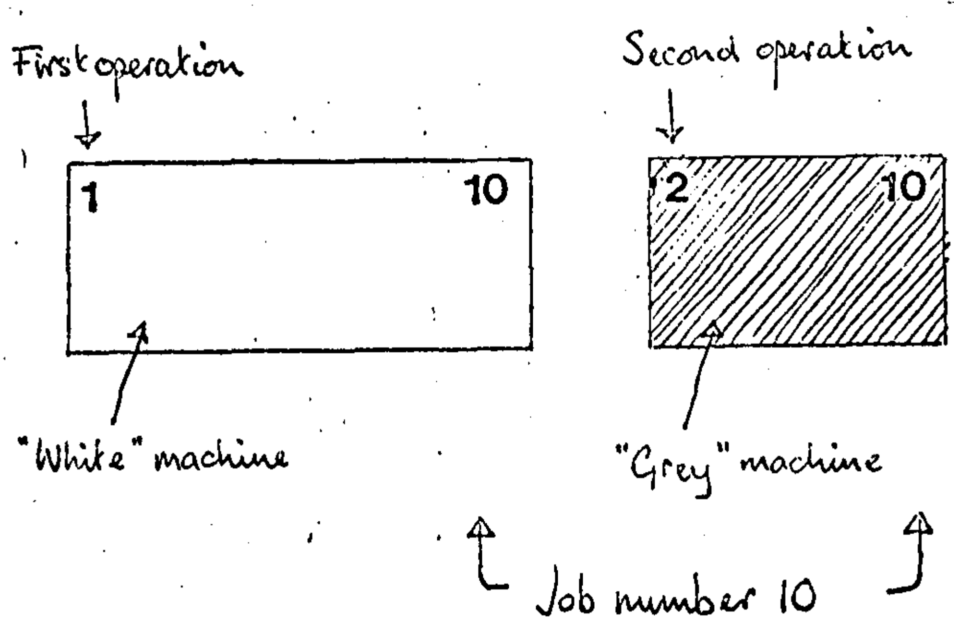
(c) A job can't be on two machines at once.

(d) A job must go through the machines in the specified order and each operation of the job must be finished before the next can begin.

(e) A machine can process only one job at a time.

COLOUR CODE

To help you produce your schedule and try out various arrangements we have represented the jobs to be processed with coloured blocks. There is one block for each operation in the job. The length of the block is the time required to carry out the operation. The colour of the block represents the machine the operation must be done on. The number on the right hand corner of the block is the job number. The number in the left hand corner shows the order of the operations - 1, 2, 3 etc. When no number is given in the left hand corner there is only the one operation in that job.



The jobs available for processing are laid out on the table. There are 55 of them. The blocks will fit into the channels on the scheduling board.

The black vertical lines divide the board into time units.

You make your schedule by:

(1) Selecting a number of jobs and fitting them onto the board.

(2) Arranging the jobs on the board so that there are no conflicts - two jobs cannot be on the same machine at the same time.

(3) Adding, replacing, removing and rearranging jobs to get the best schedule.

Remember that you are aiming to:

* Process as many jobs as possible within the time period shown by the board. (You may not be able to do them all.)

* Keep all machines as fully occupied as possible.

Machine utilisation, jobs unfinished and scheduling time scores

		Machine Utilisation	Jobs Unfinished	Scheduling Time
FULL BOARD (Prediction Span = 40 hours)	S1	79.6%	30.9%	50 mins.
	S2	78.5	34.5	30
	S3	77.9	34.5	30
	S4	77.7	38.2	25
	S5	77.1	41.8	20
HALF BOARD (Prediction Span = 20 hours)	S6	72.1	25.5	59
	S7	74.9	34.5	31
	S8	76.1	32.7	25
	S9	77.7	34.5	47
	S10	76.3	36.4	43
QUARTER BOARD (Prediction Span = 10 hours)	S11	73.9	25.5	35
	S12	73.1	32.7	45
	S13	74.5	30.9	25
	S14	70.6	30.9	43
	S15	74.5	52.7	41

