UNIVERSITY OF CAMBRIDGE

Faculty - Biology B

Department of Experimental Psychology

THE HUMAN ADAPTIVE CONTROLLER

A thesis presented for the Degree of
Doctor of Philosophy by

Brian Ronald Gaines

Trinity College

July 1971

THE HUMAN ADAPTIVE CONTROLLER

CONTENTS

0.1 Summary

The scope of this study is human learning behaviour, particularly in perceptual-motor skills, and the variables which influence it, including the nature of the environment in which learning takes place and the effect of verbal instructions. The study ranges from a general theory of adaptive behaviour based on the algebraic theory of semigroups, to specific experiments on the optimum control of learning behaviour in a perceptual-motor skill, and includes comparative studies of human and machine learning.

The first objective of the study has been to develop a rigorous and systematic account of the relations between behaviour, structure and purpose in arbitrary systems including men and machines. The second objective has been to use this account to develop an integrated approach to the problem of training, in which a knowledge of the patterns of behaviour, the structure and the desired goals of a system, may be used to formulate an optimal training strategy. The final objective has been to demonstrate the application of the theory to a realistic situation, and compare some of the theoretical predictions with experimental results.

In the theoretical studies a taxonomy of adaptive behaviour is established which enables operational and purely behavioural definitions to be provided of terms such as 'adaptive' and 'adapted'. The taxonomy is given a mathematical formulation through the algebraic theory of semigroups by deriving an algorithm for constructing a minimal and observable automaton cybernetically equivalent to a system known only through its observed behaviour. Further information about the structure of such automata for adaptive systems is obtained by analysing the influence of purpose on behaviour, in terms of the epistemological problems induced by the dual-control situation of learning about a system whilst trying to control it. These developments lead to the study of training as a control problem, and adaption as the stability of a hierarchical system.

In the experimental studies a high-order compensatory tracking task is taken as the environment and a feedback training system developed on the basis of the theoretical arguments. The viability of this system in terms of its dynamics and stability is evaluated theoretically and experimentally, using both human operators and automatic controllers. The utility of the system is investigated by an experiment with 72 RAF pilots, in which various modes of training are compared, and interactions with the form of instructions given are also evaluated. These experiments are repeated with artificial adaptive controllers as subjects, in order to enable a comparative study to be made of human and machine learning.

0.2 Acknowledgements

The research on the human operator described in this thesis was carried out at the Department of Experimental Psychology, University of Cambridge, and I am grateful to Professor O.L. Zangwill, Head of the Department, for providing me with accomodation. I am also grateful to Mr.R.L.Gregory of the Psychological Laboratory, and Mr.P.E.K.Donaldson of the Physiological Laboratory, University of Cambridge, for initiating and supervizing this research.

The work was supported financially by the Ministry of Defence (Army), and I am grateful to Dr.J.C.Penton for his ready co-operation and advice in all matters affecting the administration of the contract, and to Mr.H.C.W.Stockbridge who, through his own work, initiated the Ministry's interest in adaptive training.

The theoretical studies described in this thesis have been influenced from many sources, and, in particular, they owe much to discussions with Mr.A.Watson of the Psychological Laboratory, who supervized my undergraduate work in Psychology, and to Dr.J.H.Andreae of Standard Telecommunication Laboratories, Harlow, with whom I have spent many happy hours discussing problems of artificial intelligence and the nature of learning.

The theoretical studies are built upon the foundations laid in the works of Dr.G.Pask, Dr.W.R.Ashby and Dr.N.Wiener, from whom, indirectly, has come the intellectual stimulation for the present work.

The technological requirements of this study could not have been undertaken without access to the workshop facilities set up by Mr.R.L.Gregory and administered by Mr.S.H.Salter. Equally, the experimental studies could not have been undertaken without access to the controlled environment set up by Dr.J.L.Gedye for his own work.

Finally, I have a great debt of gratitude to the men of R.A.F. Oakington whose interest, punctuality and co-operation made the experimental studies far simpler, more pleasant and more worthwhile than they otherwise might have been.

No part of the work described in this dissertation has been carried out in collaboration.

B Gaines

0.3 Section Contents

0.4 List of Figures

CHAPTER 1 : INTRODUCTION

1.1. Motivation and Aims of Investigation

The work described originated as an investigation of 'adaptive training techniques' for human operators attempting to learn a complex perceptual-motor skill such as driving a car or flying an aircraft. The major objective has been to investigate in depth a situation in which the 'difficulty' of a task is automatically adjusted to maintain a constant level of performance for an operator learning the task. Since the automatic adjustment system acts as an 'adaptive trainer', increasing the difficulty of the task as the operator learns at a rate dependent on his learning, it is plausible that it may provide a 'teaching-machine' for perceptual-motor skills and speed learning.

Because the automatic training system is coupled through feedback to the performance of the trainee, the possibility of overall instability arises and requires both theoretical and experimental investigation. Given that an 'adaptive trainer' can be made to operate stably under reasonable conditions, its effectiveness as a teaching system and the variables that determine that effectiveness are also open to investigation.

In any experimental study involving human learning the degree of generality of the results obtained must come into question. In particular, in a study of training where 'feedback' is involved it is reasonable to expect that the 'sensitivity' of the results obtained will be reduced, not only to variations in the trainees but also to variations in the type of task used in training and to the exact nature of the training strategy itself. In the present study the sensitivity of the results to the trainees has been investigated by including automatic 'learning machines' as experimental subjects.

The sensitivity to task and training strategy raises deeper issues concerning the system-theoretical nature of the results obtained, for example, whether general results on the stability and efficacy of adaptive training can be derived for abstract systems which include the particular system investigated experimentally as a special case. In the present study this problem is investigated at a fundamental level through the formal definition of an 'adaptive' or 'learning' system, and both theoretical and experimental results are presented for systems of varying degrees of generality.

Thus the study has ranged from system-theoretical investigations of adaption and learning, through theoretical analyses of the problem of training, to actual laboratory and simulation experiments on the training of human beings and learning machines. Such a range seemed essential at the time of the study since there was neither adequate theoretical material on the nature of learning available on which to base experimental studies of training, nor a sufficient range of results on adaptive training to act as proving ground for a purely theoretical investigation. In this thesis an attempt has been made to present both theoretical and experimental studies in a unified form, and to link them together wherever possible.

1.2 Background to the Objectives

One of the most remarkable features of human behaviour is its wide range of possible variation in response to the different characteristics of the environment in which it takes place. Man, out of all the animals, has developed in the course of evolution the greatest capacity for changing his mode of behaviour to that which best achieves his goals in any new environment. Some form of adaption to circumstances is found in even the lowliest micro-organisms, however, and this capability has sometimes been taken to characterize life itself.

The characterization of life by its adaptive capability has been made less tenable in recent years by the success of control engineers in designing automatic controllers with a similar ability to modify their control policies in the event of unpredictable changes in the controlled plant. This development makes it reasonable to consider the possibility of a unified approach to the study of adaption and learning, in both animals and machines.

A unified approach to some aspects of psychology and control engineering is attractive on a number of grounds. Firstly, the well-defined and known structures of automatic controllers enable the implications of theoretical constructs linking structure and behaviour to be clarified very rapidly. Secondly, such terms as 'purpose' have to be defined clearly and operationally if they are also to be applied to machines. Thirdly, automatic adaptive controllers provide a source of identical 'subjects' for experiments on factors affecting learning. Fourthly, engineers are responsible for many systems studies and associated mathematical developments which have direct applications in psychology. And finally, it is possible that the automatic controller of the future will be a general-purpose adaptive system, simple to fabricate because of its homogeneity of structure, which will be 'trained' to implement a specific control policy. Current investigations of this possibility are just as likely to produce results relevant to psychology as they are to contribute to control engineering.

1.2.1 Behaviour, Structure and Purpose

The chief problem in formalizing such concepts as 'adaption' and 'learning' is to establish and maintain a clear distinction between the structural, behavioural and teleological connotations of these terms. The structure, behaviour and purpose of any system are intimately related and everyday language makes little distinction between them. However, in psychology it is this relationship which is to be investigated, and its individual components must be clearly separated.

For example, if a system is assumed to have a purpose such that its behaviour is directed to some goal, then it is possible to observe its behaviour and determine to what extent that goal is attained. Thus, the 'adaptivity' of a system may be defined in purely behavioural terms, given a teleological assumption. Equally, however, it is possible to observe behaviour without any pre-suppositions as to its purpose, and examine it for evidence of goal-seeking. The goals then become a property of, or a way of describing the behaviour. For example, an event may be termed 'reinforcing' if it increases the tendency for behaviour preceeding it to occur, and the 'goal' of any adapting system becomes the seeking of reinforcement.

Either approach to the analysis of adaptive behaviour is valid, evaluation of the fulfilment of an assumed goal, or determination of the goal from the behaviour. Superficially, the former is more relevant to control engineering and human operator training, and the latter to animal studies. The problem does not lie in the existence of these differing approaches, however, but in the ease with which a tacit change may be made between them. An examination of pure observations of behaviour, not influenced by any assumptions about its purpose, is biassed by the tendency of normal descriptions of behaviour to be also evaluative; acts are described by their terminal effects rather than the motions which produce them. Both aspects of an act are part of an unwieldy 'total behavioural description', from which everyday language eliminates irrelevant components. In so doing, however, the language introduces the very assumptions which are the subject matter of psychological theory and experiment.

A different type of problem arises from the relationship between the structural and behavioural connotations of the term 'adaptive'. Given full information about the structure of an adaptive system it is possible to predict what its behaviour will be in various environments. Equally, given full information about the range of possible behaviour of a system it is possible to limit its structure to some sub-set of all possible structures. The relation between the physical structure and the sub-set of behaviourally determined structures is that the former must be contained in the latter. In practice, for large, complex and irreversible systems, such as the human organism, the full details of the structure, and the full range of possible behaviour, of the system are inherently unobservable, and the structure will be limited to some sub-set by direct observations of its physical nature, and the behaviour will also limit the structure to some sub-set as before. The actual structure must lie in the intersection of these sub-sets, and hence, in this sense, behaviour can given evidence as to structure not obtained by direct observation of the structure.

Further complications arise if the gaps in observed behaviour, which are essentially unfillable by observation, are in fact filled by assumptions about the behaviour. For example, an event shown to be 'reinforcing' for some aspect of behaviour may be assumed 'reinforcing' for all other aspects. Assumptions are clearly part of any process of scientific induction, but have a peculiar status in psychology because they are inherently necessitated by the irreversible, non-replicable systems studied. The justification for particular assumptions about behaviour may come from either structural or teleological considerations - the behaviour of each of an ensemble of rats may be treated as if were the range of behaviours of a single individual, either because the rats had must the same goals in satisfying hunger, or because they have much the same physiological structure.

A good example of the interplay between structural, behavioural and teleological connotations of the term 'adaptive' lies in the justification for the use of automatic adaptive controllers in experiments designed to test the efficacy of different training techniques for human operators. The adaptive controllers are known to have been built for the purpose of attaining certain goals. Thus purpose will constrain the controllers' behaviour, and any controller with the same purpose, such as a human operator, will be under the same constraints. Thus, it is possible to use an adaptive controller as a 'subject' in an experiment on training with reasonable grounds for supposing that its behaviour will be similar to that of human operators.

The considerations put forward in this section made it reasonable to suppose that a rigorous and systematic study of learning in any animal, and indeed in arbitrary systems, might be based on a formalization of the concepts of behaviour, structure and purpose in adaptive systems and of the relations between these three. The initial objective was to analyse the way in which the adaptivity of the system might be evaluated from its behaviour, since this was a pre-requisite to the analysis of other aspects of adaption. For example, the determination of goals from behaviour seems to depend on criteria such as, 'the assumed goals for which the system is most adaptive'. The next objective was to establish the relation between the adaptive behaviour of a system and its structure, and the final theoretical objective was to establish the influence of the purpose of a system on its behaviour, independently of information about its structure.

1.2.2 Application to the Problem of Training

The normal proving ground for a theoretical analysis of adaption and learning is in a study of the lower animals, such as rats, cats and octopi. However, there is an aspect of adaption which, although present to some degree in communities of lower animals, only manifests ~~in~~telf fully in human society, and that is the process of education or training whereby a positive effort is made by some outside agency to direct the course of learning. Application of theoretical results to the problem of training is particularly attractive because an integrated approach to all aspects of adaption is required.

The problem of training may be regarded as that of varying the learning environment in such a way that the trainee is taken as rapidly as possible from his initial, naive state to one where he is competent to perform the required task. Viewed in this way, training is itself a control problem, albeit at a high level of abstraction and involving systems of great complexity. The statement of, and the solution of, this control problem requires knowledge of the structure which underlies the behaviour of the trainee. This knowledge itself may come from information about the structure, behaviour or purpose of the trainee. In practice, none of these alone is sufficient, and the diverse sources have to be integrated into a coherent basis for optimal solution of the training problem.

One advantage of taking training as the proving ground for a theory of learning is that it gives a rationale for decisions which would otherwise be arbitrary. In any study of adaption there are abitrary, but necessary, methodological decisions which must be taken, but are not determined by the theory. For example, the decision to shut a rat in a Skinner box and only observe its behaviour as a succession of bar pressings, is arbitrary and yet necessitated, not only by the practical impossibility of observing the animal's behaviour in complete detail, but also by the theoretical impossibility of utilizing such detailed observations which make every observed behaviour an isolated event unrelated to other observations. Since the objectives of a study of training are not to provide a complete account of all aspects of learning, but rather to synthesize and evaluate training techniques for particular purposes, there is an independent basis for the arbitrary decisions which have to be made in a study of adaption.

These considerations, and others of a more mundane nature made it attractive to apply the theory developed to the problem of training humans to perform skilled tasks. Hence the study of behaviour, structure and purpose in adaptive systems was directed towards establishing a formal theory of training and a basis for the synthesis of optimal training programs.