List of jobs available to be scheduled

Job Number	First Operation		Second Operation		Third Operation		Fourth Operation	
1	No Colour	4	Red	6				
2	Turquoise	19	No Colour	3	Red	5	Violet	3
3	Yellow	21	Light Blue	7				
4	Blue	8	No Colour	1				
5	No Colour	2	Turquoise	5	No Colour	1	Red	2
6	Blue	5	No Colour	1				
7	No Colour	1	Violet	6	No Colour	1	Red	1
8	Yellow	2	Violet	2				
9	Yellow	5						
10	Yellow	5						
11	Purple	3						
12	No Colour	3	Blue	6	Red	5		
13	Yellow	4						
14	No Colour	1	Red	2				
15	Turquoise	3	Blue	11	Black	2		
16	Black	4	Green	5	Red	4		
17	Blue	9	No Colour	2	Red	3		
18	Red	2						
19	Blue	10						
20	Green	7						
21	Blue	7						
22	Purple	5						
23	Violet	4						
24	No Colour	1	Red	2				
26	Red	3						
27	Purple	2						
28	Red	3						
29	Light Blue	3						
30	Green	5	Black	4	No Colour	1		
31	No Colour	6	Red	11				
32	Green	4						
33	Red	1						
34	Blue	22	No Colour	1	Red	2		
35	No Colour	3	Turquoise	8	No Colour	2	Red	3
36	Light Blue	8						
37	Green	2						
38	Light Green	2						
39	Light Green	4						
40	Purple	5						

Job Number	First Operation		Second Operation		Third Operation		Fourth Operation
41	No Colour	2	Purple	3			
42	Light Blue	15					
43	Blue	16	Red	4	No Colour	2	
44	Light Blue	3	No Colour	1	Red	1	
45	No Colour	2	Red	5			
46	Green	3	No Colour	2			
47	Yellow	2					
48	Light Blue	4	Yellow	9			
50	Green	2					
51	No Colour	2	Red	4			
52	Yellow	2					
53	Blue	2	No Colour	1	Red	1	
*54	Light Blue	22	Yellow	17			
55	Light Blue	9					
57	Blue	34					
58	Blue	27					

Notes

Each colour represents a different machine operation.

The number after each colour is the processing time in hours.

There were 55 jobs in all which were available to be scheduled (there were no jobs 25, 49 or 56).

* Indicates a job critical to good scheduling performance.

Instructions to Subjects

Thank you for agreeing to take part in this experiment. We are interested in the effectiveness of different types of display for a new distillery manufacturing a certain alcoholic product.

The distillery looks like this (Figure 16).

You can control INPUT FLOW of liquid into the kettle, OUTPUT FLOW of liquid from the kettle, and HEATING/COOLING of the kettle by means of slider controls on the Kettle Control Unit.

Temperature, volume, amount of product distilled and its strength are shown by meters on the computer screen (Figure 17).

To the right of the temperature meter you will notice a space. In certain conditions, a dotted line will appear in this space extending from the temperature pointer to its right. This is a computer prediction of how the temperature will vary during the next minute - we are interested in whether such 'predictive displays' are of any use, or not, in the control of this process.

Is there anything you would like to ask?

Detailed instructions are given on a separate sheet. Before starting the experiment we will have some practice runs so you can 'get the feel' of the system. Please do not be afraid to ask if there is anything you do not understand.

Detailed Instructions

The kettle starts off empty, with all controls set to zero.

1. Fill the kettle as quickly as possible, using the INPUT FLOW control.
   Heat is generated by this process, so you should also take care to keep the temperature reading just below 90°, or an alarm will sound.

2. When the kettle is full, stabilize the kettle volume at just below 100. Set the OUTPUT FLOW control to maximum and adjust the INPUT FLOW slider to obtain this.

3. The reaction will now start. Product will register on the product meter when strength has risen above 50% proof.

4. Now try and get the temperature reading to 50° - use the INPUT FLOW control to make large adjustments in temperature, the HEATING/COOLING control for fine adjustments. Maximum production is achieved when the temperature is reading 50° and the kettle is full without overflowing.

5. You have a total of 5 minutes to distil as much product as you can. An indication of total product is shown in the top left of the screen, time in seconds in the top right.