1.2.3 Experimental Studies of 'Adaptive Training'

Since the theory of adaption was intended to be a unified approach to any system, it was desirable in the experimental studies to choose environments in which both human and machine learning might be investigated. The problem-solving, decision-making and linguistic skills of the human operator, whilst stimulating much research in 'artificial intelligence', are far from being emulated by machines at present, whilst the continuous control skills involved in flying, driving and tracking are closely paralleled by such devices as adaptive auto-pilots and 'model-reference' process controllers. Hence, training human operators in perceptual-motor skills was taken as a suitable situation for experimental evaluation of the theory.

One particularly interesting situation which has been investigated by several workers in recent years, and which has proved troublesome both theoretically and experimentally, is 'adaptive training'. A 'self-adjusting simulator', or 'adaptive training system' for a perceptual-motor skill is a device which automatically adjusts the difficulty of a control task according to the operator's performance in an attempt to maximize his rate of learning. Such devices have been proposed by various agencies and individuals concerned with the training of human operators in control skills, such as manoeuvring an aircraft, missile or submarine, but no rigorous studies of their efficacy, or general effects on learning, have taken place.

It is a reasonable hypothesis that for any operator with a given level of skill there is an optimum level of task difficulty which maximizes his rate of learning. When the task is too difficult he generates a large amount of error and is unable to perceive the effect of his control movements, and when the task is too easy he is able to perform it well, and has no requirement for a better control strategy. Thus one might expect two distinct effects - if the required task is easy for an operator then he will learn more rapidly with training at a higher level of difficulty - whilst if the required task is very difficult for an operator then he will learn more rapidly with an easier task. Furthermore, the relative ease or difficulty of a task is a function of the operator's basic ability and state of learning, and the optimum level of difficulty would be expected to increase as the operator's skill improves. The optimal training technique should, therefore, involve feedback from the operator's state of learning to the level of difficulty of the task.

The theoretical studies already outlined from the basis for a formal treatment of the proposed advantages of adaptive training and for the design of adaptive training systems (which are called feedback trainers in this report, in order to avoid over-use of the term 'adaptive'). Hence, an experimental study of a feedback trainer, designed according to these considerations, was undertaken. This involved comparison of feedback training with other training techniques using both human operators and automatic adaptive controllers as trainees. Some auxiliary problems were discovered and investigated in this study, including the stability of adaptive trainers and the interaction of the effect of instructions with those of different training techniques.

1.3 Structure of Thesis

It has been noted in Section 1.1 that the studies reported in this thesis cover a very wide range from the theoretical to the experimental, and an attempt has been made in its organization and presentation to show the balance of the contributions to the different areas. Theoretical developments which are essential to the logical development of the results but do not contribute to their main content have been placed in appendices. Survey material which is essential background material to the arguments developed is also placed in appendices - however, material critically surveyed in order to develop a main line of reasoning is placed in the main text. Detailed experimental results are also given as an appendix but discussed in the main text. This organization has enabled the presentation to be complete whilst allowing stress to be placed on the main part of the investigations.

1.3.1 Contents of Chapters

In the following sections the main results of each chapter are outlined -

1.3.2 Axiomatic Foundations of Learning and Training

Chapter 2 commences with a critical review of previous work on behavioural definitions of adaption both in psychology and control engineering. An explicatum of the term 'adaptive' is proposed which enables many aspects of adaptive dynamics, not previously made clear, to be formalized. This is based on the concept of a 'task' as the unit of adaptive behaviour, and consideration of the variation of the satisfactoriness of the interaction between controller and environment over sequences of tasks leads to the definition of various modes of adaption. Finally the results of Appendix 3 are used to define a minimal observable structure underlying the behaviour, an 'adaption automaton', and the definitions of adaption are framed in terms of this.

1.3.3 Training as a Control Problem

In Chapter 3 the approach to learning behaviour developed in Chapter 2 is extended to provide a rigorous foundation for the analysis of training as a control and stability problem in the state-space of the adaption automaton of the trainee. The technique for selecting sequences of tasks to bring the state of the automaton into a desired region enables various modes of training to be distinguished. Before the problem of training can be 'solved' it is necessary to have some information about the structure of the adaption automaton, and two theorems on training establish the minimal and practical levels of information required for an effective trainer to be designed. Possible sources of such information are then investigated through consideration of the epistemological problems of the trainee in attempting to solve the 'dual-control' problem of controlling an environment whilst at the same time learning about it. An automata-theoretic statement of this problem is given, in which it is shown that any control policy restricts the environment to some sub-environment, and that the sub-environment generated by a naive controller may be unsuitable for learning. The basic training strategy is then formulated as maintenance of a sub-environment similar to that encountered by a controller which has learnt the problem.

1.3.4 A Feedback Trainer for a Tracking Skill

In Chapter 4 the selection of a suitable problem for the evaluation of a feedback trainer is discussed, and the results of previous chapters are applied to the design of a feedback training system for a task involving compensatory tracking through high-order dynamics. The behaviour of the training system is analysed theoretically, particularly its stability, and experimental results are given to verify this analysis for non-adapting automatic controllers and human operators.

1.3.5 Experimental Evaluation of Feedback Training

In Chapter 5 an experiment designed to test the practical utility of feedback trainers is described, which involved training 72 RAF pilots in a novel tracking task under six different training regimes. The methodology of an experimental comparison between different training techniques is discussed, and means for overcoming effects of fatigue, and auxiliary variables such as verbalization, and so on, are described. An experimental design for evaluating not only the various modes of training proposed in Chapter 3, but also the interacting effects of different forms of verbal instruction and the relative effects of stress on performance, is proposed. The results obtained with this design are analysed and the significant effects obtained are discussed in detail.

1.3.6 Experiments with Learning Machines

In Chapter 6 a range of experiments on the adaptive behaviour, and the training, of automatic controllers based on adaptive threshold logic elements are described. The first experiments involve a very simple pattern-classification problem, and illustrate the various modes of adaption and training defined previously. Other experiments involve the use of adaptive controllers in the same situation as the human operators in the experiments of Chapter 5, and a comparison is made between the effects of different training techniques and verbal instructions on human and automatic adaptive controllers.

1.3.7 Summary, Conclusions, and Recommendations

Chapter 7 contains a summary of the theoretical and experimental results, and brings together the main conclusions. The objectives of the study, and the extent to which they have been attained, are reviewed, and, finally, recommendations are made for the directions of further research on adaptive training and the theoretical foundations for the study of learning behaviour.

1.3.8 Appendices

Appendix 1 on Adaptive and Learning Controllers contains the background and reference material relevant to the studies of Chapters 3 and 6. After introducing the concepts of open-loop and closed-loop adaption, the appendix is mainly concerned with adaptive threshold logic elements (ATLEs) and their properties both as pattern-recognizers and as controllers.

Appendix 2 on the Algebraic Theory of Semigroups contains basic definitions and results relevant to Appendix 3 and Chapters 2 and 3.

Appendix 3 on the transition from Behaviour to Structure is an original study in its own right, but has been placed as an appendix since it contains mainly mathematical results bridging the gap between the behavioural approach to adaption of Chapter 2 and the structural basis for training of Chapter 3. In Appendix 3 the problem is analysed of deriving a structure for a system from its behaviour which is minimal, in that it has only sufficient complexity to account for the observed behaviour, and which is observable in that it is possible to determine the 'state' of the structure from a sufficiently long sequence of past behaviour. A procedure for determining such a structure from a complete description of all possible behaviours is derived.

Appendix 4 on the Human Controller contains a review of experimental studies of the human operator in control systems, with particular emphasis on adaptive behaviour and techniques of training. Linear and nonlinear models of the human operator are examined to determine the basic constraints upon the possible control strategies available to him. Recent studies of the adaptive capabilities of the human operator in response to changes in the controlled system are then reviewed. Finally, experiments on training the human operator, and in particular the few experiments on adaptive training, are critically examined.

Appendix 5 contains Experimental Results in detail for the study of training described in Chapter 5.

Chapter 2 : AXIOMATIC FOUNDATIONS OF LEARNING AND TRAINING

2.1 Introduction

It is an essential feature of the training situation, that the trainer is attempting to exert some control over the learning processes of the trainee. Teaching and training by human teachers involves both the concept of changing the state of the student to one in which particular behaviour may be elicited, and also the concept of feedback from the behaviour of the student to the behaviour of the teacher. These allusions to the fundamental concepts of modern control engineering suggest that a formal approach to the problems of training might be made through modern control theory, and, in the following chapter, this possibility is explored and a control-theoretical approach to training is developed.

Before any rigorous approach to the control of human learning through training can be taken, however, it is necessary to make a formal analysis of the nature of 'learning' and 'adaptive behaviour' in their own right. These terms are ones which originally arose in the biological sciences to denote the plasticity of behaviour shown by an organism in its struggle to survive in a novel or changeable environment. The same terms have been carried over into psychology to denote the goal-seeking nature of animal behaviour, and have also been applied in the engineering sciences to systems designed to optimize their performance through interaction with their environment. Since these terms, or similar ones, form an essential part of the statement of objectives of major areas of research into human behaviour, adaptive control and artificial intelligence, it is reasonable to expect them to be capable of fairly exact definition. This is not so, however, and the terms are used very loosely with tacit switches of connotation, particularly between 'structural' and 'behavioural' aspects of adaption.

Stanier, at a symposium on "Adaptation in Micro-organisms" (1953), defined the term 'adaptive' as,

'the totality of the various processes of change which confer on an organism fitness to its environment'.

This definition contains all the elements essential to a taxonomy of adaptive behaviour. There is firstly the organism, which in the present discussion will be called a controller since both natural and artificial systems are being considered. There is secondly an environment, which term has entered the vocabularies of both psychology and engineering. There is thirdly the evaluative concept of 'fitness' to the environment, which reflects a description of the organism's behaviour in terms of goal-seeking, and entails a performance measure for the interaction between controller and environment. Finally there is the concept of 'change', in that an organism is not initially fitted for its environment but becomes so through a dynamic process of adaption.

Adaption, defined in this way, may be treated as a purely behavioural concept, since there is no necessity to introduce notions of how the organism adapts to its environment, or to argue a priori what factors will cause adaption to take place. Stanier's definition is informal, and non-operational in that it does not contain a decision procedure to determine when, or what form of, adaption takes place. Most attempts to treat learning and adaption more formally have gone beyond the observed behaviour and introduced structural or epistemological considerations, generally because their aims have been to model or predict behaviour. It is possible, however, to formalize the purely behavioural concept of adaption, into a rigorous framework on which to build a theory of training.

Section 2 of this Chapter is a review of the few previous attempts to establish operational definitions of adaptive, or goal-seeking, behaviour. Sections 3 and 4 present a new formulation of these definitions through an axiomatic approach to the description of learning behaviour.

2.1.1 The 'Analytical Biology' of Sommerhof

An early attempt to give a rigorous behavioural definition of the concepts of purpose and goal-seeking in biology and psychology was made by Sommerhof in his book 'Analytical Biology' (1950). Sommerhof considers an environment which has a number of states, a regulator (controller) which also has a number of states, and a set of outcomes which are determined by a combination of the environment's state and the controller's state. Some of these outcomes are satisfactory and satisfy a 'focal condition', whilst others are not. Sommerhof proposes that goal-seeking may be said to occur when there is 'directed correlation' between the states of the controller and those of the environment such that the outcome is satisfactory. He further proposes a 'degree of goal-directedness' in terms of the set of states of the environment for which the outcome is satisfactory.

Sommerhof qualifies his definition with some constraints designed to ensure non-triviality of the goal-directedness. The state of controller must not, in itself, ensure the satisfactoriness of the interaction - that is, the controller must, in some sense, 'take note' of the environment. The state of the controller and that of the environment must be 'epistemically independent' - that is, there must not be a natural physical connection between them such that the interaction is bound to be satisfactory. Finally, single occurrences of satisfactory interactions do not show goal-directed behaviour, but the directed correlation must exist between a number of environment and controller states.

Although Sommerhof introduces the term 'state' of the controller, this is irrelevant to his criterion for goal-directedness, which depends only upon the satisfactoriness of the interaction between controller and environment, and hence is purely behavioural. In control-theoretical terms, what he proposes is a 'sensitivity analysis' (Radanovic 1966) of the interaction between controller and environment, with goal-directedness being evinced by insensitivity of the satisfactoriness to disturbances of the environment.

Sommerhof's definition clearly applies to a simple servo-mechanism, and he gives a gun-aiming servo as an example of a goal-directed system. The constraints which he places upon the nature of the behaviour in order to distinguish between the 'goal-directedness' of the simple servo, and the 'non-goal-directedness' of, for example, a pendulum, do not appear to be essential, and detract from the main argument. The 'physical law' relating position and acceleration of a pendulum is due to its constrained motion in a gravitational field, and is thus a function of the structure of the pendulum and its relationship to its environment. Similarly, the corresponding law between the position and acceleration of the armature of a servo motor is a consequence of the structure of the servo system and its connections to a load. Any behaviour of any system, no matter how complex, is physically determined. Before the advent of Newtonian dynamics, the behaviour of pendulums was equally as mysterious as that of living creatures. It is true that lack of understanding of the mechanism of a phenomenom may lead us to give it undue weight and importance, but this cannot form the basis of a logical distinction.

The main criticism of Sommerhof's approach is that it does not take into account the dynamics of adaptive behaviour. A controller does not generally change its state instantaneously according to the state of the environment, and the manner in which it changes state is a function not only of the immediate environment but also of its previous states. It is this sequential dependence, or memory, inherent in the behaviour of most adaptive systems, which gives rise to transfer effects in training, and, indeed, to most of the complexity of adaptive behaviour.