Total amount of product manufactured in the various experimental conditio

			NP	Tay	PPM
TRAIN WITH PREDICTOR	Uncertainty	S1	1321.5	1065	1346
		S2	1088	1105	1135
		S3	1046	1011.5	1163
		S4	1067	1010	1079.5
	No Uncertainty	S5	1141	995	1043
		S6	1028.5	935	1033.5
		S7	1066.5	1019	1146.5
		S8	1179.5	1032	1076
TRAIN WITHOUT PREDICTOR	Uncertainty	S9	1087	947	1050
		S10	945.5	921.5	993.5
		S11	1130.5	1062	1136
		S12	1056.5	1017.5	1053.5
	No Uncertainty	S13	1020	928.5	992
		S14	1074	999	1012.5
		S15	1127	1079.5	1072
		S16	994	1054.5	1024

Key:
- NP No Predictor
- Tay Taylor series extrapolation model
- PPM Perfect predictor model

Instructions to Subjects

Thank you for agreeing to be a subject in this experiment.

I'm interested in how well people can monitor and control a simple system under varying degrees of disturbance and computer-aiding.

On the computer screen you will see two pointers moving against vertical scales. You can view the left-hand pointer by pressing the left-hand button on the control panel, and the right-hand pointer by pressing the right-hand button.

All you have to do is to keep both pointers as near to the 50 mark, on the scale, as you can, and within the limits shown by the dotted lines either side of the 50 mark. You can do this by moving the two slider controls on the control unit, the left-hand slider to control the left-hand pointer and the right-hand slider to control the right-hand pointer. If the pointer moves up, for example, moving the control down will compensate for this.

To assist you, in some of the conditions there will be a dotted line extendi to the right of both pointers - this is a computer prediction of how the pointers will move over the next 5, 15 or 30 seconds (depending on the condition).
Is the task clear to you?

There will be six trials to enable you to become familiar with the system, and twenty-one experimental trials. Each trial lasts 5 minutes, and for tl experimental trials there will be a break of approximately 5 minutes betweer trials for you to recover, during which time I'd like you to fill in a questionnaire on your last trial.

Note: Please return the controls to zero before starting a run.

Taylor series data (time outside limits scores) for different plant gains, levels of uncertainty and prediction spans

			PREDICTION SPAN			
			NO PREDICTOR	5 SECS.	15 SECS.	30 SECS.
LOW GAIN	S1	L	22	1	1	1
		M	167	1	1	1
		H	41	1	1	1
	S2	L	142	22	35	1
		M	133	155	1	1
		H	72	134	1	1
	S3	L	18	1	10	1
		M	1	1	1	1
		H	1	30	1	1
	S4	L	155	163	64	1
		M	103	80	31	1
		H	142	1	8	1
	S5	L	72	1	1	1
		M	17	1	1	2
		H	92	43	1	1
MEDIUM GAIN	S6	L	148	95	7	21
		M	140	50	81	20
		H	107	102	146	106
	S7	L	215	94	36	90
		M	238	49	144	214
		H	247	85	27	129
	S8	L	186	195	187	77
		M	267	252	73	55
		H	254	231	187	112
	S9	L	170	71	1	1
		M	226	31	33	17
		H	159	25	1	19
	S10	L	80	204	182	168
		M	141	155	156	138
		H	187	176	200	160
HIGH GAIN	S11	L	253	199	188	173
		M	257	225	223	205
		H	283	235	200	235
	S12	L	244	231	222	239
		M	282	235	235	202
		H	266	266	262	244
	S13	L	84	101	144	157
		M	205	214	155	171
		H	200	227	246	195
	S14	L	263	184	214	235
		M	273	211	153	168
		H	288	204	196	232
	S15	L	267	250	175	237
		M	279	277	248	263
		H	256	287	274	290

Key: L = Low Uncertainty (Input disturbance)
M = Medium Uncertainty
H = High Uncertainty

PPM data (time outside limits scores) for different plant gains, levels of uncertainty and prediction spans

			PREDICTION SPAN			
			NO PREDICTOR	5 SECS.	15 SECS.	30 SECS.
LOW GAIN	S1	L	22	14	1	1
		M	167	1	1	1
		H	41	39	1	1
	S2	L	142	1	1	1
		M	133	30	1	1
		H	72	84	1	1
	S3	L	18	1	39	1
		M	1	1	1	1
		H	1	45	1	1
	S4	L	155	107	73	1
		M	103	1	1	1
		H	142	31	1	1
	S5	L	72	98	1	1
		M	17	122	1	1
		H	92	1	1	1
MEDIUM GAIN	S6	L	148	96	1	36
		M	140	11	69	123
		H	107	167	21	229
	S7	L	215	24	1	1
		M	238	1	15	4
		H	247	269	28	40
	S8	L	186	226	1	1
		M	267	240	23	1
		H	254	262	9	40
	S9	L	170	42	1	1
		M	226	12	1	75
		H	159	1	1	22
	S10	L	80	71	1	1
		M	141	184	1	1
		H	187	230	43	50
HIGH GAIN	S11	L	253	108	1	1
		M	257	244	167	144
		H	283	204	179	266
	S12	L	244	186	1	22
		M	282	253	163	258
		H	266	110	256	277
	S13	L	84	122	1	1
		M	205	184	191	198
		H	200	188	145	250
	S14	L	263	200	1	1
		M	273	169	64	150
		H	288	171	122	222
	S15	L	267	234	171	1
		M	279	269	248	237
		H	256	251	274	281

Key: L = Low Uncertainty (Input disturbance)
M = Medium Uncertainty
H = High Uncertainty

Questionnaire presented after each trial

Initials _____ Trial Number _____

These questions relate to the trial you have just run.

How easy/difficult did you find the last trial to control?

EASY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | DIFFICULT (Please tick one box)

How smooth/uneven were your control actions?

SMOOTH | 1 | 2 | 3 | 4 | 5 | 6 | 7 | UNEVEN

If you found the task difficult to control ....

Did the left or right-hand pointer present the most difficulties? [left] [right]

Was the main problem: Speed of response? [ ]
Disturbances? [ ]
Switching between displays? [ ]
Pointer drift? [ ]
Other? (specify) [ ] __________

Was a predictor trace ( ) provided? [Yes] [No]

If yes, how useful did you find it?

VERY USEFUL | 1 | 2 | 3 | 4 | 5 | 6 | 7 | NOT AT ALL USEFUL

Did you find the time covered by the predictor trace long enough? [Yes] [N

Did you use any particular section of the trace?