2.1.2 Ashby's Formulation of 'Directed Correlation'

Ashby (1962) has given a set-theoretic formulation of Sommerhof's definition of goal-seeking behaviour, which is itself an advance in the theory of adaptive behaviour. Ashby considers a set of disturbances, D, which cause some changes in the environment. He defines these changes by mapping, Φ, from D into the set of possible values of the parameters of the environment, E. A disturbance, $d \in D$, at time t_o, produces an effect in the environment, $e \in E$, at time t_1, such that -

$$e' = \Phi(d) \qquad [2.1]$$

The parameters of the goal-seeking system are similarly specified by a set, F, and its behaviour is a response to d, which may be represented by a mapping, μ, from D into F, such that f, the system response, satisfies -

$$f = \mu(d). \qquad [2.2]$$

When the disturbance has evoked responses in the environment and goal-seeking system, $\Phi(d)$ and $\mu(d)$ respectively, then these two values interact to give some final outcome at time t_2. This corresponds to a mapping, η, from the product set, E x F, into Z, where Z is the set of all possible outcomes when E and F range uncorrelatedly over all their values. Within Z is a sub-set, G, of outcomes that satisfy the focal condition and are satisfactory.

Sommerhof's 'directed correlation' is now defined as being shown by μ, in respect of D, Φ, μ and G, if and only if -

$$\forall d \in D, \quad \eta(\Phi(d), \mu(d)) \in G \qquad [2.3]$$

- that is, for any disturbance the outcome is satisfactory.

Ashby manipulates this result into a neater, and more intuitively satisfying, form, using his development of the set-theoretic terminologies of Bourbaki and Riguet (Ashby 1962). In this notation a function, such as Φ, standing on its own denotes the sub-set of the product set between its domain and range such that the functional relationship is satisfied. The inverse, Φ^{-1}, denotes the corresponding sub-set of the product set between its range and domain. The composition of two such sets, $A \subset I \times J$ and $B \subset J \times K$, is denoted by A.B, and is a sub-set of the product set $I \times K$, such that -

$$\forall (i,k) \varepsilon A.B, \quad \exists j : (i,j) \varepsilon A, \quad (j,k) \varepsilon B \qquad [2.4]$$

- the set J is eliminated by this composition.

Using this notation, Ashby remarks that the expression, [1.3], for directed correlation can be simplified algebraically by noting that the set specified by -

$$[\forall d \ \varepsilon \ D, (\Phi(d), \mu(d))] \qquad [2.5]$$

- is identical with the set $\mu.\Phi^{-1}$, with D eliminated by composition. The criterion for directed correlation may be written -

$$\forall d \ \varepsilon D, \quad (\Phi \ (d), \ \mu(d)) \subset \eta^{-1}(G) \qquad [2.5]$$

- which is thus equivalent to -

$$\mu.\Phi^{-1} \subset \eta^{-1}(G) \qquad [2.6]$$

- which in its turn may be re-written -

$$\mu \subset \eta^{-1}(G).\Phi \qquad [2.7]$$

This is an elegant formulation of Sommerhof's definition of goal-directedness, in which the essential features stand out clearly and unnecessary use of the symbolism of differential calculus and continuous variables is avoided. In particular, although Ashby introduces times, t_o, t_1 and t_2, copied from Sommerhof's original argument, they are completely irrelevant and do not appear in the final result. The 'directed correlation' defining goal-seeking is presented in a completely abstract form as an inclusion relationship between sets forming the domains and ranges of certain mappings.

Stripped of the time-dependencies, the definition appears so simple as to be almost trivial. Indeed, if one labels the parameters of the environment by the disturbances that cause them (or consider the parameter of the controller to reflect that of the environment rather than that of the disturbance), then Φ becomes an identity mapping, and the criterion is merely that -

$$\forall d \ \varepsilon \ D, \qquad \eta(\mu(d)) \subset G \qquad [2.8]$$

- which states that the evaluation of the controller's policy is uniformly satisfactory. However, the simple and clear notation is in itself an invitation to introduce a state variable into the controller's parameter mapping. For example -

$$f_{n+1} = \mu(f_n, d_n) \qquad [2.9]$$

- which extends equation [1.2] to take account of an important aspect of the behaviour of the controller in time, previously obscured by the irrelevant introduction of t_o etc., that an adaptive controller becomes satisfactory through its experience of the environment, and may possibly become unsatisfactory again - adaption is a dynamic process.

In summary, Sommerhof's analysis is a considerable step towards a rigorous definition of goal-directed behaviour, but does not take into account the dynamics of adaption. Ashby's set-theoretic formulation of this analysis is a further advance in clarity of exposition, and enables the defects, and possible extensions, of the original definition to be clearly seen.

2.1.3 Behavioural Definitions of Adaption in Control Engineering

In the 1930's, about ten years after Watson laid the foundations of behaviourism is psychology, engineers such as Bode (1960, review) and Nyquist (1932) were establishing techniques for the analysis of the structure and behaviour of automatic control systems. Requirements for such systems to operate with plant whose characteristics were not known in advance, e.g. drives with varying frictional loads, led engineers in the 1950s to develop self-optimizing controllers which changed their parameters to maximize their performance. These controllers were termed 'adaptive', since many of the concepts in both standard control systems and self-optimizing controllers were derived by analogy with biological systems.

It is interesting to note that control engineers reached similar disagreements over the definition and application of the term 'adaptive' as have behavioural psychologists. In a panel discussion among many of the eminent pioneers of adaptive control, the position was summarized by the statement (Freeman 1963),

'the best way we can define an adaptive system today is to ask 500 people and we'll get 500 different answers and perhaps if we ask the same people the next day we'll get 500 additional answers'.

It is clear on reviewing the controversy that, although engineers were trying to achieve adaptive behaviour, in the sense of a changing control policy dependent upon the environment, the criticisms as to whether a controller was, or was not, adaptive were made largely on structural grounds. Thus, Raible (1963) claims that the Moe and Murphy (1962) 'self-adaptive' controller is not adaptive because it only identifies plant parameters on-line, whilst 'search and hill-climbing' are necessary for true adaption. Since control engineers are primarily involved in the design and fabrication of control systems, it was natural for the engineer to term a controller 'adaptive' if its structure had been designed to be so, even if its actual behaviour was extremely mal-adaptive.

Occasionally a distinction between behavioural and structural definitions is noted, but not carried to its logical conclusion. For example, Clark (1963) remarks that most attempts to define adaptive systems -

'seem to be of an anatomical nature... in terms of the physical features of the control system itself... I feel that the definition might better be based on a functional definition'.

He goes on to suggest -

'where we have a plant whose dynamic characteristics vary in time... if we can design a controller that will solve our problem to the satisfaction of all who are concerned with the problem, might it not be a functional definition of an adaptive control system'.

However, although such functional, or behavioural, approaches were proposed and distinguished from structural connotations of the term 'adaptive', the vagueness in the general use of the term continued and came to be accepted. Truxal (1963), in reviewing the field of adaptive control for a major Congress, re-iterated that an adaptive control system was one -

'designed from an adaptive viewpoint',

and stated that -

'While the literature of control theory is replete with arguments re the definition, progress in adaptive control theory has not been impeded by the failure of purists to reach universal agreement on an appropriate definition'.

The attitude engendered by failure to distinguish between structural and behavioural connotations is summed up by Florentin's (1962) wry remark at the end of a paper on an adaptive control system that, viewed in a certain way, his system appears -

'to be ordinary nonlinear feedback. It seems that any systematic formulation of the adaptive control problem leads to a meta-problem which is not adaptive.'

2.1.4 Zadeh's Definition of 'Adaptivity'

The distinction between structural and behavioural connotations of the term 'adaptive' was first made, in a control engineering context, by Zadeh (1963) in a short paper entitled, 'On the definition of adaptivity', which proposed a formal and completely behavioural intension of the term. He notes that -

'Most of the vagueness surrounding the notion of adaptivity is attirbutable to the lack of clear differentiation between the external manifestations of adaptive behaviour on the one hand and the internal mechanisms by which it is achieved on the other'.

He goes on to propose a characterization of adaptive behaviour, stating that, -

'our premise is that all systems are adaptive, and that the real question is what they are adaptive to, and to what extent'.

Zadeh frames his definition, as does Sommerhof, in terms of continuous functions and functionals. He considers a system subject to one of a set of specified input time functions, $u(t)$, defined on the semi-open interval, $t \geq 0$. A family of such inputs, which may have a probability measure defined over it, he defines as a <u>source</u>, $S\gamma$. For each source, he assumes that there is a performance criterion such that the behaviour of the system when connected to the source (that is, when receiving an input belonging to the family of inputs which defines the source) is, or is not, <u>acceptable</u>. Zadeh then defines 'adaptive'in these terms, -

'A system is adaptive with respect to a family of sources, $(S\gamma)$, and a criterion of acceptability, if it performs acceptably well with every source in the family $(S\gamma)$.'

This definition is very close indeed to that of Sommerhof, and lack of reference to the latter illustrates the neglect of his work whose publication proceeded that of Zadeh by some thirteen years. Zadeh does not attempt to give the term 'adaptive' an absolute meaning and exclude trivial cases, but states quite explicitly, -

'Under this definition every system is adaptive with respect to some set of sources and performance criterion. Thus what matters is not whether the system is adaptive, or not, but what are the sources and performance criteria to which it is adaptive.'

The same criticism may be leveled at Zadeh's definition as at Sommerhof's, that he does not take account of the dynamics of adaptive behaviour. Indeed, this possibility is excluded by defining inputs on the interval, $t \geq 0$, rather than some finite interval. Zadeh takes into account the time variation of performance with learning only through the suggestion that, -

'it is appropriate to use a performance function which assigns low weight to the performance in the initial stages of the learning process.'

It is interesting to note that Donalson and Kishi (1965) in a textbook on automatic control, give a version of Zadeh's definition of 'acceptability' in which they modify the input sources to be defined over a finite time interval. The reason for this is not stated but is inherent in their discussion, for they require an adaptive controller to modify its control action in an attempt to become an acceptably performing system. The acceptability of performance, in their structural definition, becomes something monitored by the controller, and hence regularly evaluated. This change of meaning, from an overall evaluation of the controller's ability to cope with its environment, to a local evaluation of whether it is, or is not, yet coping, it very important in the context of adaptive dynamics. Whilst their modification was necessary to the formulation of a structural definition, it could equally form the basis of a behavioural definition in which performance is monitored not by the controller, but rather by some outside observer.

Both Zadeh's and Donalson and Kishi's definitions of 'acceptability' are important and necessary to a theory of adaption, but a distinction must be made between them. In this thesis, the term 'acceptable' is applied to the interaction between controller and environment in a global sense, similar to that of Zadeh. Whilst the term 'satisfactory' is used in a local sense equivalent to Donalson and Kishi's 'acceptability'.

2.1.5 The Dynamics of Adaptive Behaviour

The adaptive control systems studied in engineering have not been very complex, and, typically, they have had the property that either they are able to adjust their parameters to a level suitable to their environment and maintain them there, or they are not, and the ones which are not able to do so tend not to be reported. Any further complexity of behaviour, such as sometimes adapting successfully and at other times not, or becoming satisfactory for a while and then becoming unsatisfactory again, would be regarded as a defect and removed if possible. Equally, complex behaviour based on interactions between environments, such as success in adapting to one environment only after having adapted to another, would not be noticed since the parameters of the controller would usually be reset upon changing the environment. Thus Zadeh's definition in terms of total success or total failure of adaption, for all time, is comprehensible in the light of engineering experience with adaptive systems.

In psychology, on the other hand, experience has been of systems whose level of complexity engineers are still many years from synthesizing. The human operator most certainly does not have a dichotomous capability in most tasks, and his performance of one is greatly affected by his experience of others. All animals show a similar complexity of behaviour, and it is interesting to compare some definitions of 'learning' by behavioural psychologists with those more formal definitions discussed so far.

Guthrie (1952) gives a very simple definition -

'These changes in behaviour which follow behaviour we call learning'.

Hunter (1934) attempts to exclude some behaviour, which is due to changes in peripheral structures, from the definition -

'We may say that learning is taking place whenever behaviour shows a progressive change or trend with a repetition of the same stimulating situation and when the change cannot be accounted for on the basis of fatigue of receptor and effector changes'.

McGeoch and Irion (1952) require a change in the evaluation of the behaviour, rather than merely a change in the behaviour itself -

'Learning, as we measure it, is a change in performance which occurs under conditions of practice'.

Thorpe (1956) requires these changes to be adaptive -

'We can define learning as that process which manifests itself by adaptive changes in individual behaviour as a result of experience'.

Whilst Bush and Mosteller (1955) deliberately exclude this requirement -

'We consider any systematic change in behaviour to be learning whether or not the change is adaptive, desirable for certain purposes, or in accordance with any such criteria'.

There is one important aspect of adaptive behaviour, common to these diverse definitions, which does not appear in Sommerhof's or Zadeh's formulations, and that is the nature of learning as a change in behaviour with practice or experience. There is one engineering paper where the essential nature of the dynamics of adaption is clearly stated. Martens (1959) has proposed a definition of 'machine learning' which is intended to provide a test to determine whether learning really does occur. He requires that a machine be able to adapt to two mutually incompatible criteria. Thus, in terms of Zadeh's definition, the controller would have to have an acceptable performance with each of two source/performance-criterion pairs, such that it is physically impossible for a system with a fixed control strategy to have an acceptable performance with both.

This distinction is not as fundamental as it might appear, since it is clear that the machine must have some means of distinguishing between the two source/performance-criterion pairs, and if this information is considered as part of the source then the adaption is once again trivial on Marten's criterion. More important, however, is the rationale behind the situation which Martens requires to be established as a criterion of real learning. This is such that the controller cannot be adapted to both objectives at the same time, but must become adapted to one after being adapted to the other, and vice versa. This connotation of adaptive dynamics is clear enough in structural definitions of adaptive controllers, and in the definitions of learning by behavioural psychologists, but it seems to have been omitted in more formal analyses of adaptive behaviour.