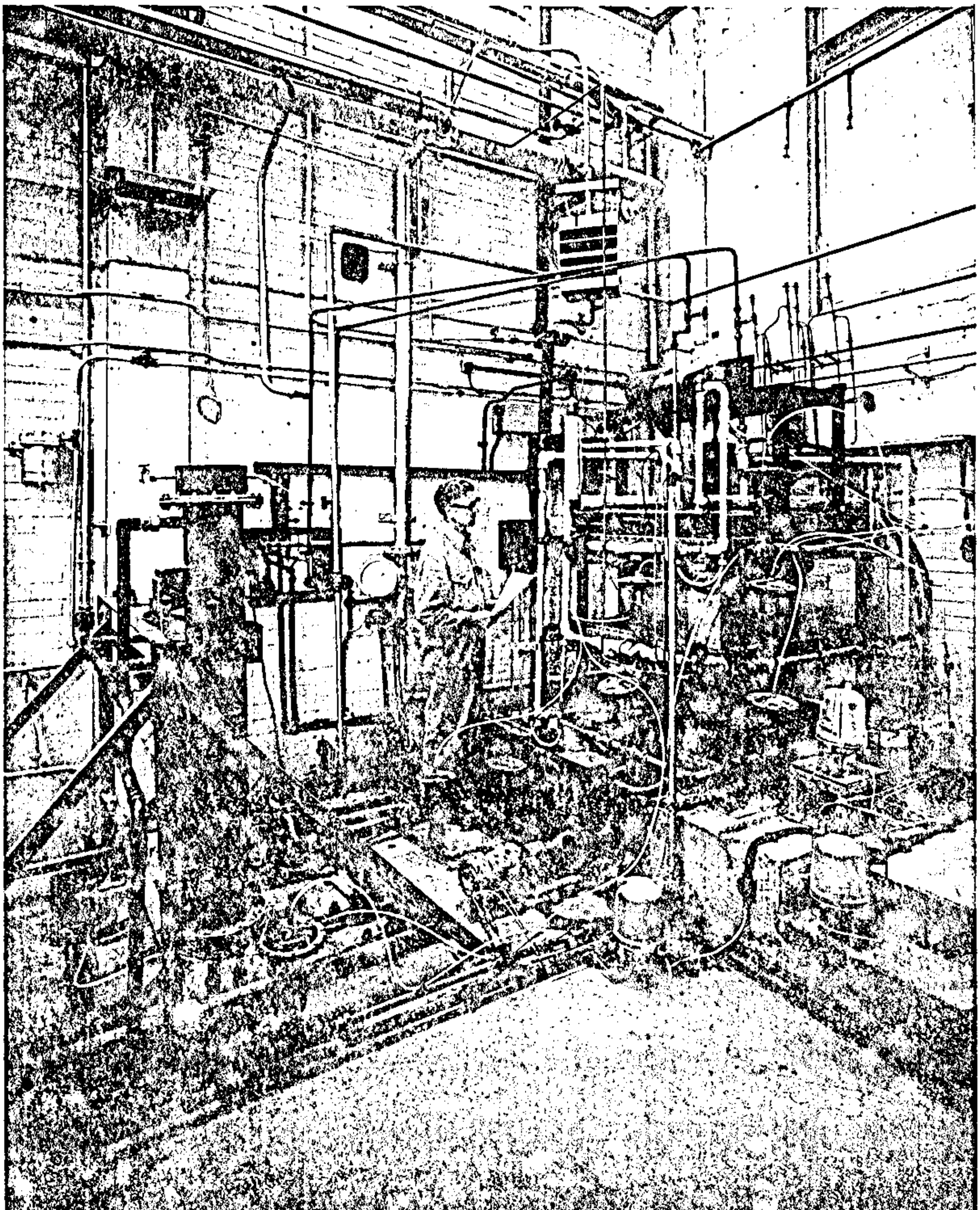
First segment? [ ]
Middle segment? [ ]
End segment? [ ]

What strategy did you use to control?