The objective of the theoretical studies described in the remainder of this chapter has been to extend the purely behavioural approach to the definition of adaption proposed by Sommerhof and Zadeh to a taxonomy of adaptive behaviour which takes full account of the dynamics of adaption, and in particular those aspects of the dynamics relevant to training.

2.2 An Axiomatic Basis for a Theory of Adaptive Behaviour

Given a situation which is to be regarded as that of a controller interacting with an environment, there are a number of decisions to be taken before a discussion of the adaptivity of the controller can begin. Within the theory of adaption these decisions have to satisfy certain logical constraints, but they are otherwise arbitrary. Within the context of human or animal behaviour, and its function, or of adaptive control practice, these decisions will certainly not be arbitrary, for the utility of adaptive concepts in specific situations will depend on them. Some of the decisions will be obvious and unmentioned, and others will be made explicit. Much of the early controversy over conflicting usage of the term 'adaptive' in engineering arose because the 'obvious', tacit decisions of one engineer were not those of another, or because disagreement over specific decisions was wrongly ascribed to the definition of adaption itself.

In attempting to quantify and formalize the application of the term 'adaptive' in psychology, it is reasonable to expect that confusion is likely to arise over similar issues, and particular emphasis is placed in the following discussion on a clear exposition of those aspects of adaption which are arbitrary and depend on agreed definitions. In the development proposed here, the arbitrary decisions are localized at the point where the interaction between controller and environment is divided into 'tasks', but in practice such extreme localization is unlikely to be apparent and the arbitrariness is spread more thinly.

Although the terminology used, of controller and environment, applies particularly well to perceptual-motor skills, the theory applies, and is intended to be applied, to all aspects of human behaviour involving purposeful interaction with some other system; the illustrative examples are chosen to emphasize this. It also applies to the analysis of any form of adaptive behaviour by animal or machine, and represents a unified approach to the study of adaption and learning.

2.2.1 The Adaptive Control Situation

The three elements of a control situation are a <u>controller</u> interacting with an <u>environment</u> for a <u>purpose</u>. In biological systems none of these may be well-defined, and the reasonable assignment of roles and purposes is the subject of experimental study and theoretical analysis (Gaines 1966 p.339). Even man-made systems do not necessarily split into controller and environment in an obvious way, especially when the whole system is synthetic and contains local regulatory control loops. In the case of the human operator, a similar consideration applies, on the motor side, to such systems as the spinal reflexes controlling muscle contraction, and, on the perceptual side, to the dark/light adaption of the eye. In considering learning within the central nervous system, these peripheral systems are often better thought of as part of the 'environment'.

However, the separation between controller and environment can generally be agreed upon, and they are, for the purposes of the theory, defined as black boxes with input and output such that the inputs of one are the outputs of the other. The following examples of environments for human controllers illustrate this concept together with the possible associated problems -

(i) The environment is a vehicle on a road. Its inputs and outputs clearly depend on whether it is being driven or being serviced. For a driver, the inputs to the environment will be through the steering wheel, clutch, etc., and the outputs from it will be visual images of the road ahead, speedometer, and so on, together with acceleration forces, etc.

(ii) The environment is a line figure in two dimensions subject to the transformations of Euclidean geometry. Its inputs are usually regarded as mathematical operators transforming the figure, and its output as a mathematical description of the transformed figure. Consideration of input/output at this level of abstraction is an example of neglecting peripheral dynamics, and can create problems. Although the mode of presentation is mathematically irrelevant, a figure presented pictorially may have an entirely different psychological effect from one presented as a set of incidence relationships. Whether the hand drawing a transformed figure is regarded as part of the environment of the human controller, or part of the controller itself, is an arbitrary decision, but it will have some effect on the analysis of the learning process in geometry.

(iii) The environment is another human, subject to natural language communication. Its inputs and outputs may be thought of as linguistic utterances, subject to the same considerations as in (ii).

In the context of a unified approach to natural and artificial systems, it is interesting to note that machines are also available for interaction with environments similar to these three. The first is the type of situation in which one reasonably expects to find some form of servo-mechanism with similar characteristics to the human operator. The second is a typical environment for Newell, Shaw and Simon's (1959) General Problem Solver, and the third is within the scope of Weizenbaum's (1966) ELIZA, although the level of conversation so far has been fairly mundane.

The purpose of each of the controllers in the environments above may be defined intensively, as motion from one location to another, or a proof that two angles are equal, or persuasion to perform some act. However, from the point of view of system behaviour, it is not necessary to define the purpose but only to give some prescription for saying when it has been achieved - that is, some performance measure is required. Sommerhof and Zadeh both define performance measures in numerical terms, and often they will be in that form, but, in fact, the only feature they use of the measures is a dichotic one - is the controller's performance up to the required standard or not. Thus the minimum requirement for the definition of a controller's purpose is a decision procedure which determines, at least occasionally, whether some segment of the interaction between controller and environment has been satisfactory or not.

2.2.2 Segmentation of the Interaction into 'Tasks'

For purposes of the behavioural definition of adaption, we are concerned with the manner in which the evaluation of a controller's interaction with its environment changes as a function of that interaction. The expected behaviour of an adaptive controller when coupled to an environment is that, if its control poicy is not satisfactory for the environment, then it will eventually become so. Thus it must be possible to segment the interaction between controller and environment into at least two phases, in the first of which it is not satisfactory and in the second of which it has become so.

This segmentation of the interaction between controller and environment is inherent in the concept that an adaptive controller 'becomes satisfactory', rather than just 'is satisfactory', and is fundamental to the analysis of adaptive dynamics. A further dynamic feature of the behaviour is that the controller should remain satisfactory once it has become so; that is, it should reach a stable condition of satisfactoriness.

The segmentation is clearly present in Sommerhof's concept of a 'focal condition' which must be attained, but he is only concerned with whether it is attained and not with what happens before or afterwards. This omission probably occurs because his theory develops from a cosnideration of such problems as firing a gun at a target, where the line of sight has to satisfy a terminal, or 'focal', condition only at the moment of firing. Zadeh, on the other hand, although his definition strongly resembles that of Sommerhof, defines satisfactoriness over a semi-infinite interval rather than as a terminal condition, and makes some concession to the dynamics of adaption by suggesting that the performance criterion should attach less weight to the early stages of learning. These two, taken together, imply that the controller is expected to become adapted, and is not 'acceptable' unless it remains so.

In order to consider controllers which become satisfactory and the relapse, or to consider the effect of learning in one environment on later learning in another, it is necessary to extend the basic segmentation of the interaction which is inherent in the concept of becoming adapted, and analyse the changing evaluation of the controller's performance in greater detail. Donalson and Kishi, in their unmarked variation of Zadeh's semi-infinite interval to a finite interval, do just this, primarily because they are considering the structure of a controller which is monitoring its own behaviour. The introduction of a defined time interval as a criterion for segmenting the interaction is unnecessarily restrictive, however, and the following definition of a 'task' is a weak extension of the segmentation which is sufficient to form the basis of a taxonomy of adaptive behaviour.

2.2.3 Definition of a 'Task'

A task is a segment of the interaction between controller and environment for which it is possible to say whether or not the controller has performed satisfactorily. Equivalence relations between tasks (so that it is possible to say, for instance, that an interaction consists of the same task repeated several times) are arbitrary, but will generally follow the natural relationships between different types of controlled system. A task will typically consist of some specification of plant parameters, initial conditions and period of interaction, together with a tolerable performance level above which a control policy is considered satisfactory.

Within the theory of adaption a 'task' is restricted only by the necessity for some procedure to determine whether a particular interaction between controller and environment which constitutes a task is, or is not, satisfactory. In practice, it is convenient to choose the set of tasks in such a way that the segmentation of the interaction into tasks is unique. Any interaction may then be regarded as the performance by the controller of a set of well-defined tasks. Since the theory of adaption deals only with these tasks and the satisfactoriness of the controller in performing them, it is obviously of practical importance that they should be chosen to give adequate information about those aspects of the adaptive behaviour which are of interest. These are meta-theoretic considerations, however, and there is no postulate that any system behaviour can be mapped uniquely onto some member of the free semi-group generated by a set of tasks. The utility of the theory in any problems of practical interest is an empirical finding, and hence the informal justification given for the steps taken so far and the examples given later.

The segmentation of an interaction into tasks may be performed in many ways - the time of interaction between controller and environment may be fixed - a criterion for the termination in terms of the behaviour itself may be given - the interaction may be terminated as soon as a decision can be made about the satisfactoriness of the controller. The 'termination' itself may be purely conceptual, a convenient division of a continuous sequence of behaviour into separate sub-sequences, or it may have a physical reality in that the plant is modified at the termination of a interaction.

The following sections contain two examples of the segmentation of an interaction into tasks, and the second example includes a description of a simple adaptive system whose behaviour is used later to illustrate the definitions.

2.2.4 A Set of Tasks for a Single-Input, Single-Output System

Consider the stable, noiseless, second-order plant shown in Figure 2-1, consisting of two integrators in cascade with feedback from the output. Its parameters are the gain, undamped natural frequency and damping ratio. Let a task be defined by ascribing values to these three parameters and to the initial values of the integrators,

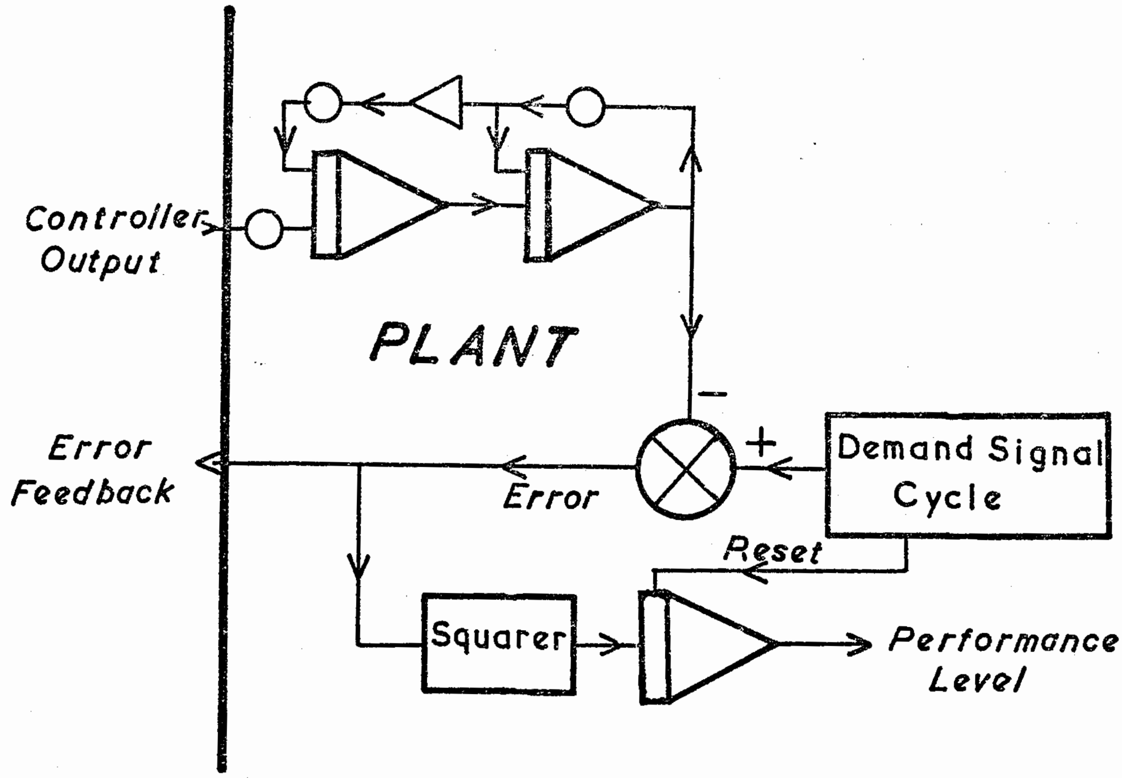


Figure 2-1 A 'Task' Generator for an Adaptive Controller

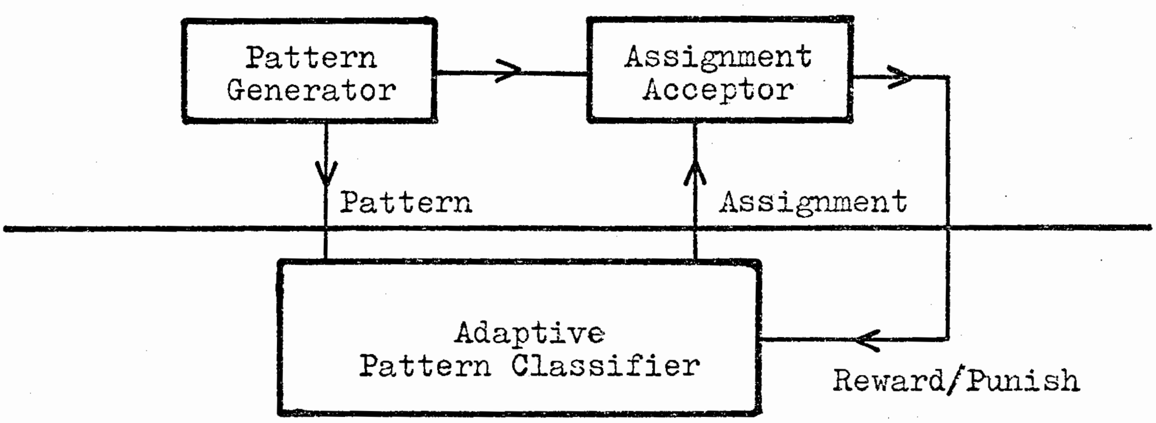


Figure 2-2 A 'Task' Generator for a Pattern Classifier

together with a time-varying demand signal, $f(t)$ for $0 \leq t \leq T$, and a decision procedure such that an interaction is satisfactory if, and only if, -

$$\frac{1}{T} \int_{t_o}^{t_o + T} \left[f(t-t_o) - x(t)\right]^2 dt < E^2 \qquad [2.10]$$

- where t_o is the time at which the interaction starts, $x(t)$ is the output of the plant, and $E > 0$ is some tolerance on the r.m.s. error.