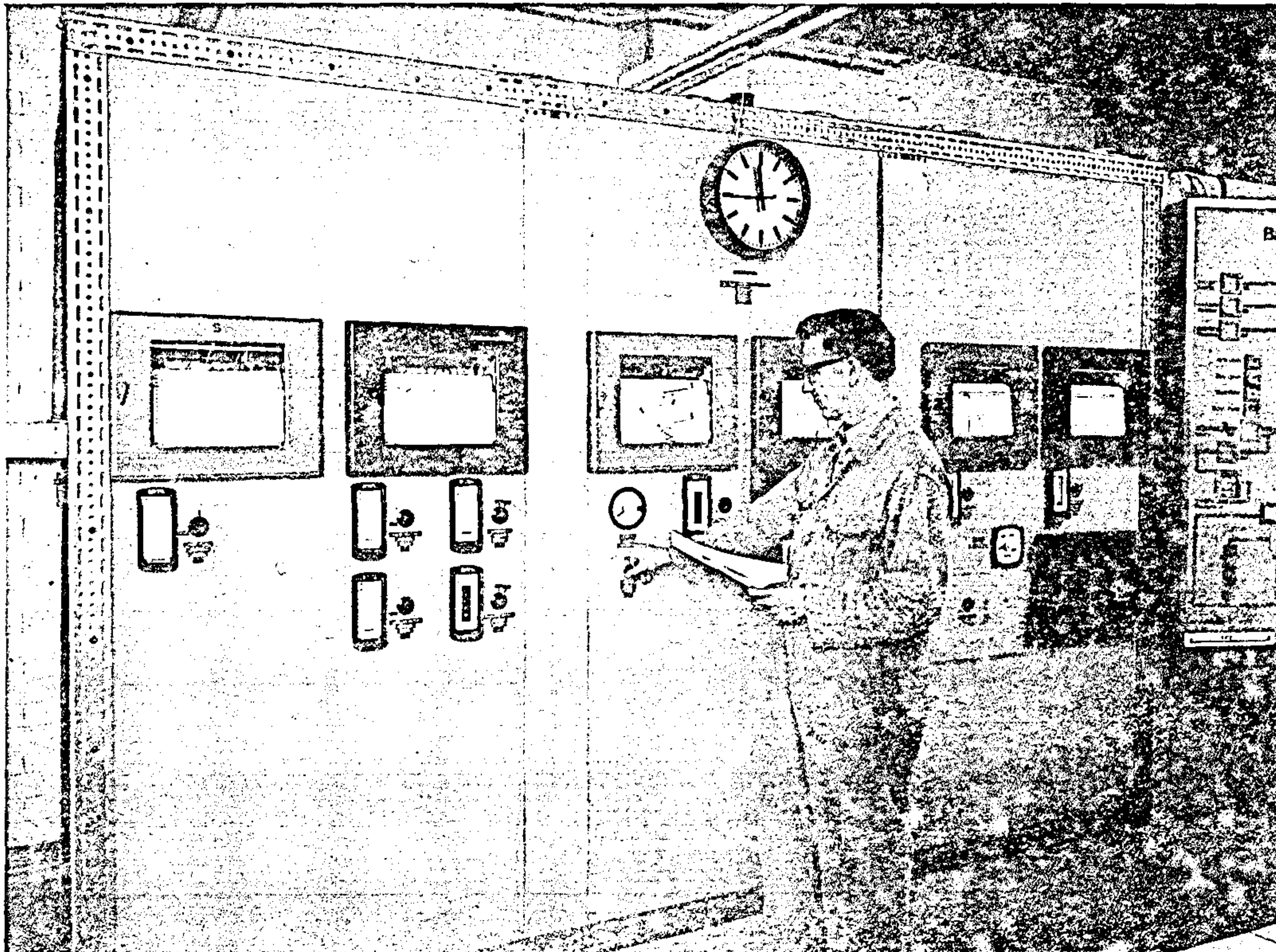
Did you anticipate the pointer's future movements? [Yes] [No]

If yes, how far ahead were you anticipating? _____ seconds

Any other comments you feel may be relevant?



The Warren Spring Batch Kettle Plant (photograph by courtesy of Warren Spring Laboratory).



The Batch Kettle Control Panel (photograph by courtesy of Warren Spring Laboratory).

Listing of Taylor series subroutine implemented on Argus 500 computer

```
.
.
. CALL PRED (ITEMP,IFUT)
  Machine code instructions follow to transfer predicted values
  to chart recorder
.
.
.
   SUBROUTINE PRED (IX,JJ)
   DIMENSION JJ (3), FUTURE (3)
   COMMON/DATA/AX (192), AY (44), DNEW (10), DOLD (10)

   NP = 3
   ND = 2
   DNEW (1) = FLOAT (IX)
   DO 10 I = 2, ND
10 DNEW (I) = DNEW (I-1) - DOLD (I-1)
   DO20 I = 1,ND
20 DOLD (I) = DNEW (I)
   DO30 J = 1,NP
   COEFF = 1.0
   FUTURE (J) = DNEW (1)
   DO30 I = 2,ND
   COEFF = COEFF * FLOAT (J)/FLOAT (I-1)
30 FUTURE (J) = FUTURE (J) + DNEW (I) * COEFF

   DO31 I = 1,NP
31 JJ (I) = FUTURE (I)
   RETURN
   END
```

Comments

PRED (IX,JJ) forms a subroutine within the control program HFPSM2. IX inputs the data value, JJ returns the predicted values.

NP is the number of pens.

ND is the number of derivative terms.

DNEW (I) and DOLD (I) are used to calculate derivative terms.

COEFF is the coefficients of the Taylor series expansion.

FUTURE (1,2,3) are the predicted values at times 10, 20, and 30 secs. ahead.