To test the adaptivity of a controller to the plant and demand conditions specified by such a task, it is connected to the plant and the demand signal cycled with period T. After every cycle the task is complete, and the r.m.s. error during that cycle determines whether or not the controller has performed satisfactorily. If the controller is adaptive the r.m.s. error in each cycle might be expected to decrease, and hence (by suitable choice of E) the controller will be unsatisfactory initially, but after a number of repetitions of the task it will become satisfactory and remain so.

Many other forms of adaptive behaviour might arise, however - the controller could be always satisfactory or always unsatisfactory - it could start by being satisfactory and become unsatisfactory, never settling at one or the other. If other tasks with different values of the plant parameters or demand signals were interpolated, then the range of possible behaviour would become far greater still. It is the description of this variety of possible behaviours which concerns the behavioural theory of adaption.

2.2.5 Example of a Set of Tasks for an Adaptive Pattern-Classifier

The two integrator environment described in Section 2.2.4 is a typical continuous control system for both the human operator and automatic control systems. The analysis of adaption through segmentation of the interaction between controller and environment into 'tasks', however, applies equally well to discrete, problem-solving environments, such as those involved in pattern-recognition. The following example of a simple, perceptron-like pattern-classifier, learning to dichotomize patterns represented by binary vectors, not only exemplifies an alternative form of task, but also demonstrates sufficient variety and complexity of adaptive behaviour to illustrate later definitions of modes of adaption.

The problem of the pattern-classifier is to assign each of the patterns in an input stream to one of two categories. During the learning phase, it is supplied with information (reward/punishment) as to whether each of its assignments is correct, and uses this to change its categorization policy. Thus the environment of the machine, as shown in Figure 2-2 , consists of a generator of patterns at its input, an acceptor of assignments at its output, and a source of performance feedback at its reward/punishment input. The performance measure for a period of interaction will normally be based on the correctness of its decisions, e.g. the proportion correct.

A typical 'task' for the pattern-classifier may be defined by a set of input patterns, such as pattern_A followed by pattern_B, which may be written - $T_1 = (B,A)$, together with a performance criterion. There are four possible, non-trivial, performance criteria, that the performance for T_1 is satisfactory if - the category to which A is assigned is correct - that to which B is assigned is correct - either is correct - or both are correct. Given a sufficient set of pattern sequences defining tasks, any stream of input patterns may be split up into segments corresponding to these tasks, and hence the interaction between the pattern-classifier and its environment may be segmented as previously described.

This form of pattern-classifier has been used in the present study to elucidate some of the phenomena of adaptive behaviour, whilst the more continuous control task described in Section 2.2.4 typifies another experimental situation used in the study to investigate human and machine learning under various training regimes.

2.2.6 Probability and Indeterminacy in the Definition of a Task

It has been tacitly assumed in the Section 2.2.4 that the initial conditons of the plant may be dropped from the definition of a task, even though this makes the satisfactoriness of the interaction indeterminate. It is often the case that certain potential parameters of a task may be regarded as irrelevant to its description, because they do not appreciably affect the outcome of the decision procedure for evaluating the performance of the controller in the task. In the last example, if $\hat{T}$ is long compared with the period of the plant, then the initial conditions will have very little effect on the performance integral of inequality [2.10] . In any practical application there will always be an effect of experimental error on the decision procedure, and indeterminacies having effects comparable with this may be neglected.

It may, however, be desirable to regard as irrelevant parameters which do affect the outcome of the decision procedure - for example, a task might be defined, in a similar way to Zadeh's 'source', as a set of fixed plant parameters together with a class of possible demand signals, perhaps with a probability measure. The decision as to a controller's satisfactoriness then becomes indeterminate or may be based on an ensemble measure of its performance - for example, the maximum or expected r.m.s. error. Indeterminacy may be accepted as a third form of evaluation, or may be removed by deciding that any unsatisfactory evaluation in an ensemble gives rise to an overall unsatisfactory evaluation. Any ensemble evaluation, although acceptable in theory, leads to acute experimental problems in the evaluation of adaptive behaviour, however, because it is impossible, in general, to reset the state of an adaptive system and repeat the same experiment a number of times. Attempts are made to overcome these problems in practice by the use of populations consisting of different adaptive systems assumed to be exemplars of a single system. This has obvious dangers in itself, but is again a meta-theoretic procedure.

Thus, the concept of a 'task' encompasses specific descriptions of the possible range of controlled systems, based on distributions over plant parameters, demand signals, disturbances and so on, and, through the segmentation of interactions, it enables the behaviour of a controller coupled to an environment to be regarded as the performance of a sequence of tasks, for each of which it is, or is not, satisfactory. A calculus of adaptive behaviour based on these concepts may be used to evaluate the relative adaptivity of various controllers only in terms of their ultimate satisfactoriness for given tasks, and the path-length (number of tasks) before this is achieved. More complex cost-functions for trajectories of adaptive behaviour could readily be defined, but at our present state of knowledge they are not justified.

2.3 The Definition of Modes of Adaption

The basic concept of a "task" introduced in the previous sections may be used as the foundation for definitions of different modes of adaptive behaviour and relationships between them. The fundamental situation with which an adaptive controller is expected to cope is to be coupled to a fixed environment and learn to control it satisfactorily. The interaction is equivalent to the controller performing a sequence of tasks consisting of the same task repeated indefinitely. Such repetetive task sequences play a special role in the theory of adaptive behaviour because they correspond to the characterization of an adaptive system by its changing response to the same situation. The repetetive task sequence takes the place of the static 'source', or environment, in Zadeh's definition of adaptivity, but is exactly equivalent to it with the addition of segmentation. The importance of this segmentation in the study of adaptive dynamics begins to become apparent in the following version of Zadeh's definition of 'acceptability'.

2.3.1 Definition of an Acceptable Interaction

An interaction between controller and environment consisting of the repetition of a single task is acceptable if it is eventually always satisfactory.

Thus, in an acceptable interaction, the initial performance of the controller does not matter, and for a number of repetitions of the task it may be satisfactory, unsatisfactory, or waver between the two. However, it must eventually become satisfactory and remain so - an acceptable interaction is one which reaches a stable condition of satisfactoriness. In this condition we may say that the controller has become 'adapted' to the task.

2.3.2 Definition of a Controller Adapted to a Task

An interaction between controller and environment consisting of the repetition of a single task is immediately acceptable if it is always satisfactory - an immediately acceptable interaction is obviously acceptable. A controller in such a condition that it would have an immediately acceptable interaction with a task is adapted to that task.

2.3.3 Dynamics of Adaption

The concepts of acceptability and adaptedness concern the performance by the controller of a single task, but the adaptivity of the controller is generally advantageous because the particular task it must perform is incompletely specified or may change. The possible varieties of adaptive behaviour in a control situation involving performance of any one of a number of tasks are many, but there are three modes of adaption of particular interest which are defined in the following sections.

Very often an adaptive controller will be required to perform a single task which will not change, but whose characteristics cannot be specified in advance. It must be capable of having an acceptable interaction with any of a range of tasks, but need not necessarily be capable of adapting to a sequence of different tasks. This mode of behaviour is characterized in the following definition.

2.3.4 Definition of a Potentially Adaptive Controller

A controller in such a condition that it will have an acceptable interaction with any one of a set of tasks is potentially adaptive to that set of tasks.

A potentially adaptive controller fulfils one function of an adaptive system in compensating for ignorance about the nature of its environment. It will not necessarily fulfil another major function by performing satisfactorily in a changing situation, since there is no implication that, having adapted to one task, it remains potentially adaptive to others. Potential adaption is implied in statments like, 'a shoe adapts itself to the shape of a foot', and is the weakest form of goal-attainment to merit the designation 'adaptive'.

A controller which has to perform satisfactorily in a changing situation must not only adapt to its immediate task, but must also remain potentially adaptive to the other tasks it may meet. This mode of behaviour is characterized in the following definitions of compatible adaptivity.

2.3.5 Definition of a Compatibly Adapted Controller

A controller is compatibly adapted to one task with respect to set of tasks if, in an interaction consisting of the repetition of that task, it is not only always satisfactory but also potentially adaptive to the set of tasks.

A controller which is compatibly adapted to one task with respect to a set of tasks is clearly 'adapted' to the first task, and, whilst performing it, never loses its capability for adapting to any of the set of tasks. If the controller is to function in a changing environment, it must always, when adapting to one task, become compatibly adapted with respect to the remaining tasks it may meet - that is, it must be potentially adaptive to all its possible tasks no matter which of them it has previously performed.

2.3.6 Definition of a Compatibly Adaptive Controller

A controller is compatibly adaptive to a set of tasks if, given any sequence of tasks from that set, it remains potentially adaptive to the set of tasks.

Thus, a controller which is compatibly adaptive to a set of tasks will have an acceptable interaction with any one of them, and, no matter how it becomes adapted to one of them, it will be compatibly adapted with respect to the remainder.

The phenomenom of potential, but not compatible, adaptivity is very interesting and quite common in biological adaption, both phylogenic and ontogenic, and in animal, even human behaviour. A microbe may adapt to a new culture to such an extent that it becomes dependent on it and dies when returned to its former environment. A species may evolve under the influence of a climatic change but come to an evolutionary dead end such that it cannot reverse the process when the climate returns to its original form. A human subject may be able to learn to solve either of two problem types equally well, but, after having learnt one and performed it for some time, he acquires a 'set' towards the processes involved in its solution and, when given the other, cannot discover how to solve it.

A controller which is compatibly adaptive to a set of tasks is not necessarily able to become adapted simultaneously to all of them. It is, however, quite possible for two tasks to be so similar that a controller which is adapted to one may also be adapted to the other. This mode of behaviour is characterized in the following definitions of joint adaptivity.

2.3.7 Definition of a Jointly Adapted Controller

A controller is jointly adapted to a set of tasks if, given any sequence of tasks from that set, it remains adapted to every member of the set.

Thus, a controller which is jointly adapted to a set of tasks will be always satisfactory given any sequence of those tasks. This is a very strong condition, and an even stronger one is that when a controller adapts to any one of a set of tasks it should eventually become jointly adapted to all of them.

2.3.8 Definition of a Jointly Adaptive Controller

A controller is jointly adaptive to a set of tasks if it is both compatibly adaptive to the set and, during an acceptable interaction with any task in the set, it eventually becomes jointly adapted to the whole set.

Joint adaptivity would obviously be expected to be far rarer than compatible adaptivity, since it requires either that the same control policy be appropriate to two different environments, or that a very rapid change of policy be made when the environment changes. The former phenomenom arises, fairly trivially, when the environments are themselves very similar - for example, riding a scooter and then a motor-bike - and the latter is of greater interest. The occurrence of joint adaption when it is logically impossible for the same control policy to be satisfactory in both environments - for example, when the joy-stick in a tracking task is reversed in sense - implies that the learning of one skill does not greatly interfere with the skill already acquired.

2.3.9 Inter-relationships Between Different Modes of Adaption

The preceeding definitions of different modes of adaptive behaviour have been given in order of increasing strength, for if a controller is jointly adaptive to a set of tasks it is also compatibly adaptive to them, and if a controller is compatibly adaptive to a set of tasks it is also potentially adaptive to them -

Jointly Adaptive $\rightarrow$ Compatibly Adaptive $\rightarrow$ Potentially Adaptive.

These three modes of adaption are by no means exhaustive, and many variations are possible, defining other modes of adaptive behaviour. However, many of these would be regarded as pathological, serving no useful function and having no correspondence to a well defined type of adaptive behaviour. Other, more interesting varieties of adaptive behaviour may be described in terms of the modes already defined. The questions which are usually of interest are whether a controller is potentially adaptive to a set of tasks, and whether it has any states in which it is compatibly or jointly adaptive to them. The modes of adaption which have been defined form explicata for all the common stereotypes of adaptive behaviour.

2.3.10 Inter-relationships Between Tasks

The definitions of adapted, compatibly adapted, jointly adapted, and potentially, compatibly and jointly adaptive, may be used to define binary relations on the set of tasks relative to a given controller - for example, $task_1$ is related to $task_2$ if, and only if, the controller is compatibly adapted to $task_1$ with respect to $task_2$. All six relations are reflexive, and only that induced by 'compatibly adapted' is not symmetric. However, only 'adapted' and 'potentially adaptive' induce relations which are also transitive (and hence are equivalence relations). For instance, a controller may be jointly adapted to $task_1$ and $task_2$, and also jointly adapted to $task_2$ and $task_3$, but given a sequence containing both $task_1$ and $task_3$ there is no reason why even its potential adaptivity to both tasks should not disappear. It is this lack of equivalence relations which gives adaptive behaviour its extraordinary richness. A controller which showed no 'pathological' behaviour would be very rare, although the more drastic forms would not be expected to occur. For example, the relation induced by 'compatibly adaptive' ought to be one of equivalence, because no sequence of normal tasks should be able to destroy a controller's ability to adapt to one of them.

The binary relations over tasks, generated by consideration of a controller's adaptive behaviour, may be used as the basis for a taxonomy of environments according to the problems involved in adapting to them. For example, if a family of controllers, in adapting to one task, became jointly adapted to another, then it would be reasonable to suppose that the two tasks were very similar in the control strategy required. If it were found that this was not so, but that potential adaption to one task implied potential adaption to the other, then it would be reasonable to suppose that the two tasks were related in the adaptive capabilities which they required - this is the basis of some 'intelligence tests'.

2.3.11 Arbitrariness and Triviality in the Definitions of 'Adaptive'

The behavioural definitions of 'adapted' and 'adaptive' contain an arbitrary element because the classification of 'tasks' is left undefined, and the segmentation of the interaction between controller and encironment is at will within the (very weak) constraints of Section 2.2.3. This arbitrariness need not cause difficulty in the analysis of adaptive behaviour, provided it is accepted that at some stage in the discussion of an adaptive controller and its behaviour this classification must be agreed. Much of the early controversy over the application of the term 'adaptive' arose because the 'obvious', tacit classification of one psychologist, or engineer, was not that of another, or because disagreement over such a classification was wrongly ascribed to the definition of adaption itself.