Operator performance scores with prediction on temperature

	NO PREDICTOR			M.P.R.		
Subject	(a)	(b)	(c)	(a)	(b)	(c)
OPERATOR 1	97.4%	2.744	£ 84.0	99.2%	2.502	£167.8
	96.6	3.792	75.5	99.3	2.626	141.2
	97.7	2.483	81.9	98.5	1.443	117.3
$\bar{x}$	97.2	3.006	80.5	99.0	2.19	142.1
OPERATOR 2	97.7	1.884	136.8	98.7	0.61	160.5
	98.7	1.051	149.4	97.1	2.182	133.3
	97.8	1.414	148.8	99.4	0.808	209.2
	97.9	2.875	- 8.2	98.8	1.339	184.6
$\bar{x}$	98.0	1.806	106.7	98.5	1.235	171.9
OPERATOR 3 †	95.7	8.518	48.6	97.8	4.187	98.7
	95.7	3.298	81.0	98.6	5.177	122.7
				98.6	6.705	131.5
$\bar{x}$	95.7	5.91	64.8	98.3	5.36	117.6
OPERATOR 4 †	98.8	5.572	92.5	97.8	3.051	106.2
	97.1	2.885	95.9	99.2	4.688	137.8
	92.6	3.439	16.2			
$\bar{x}$	96.17	3.96	68.2	98.5	3.87	122.0
GRAND AVERAGE	96.8%	3.67	£ 80.1	98.6%	3.16	£138.4

Key:

(a) Percentage conversion of reagents to product

(b) Standard deviation around the predicted variable

(c) Calculated profit in £

† indicates prediction trials were undertaken first

Operator performance scores with prediction on pH

	NO PREDICTOR			M.P.R.		
Subject	(a)	(b)	(c)	(a)	(b)	(c)
OPERATOR A	96.0%	0.745	£ 91.4	99.2%	0.362	£204.4
	97.1	0.627	134.8	98.2	0.442	168.1
	96.9	0.653	99.7	98.4	0.421	177.9
	96.8	0.793	138.5	97.5	0.442	146.0
	97.7	0.6	155.5	94.9	0.679	69.3
				97.5	0.32	150.4
				97.2	0.233	143.0
$\bar{x}$	96.9	0.686	124.0	97.56	0.414	151.3
OPERATOR B	96.9	0.67	77.5	93.1	0.975	32.7
	90.8	0.649	- 2.6	92.2	0.643	27.7
	92.1	1.152	9.3	94.6	0.755	54.6
	91.5	0.97	13.8	92.7	0.691	14.6
$\bar{x}$	92.83	0.86	24.5	93.15	0.77	32.4
OPERATOR C †	92.3	0.95	7.5	97.7	0.593	122.1
	98.3	0.42	116.9	96.2	0.555	76.7
				94.2	1.017	27.9
				96.5	0.748	65.1
				94.2	0.878	18.9
$\bar{x}$	95.3	0.69	62.2	95.8	0.76	62.1
OPERATOR D †	93.7	0.921	46.8	94.8	0.956	47.6
	96.5	0.706	95.6	97.5	0.739	90.9
$\bar{x}$	95.1	0.814	71.2	96.25	0.849	69.25
GRAND AVERAGE	95.0%	0.76	£ 70.5	95.7%	0.70	£ 78.8

Key:

(a) Percentage conversion of reagents to product

(b) Standard deviation around the predicted variable

(c) Calculated profit in £

† indicates prediction trials were undertaken first

Details of profit figure calculation

Item	Cost
Process time	£10/hour
Caustic	£ 2/gallon
Reactant	£ 1/gallon
Feed	£ 1/gallon
Steam	£ $\dfrac{\int \text{valve position.dt}}{60}$
Cooling water	£ $\dfrac{\int \text{valve position.dt}}{240}$

Product value

(1) If conversion $>$ 90% then

product value = £ $(30 + (\%\text{Conversion} - 90)^2 \times 0.3)$/gall.

(2) If 70% $<$ conversion $<$ 90% then product has to be reprocessed

product value = £ $\dfrac{(\text{Original } \% \text{ conversion} \times 37.5)}{100}$/gall

reprocessing cost = £50

(3) If conversion $<$ 70%

product value = £0.