Even when the arbitrariness in the definitions is accepted, there remains the possibility that some types of adaptive behaviour may be 'trivial'. For example, 'jointly adaptive' is an apparently very strong conditon which may be quite trivial in reality - for example, the tasks to which a controller becomes jointly adapted may be completely equivalent and need not be distinguished. 'Potentially adaptive' is a very weak condition which may often be regarded as trivial, because it is shown by systems undergoing an irreversible descent to equilibrium. 'Compatibly adaptive' adds the requirement of reversibility, and is closest to what is commonly regarded as being 'really' adaptive. However, although a compatibly adaptive controller shows all the behaviour which one would expect of a 'really' adapted to all of its tasks all the time, and hence show no adaptive dynamics - it is just a very good, but static, controller!

In testing the behaviour of an animal, or automatic controller, for 'adaption' or 'learning' it may be desirable to eliminate this 'trivial' adaption. It is meaningful, for example, to ask whether an animal performs well in two different situations because it has a policy suited to both, or because it changes its policy according to the situation. To force a controller to show adaptive dynamics, one might say that it is 'really adaptive' to a task if it was an acceptable, but not immediately acceptable, interaction consisting of the repetition of that task. The controller cannot then be a static system which happens to have a suitable control policy.

Sommerhof (Section 2.1.1) formulates this non-triviality requirement by requiring that the state of the controller must not, in itself, ensure the satisfactoriness of the interaction. A particularly neat, and operational, formulation of the same requirement is given by Martens (1959) in his definition of 'machine learning'. He requires that the controller have acceptable interactions judged by two incompatible performance criteria. It cannot then have a policy which satisfies both criteria at once, and hence it is not immediately acceptable for one task when it is adapted to the other. This may be re-phrased that a controller is 'really adaptive' if it is compatibly adaptive to a pair of tasks such that it is impossible for any controller to be jointly adapted to them both. However, it will be noted that all these specific formulations are encompassed in the present one.

2.4. An Automata-Theoretic Formulation of Adaptive Behaviour

In Sections 2.2 and 2.3 an axiomatic formulation of the nature of adaptive behaviour has been given in terms of tasks and the satisfactoriness of their performance. This formulation has been deliberately maintained in non-mathematical form, partly for purposes of presentation, but also because premature mathematization obscures the basic logical nature of the definitions which are independent of the notation used in their expression. However, to take the approach established in these sections a stage further and build on it a theory of training, some conciseness of notation is necessary, and this best comes from an automatic-theoretic definition of the minimum-state, observable structure which is cybernetically equivalent to the system showing adaptive behaviour (using the terminology and results established in Appendix 3). It is important to emphasize that the results of Appendix 3 justify the remark that the analysis is still purely behavioural, in that the automation used is one derived purely from descriptions of behaviour and acts only as a convenient basis for discussing these behaviours.

2.4.1 Adaption Automata

From Section 2.3, the adaptive behaviour of a system is completely described, in terms of this analysis, by a sequence of descriptions of the task given and the satisfactoriness of the controller in performing the task. Thus, a minimum sufficient set of descriptors for adaptive behaviour is the set -

$$D = \{(t,p^+), (t,p^-) : t \in T\} \qquad [2.11]$$

- where t is a task belonging to the set of all tasks which may be given, T, and p, p^-, are the two possible outcomes, 'satisfactory', and 'unsatisfactory', respectively.

The free semigroup, F_D, generated by this set of descriptors, consists of all possible sequences of tasks, together with all possible outcomes in terms of the satisfactoriness of the interaction. The sub-set of this semigroup, which constitutes the observed behaviour of some adaptive system, clearly satisfies postulates (i) and (ii) of Section A3.2. It will be assumed that it also satisfies postulate (iii) of Section A3.3 - any real behaviour is necessarily finite and the derived automaton an approximation (Section A3.3.6). Hence, using the construction of Section A3.3, an automaton may be derived which is cybernetically equivalent to the observed system - this will be called the adaption-automaton of the adaptive system defined by its behaviour. The adaption-automaton will give rise to the same sequence of satisfactoriness as the adaptive system when it performs the same sequence of tasks. In general, it will be an indeterminate automaton, and even if its state is known its next state cannot necessarily be predicted from the task given, but it is always observable, in that its present state can be determined provided suggicient of its past behaviour has been observed.

An adaption-automaton is a state-restricted automaton with a possibly infinite set of states, S, probably a finite set of inputs, T, and two possible outputs, $P \equiv \{p^+, p^-\}$. If the automaton is in state, $s \in S$, is given the task, $t \in T$, then its next state, $s' \in S$, and its output, $p \in P$, belong to the transition, and output, sets, respectively:-

$$s' \in \sigma(s,t) \tag{2.12}$$

$$p \in \pi(s,t) \tag{2.13}$$

Since it is the effect of sequences of tasks, especially those generated by the repetition of a single task, which are of interest, it is important to have a clear notation distinguishing between tasks, sets of tasks, sequences of tasks, and sets of sequences of tasks. A sequence consisting of the task t_1 followed by the task t_2 will be written t_1t_2 (with the obvious extension to longer sequences), a sequence consisting of the task, t, repeated n times will be written t^n. A typical set of tasks will be represented by the letter, T(T_1, T_2, etc); a typical sequence of tasks will be represented by the letter, u (u_1, u_2, etc). The set of sequences generated by the set of tasks, T, as free generators will be written U(T) - that is, the set of all possible sequences of tasks which may be formed using members of T. The function, σ, has an obvious extension from tasks to task sequences - if $u = tu'$, then:-

$$\sigma(s,u) \equiv \{\sigma(s',u') : s' \varepsilon\ \sigma(s,t)\} \qquad [2.14]$$

Having established this structure, it is possible to give set-theoretic definitions of all the various modes of adaption defined in Chapter 2. These may now be characterized by the sets of states of the automaton in which the behaviour can arise which satisfies the constraints of the previous definitions.

2.4.2 Adaption Sets

Let $$W(T) = \{s : \forall t \varepsilon\ T,\ \pi(s,t) = p^+\} \qquad [2.15]$$

that is, W(T) is the set of states in which the controller will have a <u>satisfactory interaction</u> given any task from the set, T.

Let $$A(t) = \{s : \forall n,\ \sigma(s,t^n) \subset W(t)\} \qquad [2.16]$$

that is, A(T) is the set of states in which the controller is <u>adapted</u> to the task, t, because, given a sequence of task consisting of t repeated, the interaction is always satisfactory.

Let $$P(T) = \{s : \forall t \varepsilon\ T, \exists N : \sigma(s,t^N) \subset A(t)\} \qquad [2.17]$$

that is, P(T) is the set of states in which the controller is <u>potentially adaptive</u> to the set of tasks, T, because, given any sequence of tasks consisting of a member of T repeated, the interaction is eventually always satisfactory.

Let $$C_A(t,T) = \{s : \forall t\ \varepsilon T,\ n,\ \sigma(s,t^n) \subset W(t) \cap P(T)\} \qquad [2.18]$$

that is, $C_A(t,T)$ is the set of states in which the controller is <u>compatibly adapted</u> to the task, t, with respect to the set of tasks, T because, given a sequence of tasks consisting of t repeated, the interaction is always satisfactory and the state remains with the sub-set which is potentially adaptive to the set of tasks, T.

Let $$C(T) = \{s : \forall u\ \varepsilon U(T),\ \sigma(s,u) \subset P(T)\} \qquad [2.19]$$

that is, C(T) is the set of states in which the controller is <u>compatibly adaptive</u> to the set of tasks, T, because, given any sequence of tasks consisting of members of T, its state remains in the sub-set which is potentially adaptive to the set of tasks, T.

Let $$J_A(T) = \{s : \forall t \varepsilon\ T,\ u \varepsilon U(T),\ \sigma(s,u) \subset A(t)\} \qquad [2.20]$$

that is, $J_A(T)$ is the set of states in which the controller is <u>jointly adapted</u> to the set of tasks, T, because, given any sequence of tasks consisting of members of T, its state remains in the sub-set which is adapted to each member of T.

Let $J(T) = \{s: \exists N, \forall t\varepsilon T, u\varepsilon U(T), \sigma(s,ut^N) \subset J_A(T)\}$ [2.21]

that is, J(T) is the set of states in which the controller is jointly adaptive to the set of tasks, T, because, given any sequence of tasks consisting of members of T, in further adapting to any one of these tasks it becomes jointly adapted to them all.

The definitions of this section include all the modes of adaption previously described in Section 2.3. The inter-relationships between the various modes, briefly discussed in Section 2.3.9, now appear as inclusion relations between the adaption sets. The most important of these are:-

$$W(T_1 \cup T_2) = W(T_1) \cap W(T_2) \quad [2.22]$$

$$P(T_1 \cup T_2) = P(T_1) \cap P(T_2) \quad [2.23]$$

$$C(T_1 \cup T_2) \subset C(T_1) \cap C(T_2) \quad [2.24]$$

$$J(T_1 \cup T_2) \subset J(T_1) \cap J(T_2) \quad [2.25]$$

$$A(t) \subset W(t) \cap P(t) \quad [2.26]$$

$$C_A(t,T) = A(t) \cap P(T) \quad [2.27]$$

$$P(t_1 \cup t_2) \supset A(t_1) \cap A(t_2) \supset J_A(t_1 \cup t_2) \quad [2.28]$$

$$P(T) \supset C(T) \supset J(T) \quad [2.29]$$

2.4.3 Trajectories in the State-Space of an Adaption-Automaton

These relations, and the process of adaption itself, appear most clearly in diagrams of the adaption sets as sub-sets in the state-space of the adaption-automaton. In the same diagram, the dynamic behaviour of the adaptive system may be shown as the trajectory of states through which it passes in adapting to a sequence of tasks. The set of all states of adaption-automaton is shown as a rectangular region in Figure 2-3. From an intial state, such as Y, repetition of a task, t, generates a trajectory of states within this region - the states are assumed to be assigned to points in the region such that, without loss of generality for a single task, the trajectory appears as a connected path. Within the region are delimited those states, W(T), for which the controller is satisfactory when given the task, t. The states, A(t), for which the controller is adapted to the task, t, from a sub-set of these, since a trajectory starting in the adapted region must always remain satisfactory. The states, P(t), for which the controller is potentially adaptive to the task, t, form a third region, enclosing the one in which it is adapted to t. Figure 2-3 includes all

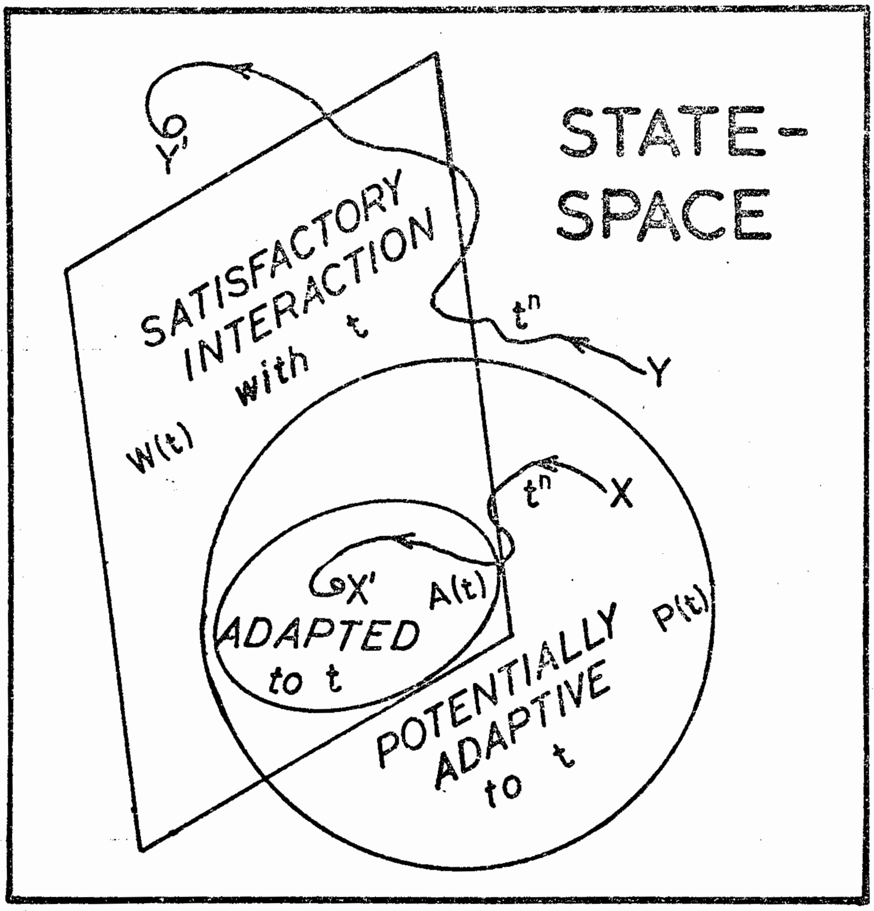


Figure 2-3 Adaption to a Single Task

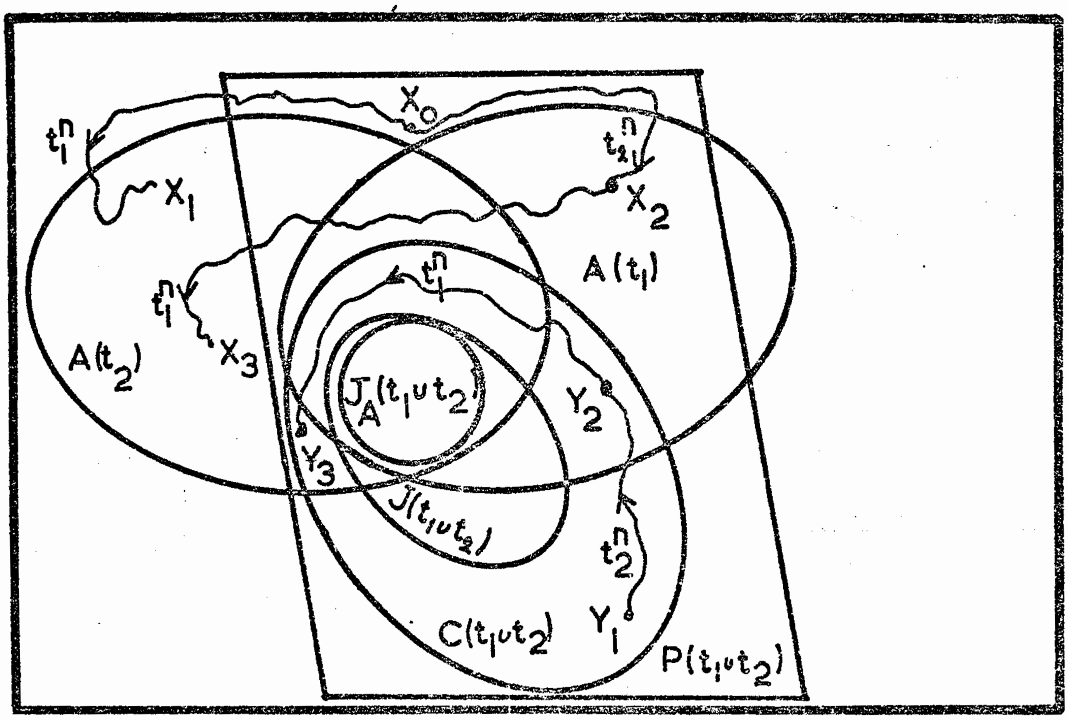


Figure 2-4 Adaption to a Pair of Tasks

the adaption sets relevant to a controller adapting to a single task.

A trajectory through the state-space, generated by giving the controller the task, t, many times in succession, t^n, will show the following behaviour:-

started outside the potentially adaptive region, at Y, it may enter the region of satisfactory interaction, but must eventually always leave it;

started within the potentially adaptive region, at X, it will remain that region, perhaps oscillating between being satisfactory, and being unsatisfactory, for a while, but eventually entering the adapted region, where it is always satisfactory, and never leaving it.

Some of the relations between potential, compatible, and joint adaption for two tasks, t_1 and t_2, are illustrated in Figure 2-4. The space of all states has been split into the regions, $A(t_1)$, $A(t_2)$, $P(t_1 \cup t_2)$, $C(t_1 \cup t_2)$, $J(t_1 \cup t_2)$ and $J_A(t_1 \cup t_2)$, and inclusion relations, [3.28] and [3.29] are clearly illustrated. Since two tasks are involved, trajectories in the state-space may show a more complex variety of behaviour than for a single task. For example, since the state, X_0, is within the potentially adaptive region for both tasks, $P(t_1 \cup t_2)$, but outside the compatibly adaptive region, $C(t_1 \cup t_2)$, trajectories generated by the task sequences, t_1^n or t_2^n, will eventually enter the adapted regions, $A(t_1)$ or $A(t_2)$, respectively, but, in so doing, they must also eventually leave $P(t_1 \cup t_2)$. This loss of potential adaptivity to one task may not take place immediately on first adapting to the other (as in the trajectory from X_0 to X_1), but may be dependent on adaption taking place to both taks. For example, the trajectory from X_0 to X_2 under t_2^n enters $A(t_2)$ within $P(t_1 \cup t_2)$, and hence is within $C_A(t_1 \cup t_2)$. On taking advantage of the residual adaptivity to t_1, however, the trajectory proceeds to X_3 which is not within $P(t_1 \cup t_2)$, so that the controller cannot again become adapted to t_2.

A trajectory starting from Y_1 within the compatibly adaptive region, $C(t_1 \cup t_2)$ cannot leave the region of potential adaption to both tasks, $P(t_1 \cup t_2)$, no matter what sequence of tasks of the form, $(t_1+t_2)^*$, are given. Trajectories originating from within $C(t_1 \cup t_2)$ by t_1^n or t_2^n will drive the state of the adaption-automaton back and forth between the regions, $A(t_1)$ and $A(t_2)$, respectively. If such trajectories should remain within the region $A(t_1) \cap A(t_2)$, so that the controller is always adapted to both tasks, then they are within the region of joint adaption, $J_A(t_1 \cup t_2)$, and may have originated in the region of joint adaptivity, $J(t_1 \cup t_2)$. They need not necessarily have done so, since, whilst all trajectories generated by sequences of the form $(t_1+t_2)^*$ from within $J(t_1 \cup t_2)$ must eventually enter $J_A(t_1 \cup t_2)$, there are states outside $J(t_1 \cup t_2)$ from which some trajectories of this type will enter $J_A(t_1 \cup t_2)$.

For a given adaption-automaton, certain of the adaption sub-sets will generally be empty, and the definitions do not imply that the corresponding states exist, but merely classify them should they be present. The two Figures, 2-3 and 2-4, are drawn as if the state-space had a convenient topology under the action of t_1 or t_2, such that the transitions occur to neighbouring points (In the Euclidean topology of the plane). A plot of trajectories in an arbitrary state-space will not show this convenient form in general, and the individual adaption sets may be partitioned into disjoint sub-sets, subject to the various inclusion relations. Neither do these figures illustrate the possible indeterminacy of the adaption-automaton - trajectories may fork in practice.

2.5 Summary and Conclusions

In this Chapter, a critical discussion of the nature of adaptive behaviour has led to an axiomatic formulation of a theory of this behaviour based on the concept of a "task". Upon this concept has been built a calculus of adaptive behaviour, initially in terms of logical definitions, but finally in terms of semigroup-theoretic definitions based on the derived concept of an adaption-automaton.

The possible modes of adaption are many, and only the most important in the analysis of a system's adaptive behaviour have been singled out for definition. However, any form of adaptive behaviour may now be treated as a trajectory in the state-space of the adaption-automaton, and this is both a convenient and intuitively satisfying representation. In the following Chapter the problem of training is analysed in terms of the trajectories in state-space induced by the sequences of tasks used for training purposes.

CHAPTER 3: TRAINING AS A CONTROL PROBLEM

3.1 Introduction

In this Chapter the approach to learning behaviour developed in Chapter 2 is extended to provide a rigourous foundation for the analysis of training as a control problem, and to enable different modes of training to be defined In this way it is possible to regard training as a problem of control and stability in the state-space of the adaption-automaton of the trainee. Before problems can be 'solved' in any sense, however, it is necessary to have some information about the adaption-automaton.

It is not reasonable to suppose that the adaption automaton is completely known in advance, and neither is it reasonable (because the problem would be insoluble) to suppose that no information is available. Hence, the latter part of this Chapter is concerned with factors influencing the structure of the adaption automaton, and with the minimal forms of information about its structure which make the training control problem soluble.

3.2 Training as Control of the Adaption-Automaton

The formal analysis of learning behaviour in Chapter 2 (together with the results of Appendix 3) show that it is possible to associate with any learning system a structure, an 'adaption-automaton', which is cybernetically equivalent in its behaviour to the adaptive behaviour of the learning system. In particular it has been shown (Section 2.4.3) that any learning sequence of behaviour involving an interaction between the learning system and its environment may be represented as a trajectory in the state space of the adaption automaton. By considering the effects of one form of interaction, or 'task sequence', upon the adaptivity of the learning system to particular 'tasks' it is possible to formalize the concept of training and associated concepts such as, 'negative and positive transfer'.

Consider again Figure 2-3 which shows the regions of satisfactory interaction, W(t), adaption, A(t), and potential adaption, P(t), within the state space of the adaption automaton for a particular learning system. If the objective of training is satisfactory performance of the task, t, then it is required that the state of the adaption automaton should finally lie within the region A(t). However, it is adequate for training purposes that the state of the adaption automaton should lie within the region P(t), since trajectories generated by the task, t, from within this region always eventually enter, and remain in, A(t); that is, the learning system will itself adapt to the required task when the state of its adaption-automaton is initially within P(t).

Thus the training control problem only becomes non-trivial when the initial state of the adaption-automaton is outside the region P(t), and the objective of any training strategy must be to bring the state of the adaption-automaton within this region. When the initial state of a controller's adaption-automaton is outside the region of potential adaption to a task, then successful learning will not take place if it is given that task alone. Given some other sequence of tasks, however, the controller will adapt to them, and may, in so doing, become potentially adaptive to the original task - the sequence of tasks may be said to have trained it for the original task.

In Figure 3-1, for example, the point A is outside the region of potential adaption to the task, t, and repetition of the task does not lead to stable satisfactory performance. The sequence of tasks, u, however, gives rise to a trajectory in the state-space which terminates at A_1, within the region of potential adaption - hence, from A, a sequence of tasks of the form, ut, causes the controller to become adapted to t, whereas a sequence of the form, t^n, does not. If the training sequence, u, consisted of another task, t', repeated many times, then it would be said that there had been transfer of training from t' to t.

The training sequence, u, will not necessarily be suitable for all initial conditions of the controller, and, for example, the trajectory induced by u from the point B in Figure 3-1 terminates at B_1 which is still outside the region of potential adaption. It is possible that the training sequence, u, may not only be ineffective in this way, but may also be detrimental in certain circumstances - for example, the point C is within the region of potential adaption, and yet the trajectory induced by u from C terminates at C_1 which is outside that region. In this case, if $u = (t')^n$, then it would be said that there had been negative transfer from t' to t.

Even when u is ineffective as a training sequence it may be possible to find an alternative sequence with the desired effect - for example, in Figure 3-1 the sequence, v, induces a trajectory from the point, B, which does terminate in the region of potential adaption. In the case of point C, it is clear that no training sequence is required, and the task, t, itself induces a trajectory within the region of potential

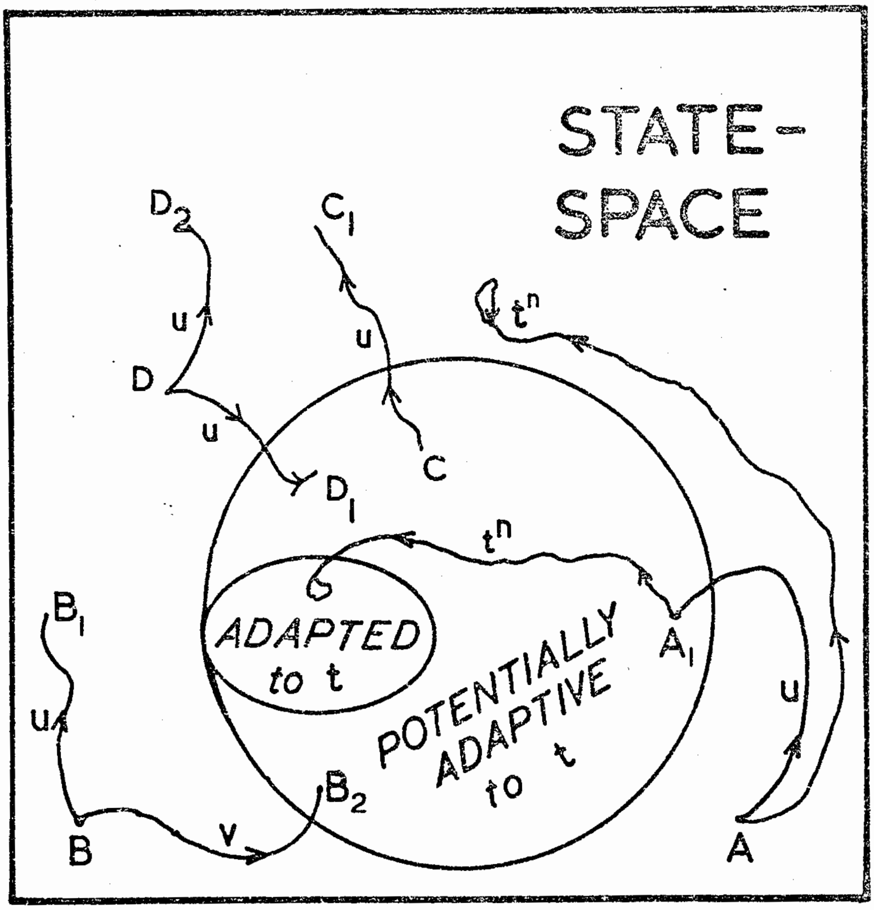


Figure 3-1 Effects of Open-Loop Training Sequences

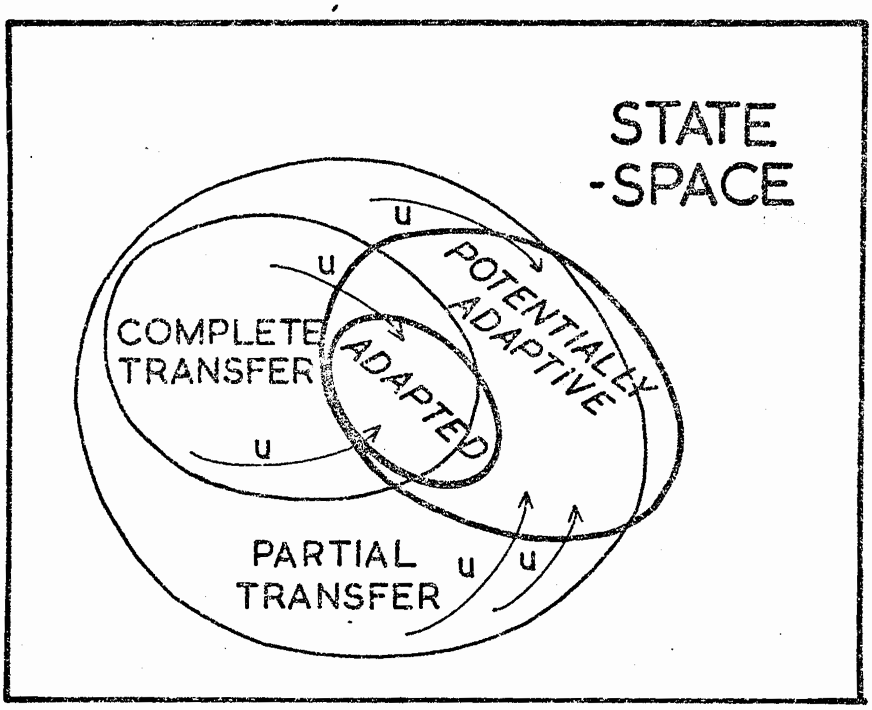


Figure 3-2 Open-Loop Training Spaces

adaption. Thus, even when the controller is not potentially adaptive to its required task, it may be possible to cause it to become so by giving it an appropriate training sequence. To chose between training sequences, however, and to decide whether one is required, it is necessary to have some information about the initial state of the adaption-automaton.

Hence, training may be analysed as the control problem:- given an adaption-automaton structure, representing the possible adaptive behaviour of the trainee, generate a sequence of tasks to serve as an input to the automaton, such that its state terminates in a region where the controller is potentially adaptive to the required task. In general, the initial state of the automaton will be unknown, and information about the past behaviour of the trainee must be used in order to determine it. Since the adaption-automaton may be indeterminate in its state-transitions, an open-loop training sequence will not necessarily be adequate, and continuous feedback from observations of the behaviour and performance of the trainee may be necessary. In the following sections, different strategies for training are distinguished according to these considerations.

3.2.1 Fixed Training

It is convenient to distinguish one strategy for 'training' in which no attempt is made to solve the control problem described in the last section. In Section 2.4, it is stated that 'the fundamental situation with which a controller is expected to cope is to be coupled to a fixed environment and learn to control it satisfactorily'. This is, in a weak sense, a training situation, in that the trainee is given an opportunity to learn, and the sequence of tasks corresponding to it, a repetition of a single task, will be called a fixed training sequence for the environment corresponding to that task.

The success of fixed training depends on the trainee being at least potentially adaptive to the task he has to perform, or, more strongly, compatibly adaptive to the set of tasks he may have to perform. The selection of tasks by the trainer involves no observation of the condition of the trainee's adaption-automaton, and he is given only those tasks which he is required to perform satisfactorily.

3.2.2 Open-Loop Training

In open-loop training, the trainer still does not observe the state of the adaption-automaton of the trainee, but prepares him for adaption to the main task by giving an initial sequence of auxiliary, or training, tasks for which the trainee is not necessarily required to be satisfactory. For example, the sequence u from point A of Figure 3-1, or the sequence v from point B, both induce trajectories which bring the state of the adaption-automaton within the region of potential adaption to the task, t. Fixed training on the task t alone is inadequate if the initial condition of the trainee corresponds to A or B, but the initial training sequences, u or v, respectively, enable adaption to take place, given further fixed training.

In open-loop training, however, there is no information available as to the state of the adaption-automaton, and, if u were chosen to be the training sequence, then it would be given under all conditions. From Figure 3-1, it is clear that u would not be an effective training sequence from B or C, and it becomes important to determine those regions of state-space in which u has the desired effect. In Figure 3-2 are delimited those states from which training with an initial sequence, u, causes the trainee to become adapted, or potentially adaptive to the task, t. If u were itself a single task, t', repeated, then the first region might be called one of 'complete transfer' from t' to t, and the second region one of 'partial transfer'. It is interesting to note that neither the region where the controller is potentially adaptive, nor that where it is adapted, need be contained in the region of partial transfer - the region of potential adaptivity, but no transfer from u, may be called one of 'negative transfer'. Clearly, the utility of u as a training sequence depends on maximizing the region of partial transfer, and minimizing that of negative transfer.

3.2.3 Conditional Adaption Sets

The behaviour of a controller relative to training by a given task sequence, and the various phenomena of transfer, may be defined by sets of the form shown in Figure 3-2, and set-theoretic definitions of trainability may be given similar to those for adaptivity. There is obviously a far greater variety of possible behaviour in training, where adaption is conditional upon previous learning in a training sequence which may itself have definable structure, and many phenomena are best analysed directly in terms of trajectories in state-space. A few conditional adaption sets are of sufficient general importance to be formally defined.

Let $$A(t:u) = [s: \sigma(s,u) \subset A(t)] \qquad [3.1]$$

that is, A(t:u) is the set of states in which the training sequence, u, causes the controller to become adapted to the task, t. This is the set forming the region of 'complete transfer' in Figure 3-2, and the trainee may be said to be completely open-loop trainable by the task sequence, u, for the task, t, when the state of his adaption-automaton is in A(t:u).

Let $$P(T:u) = [s: \sigma(s,u) \subset P(T)] \qquad [3.2]$$

that is, P(T:u) is the set of states in which the training sequence, u, causes the controller to become potentially adaptive to the set of tasks, T. A trainee whose adaption-automaton is in one of these states may not adapt to a task, $t \in T$, when given the fixed training sequence, t^n, but will do so when given the open-loop training sequence, ut^n; a trainee in such a condition may be said to be potentially open-loop trainable by the task sequence, u, for the set of tasks, T.

Similar conditional-adaption sets may be defined for compatibly open-loop trainable -

$$C(t:u) = [s: \sigma(s,u) \subset C(T)] \qquad [3.3]$$

and for jointly open-loop trainable -

$$J(T:u) = [s': \sigma(s,u) \subset J(T)] \qquad [3.4]$$

Apart from the greater variety of the conditional adaption sets themselves, there is also a far wider range of inclusion relations possible, and these will not be described in detail. The interesting inclusion relations are not those logically entailed, as are [2.22] through [2.29], but those which are a function of the controller and tasks. For example, the region of negative transfer has already been defined as those states in which the controller is potentially adaptive to t, but not potentially open-loop trainable by u. For u to induce no negative transfer, we must have:-

$$P(t:u) \supset P(t) \qquad [3.5]$$

Another region of interest is that part of A(t) which is outside A(t:u) - in this region the controller is adapted to t, and this adaption is destroyed by giving it u. However, provided -

$$P(t:u) \supset A(t) \qquad [3.6]$$

the controller is able to re-adapt to t, and u has acted only as a transient distrubance to its performance. In this event we might say that the learning of the task, t, was robust to the influence of the sequence u.

Clearly consideration of the robustness of learning leads to an even more complex structure of adaption sets - A(t) may be split into disjoint parts, $A_1(t)$, $A_2(t)$, with corresponding potential adaption sets, $P_1(t)$, $P_2(t)$, such that it shows robustness to u in $A_1(t)$ but not in $A_2(t)$, and we may consider the problem of training a controller to, "adapt to t robustly with respect to u". Thus, the conditional adaption sets provide not only the means for evaluating various open-loop training sequences, but also a calculus for defining the effects of performing one task, or task sequence, when the controller is in various stages of adaption to another; the conditional adaption sets enable the stability of adaption to be evaluated.

3.2.4 Feedback Training

An open-loop training sequence would be chosen to make the conditional-adaption sets as large as possible, and clearly the optimum sequence would be such that at least P(t:u) includes all possible initial states of the adaption-automaton. If this is not possible, some general restrictions, such as $P(t:u) \supset P(t)$, may be applied to ensure that training does not destroy adaptivity which is already present. However, it may not be possible to find a single training sequence which has all the desired properties, and hence it may be necessary to apply different training sequences dependent on the condition of the trainee.

So far, the training sequences have been taken as fixed sequences of tasks, but, since the adaption-automaton may be indeterminate in its behaviour, it is possible that an effective training sequence cannot be determined in advance even when the initial state of the automaton is known. For example, from D in Figure 3-1, the training sequence, u, might induce a trajectory which terminates either at D_1 within the region of potential adaption, or at D_2 which is outside the region; a further examination of the condition of the controller after having been given u is necessary to determine whether the training has been successful.

Even if the adaption-automaton is determinate, it's initial state will generally be unknown, and observations of the controller's interaction with some environment will have to be made in order to determine the state of the automaton. The task sequence corresponding to this initial interaction clearly need not be a training sequence, and might rather be a testing sequence designed solely to obtain information about the controller prior to training. Such a testing sequence, or 'probe', differs from a training sequence only in that it is designed for rapid observation, rather than to change the state of the adaption-automaton appreciably, and testing may be analysed within the same theoretical, and conceptual, framework as training.

Both when the adaption-automaton is indeterminate, and when it is to be determined by observation, the training sequence is contingent upon feedback from observations of the interaction between trainee and environment, and a system for implementing this feedback will be called a feedback trainer. For the feedback trainer, training has itself become a control problem: there is an 'environment', physically the controller, conceptually its adaption-automaton, whose inputs are tasks and one of whose outputs is the satisfactoriness of the previous interaction; the control problem is to take the automaton from an initial state in which the controller (trainee) is not adapted to the task, to a final state in which it has become adapted to the task (or potentially, compatibly, or jointly, adapted to a set of tasks); the performance measure for this control problem may be based on the number of tasks given before the controller becomes adapted, or it may be a more complex cost function based on the cost of giving irrelevant tasks, and so on. In the following section the "solution" of the feedback training control problem is analysed in more detail.

3.3 The Nature of the Training Control Problem

The derivation of open-loop training sequences and feedback training algorithms, given the structure of the adaption-automaton, is amenable to 'solution' by some of the techniques of modern control theory, such as 'dynamic programming' (Bellman 1957). However, it is totally unrealistic to consider the application of such optimal control algorithms to training in any real situation since the adaption-automaton will never be known in sufficient detail. It is noted in Appendix 3 (A3.3.6) that the derivation of structure from behaviour as there described is a solution to the problem of complete induction. In practice, however, due to the irreversibility of most learning systems it is not possible to collect descriptions of all possible behaviour and only partial information is available. Essentially the adaption-automaton will be only partially known, an approximate model of the adaptive behaviour.

Thus the training control problem is comparable with the majority of real control problems where there is only partial information about the plant and controllers are designed to perform reasonably well with a range of possible plants rather than optimally with a particular plant. In these circumstances feedback control is essential (because the state-transitions of the partially-known plant cannot be completely predicted in advance), and the prime initial problem is to ensure stability in control - that is, that the controller should bring the plant to (global stability), and maintain it in (local stability), a prescribed region of its state-space. In the following sections some weak conditions on the adaption-automaton will be established which ensure the existence of globally stable feedback trainers.

3.3.1 First Training Theorem

A necessary requirement for training to be possible is that there should be available training sequences to transfer the state of the adaption-automaton from any state it might reach during training into at least the potentially-adaptive region. Since an adaption-automaton is generally indeterminate it is also necessary to ensure that these sequences do actually cause the required transition, if not always at least 'frequently' (in the mathematical sense). Given these postulates it is at least feasible to cause the learning system to adapt through training.

However, the training algorithm itself is not trivial given that the only feedback available is the performance of the learning system. A suitable (necessarily weak) training strategy would be to give the required task, t, only once the state of the adaption-automaton is within the potentially adaptive region, P(t), and to chose a training sequence at random if the state is outside P(t). The random choice ensures that there is a finite probability of entering P(t) and, since there is zero probability of leaving P(t), entails convergence of the state to within A(t). Unfortunately, whether the state of the automaton is within P(t) is not known since its performance may be unsatisfactory within P(t) and it may be satisfactory outside it (within W(t) - P(t)).

A suitable training procedure given only performance feedback is to give the task t until the performance is unsatisfactory, and then to select either t or one of the possible training sequences at random, then give the task t again until the performance is unsatisfactory, and so on. Clearly if t is given outside P(t) but within W(t) then, by the definitions of these sets (Section 2.4.2), the performance must eventually become unsatisfactory and some training sequence will be given with a finite probability of causing the state to enter the region P(t). Equally if the state is within P(t) but outside W(t) there is a finite probability of t being given long enough for the state to enter A(t). Convergence under this 'training strategy' is clearly a weak result, but the postulates are so weak that no stronger result can be expected - the weakest possible conditions that will ensure the existence of a convergent feedback training algorithm, and hence a globally stable feedback trainer, are an important starting point from which to discuss stronger results based on stronger postulates.

For purposes of the theorems on training, the adaption-automaton will be taken to be finite-state. This is unnecessarily restrictive but makes the proof more transparent and meaningful in practical terms.

Theorem 3-1 Given a finite-state adaption-automaton and associated sets as defined in Section 2.4.1, a specified task $\tau \in T$, and a set of task sequences, $V \subset U(T)$, satisfying the conditions specified below, then there is a feedback training strategy (based on the performance feedback of equation [4.3] only) such that the probability of the output of the adaption-automaton being satisfactory tends to unity.

Conditions on task sequences

If Σ is the state semigroup (Appendix 3 - Section A3.4) generated by the following constraints on the adaption-automaton :-

a) initial state belongs to some subset $S_o \subset S$

b) the sequence of tasks input to the automaton is constructed of segments of the form:-

(i) initial segment is τ

(ii) if the output of the automaton was p^+ at the end of the previous segment then next segment is τ

(iii) otherwise next segment is $v\tau$, where v is any task sequence such that $v \in V \cup \tau$

then -

$\forall\, s \in \Sigma \cap S,\ \exists\, v \in V : \pi(s,v) \in P(\tau)$ frequently where 'frequently' implies that the sequence of events in which the state s occurs and the input sequence v is given either terminates after a finite number of occurrences, or after any occurence of the pair (s,v) there always exists some future occurrence of the condition $\pi(s,v)\varepsilon\ P(\tau)$. This ensures that training is always possible taking into account the possible indeterminacy of the adaption automaton.

Proof The proof is constructive - consider the training strategy defined by the use of the task segments described in condition (b), subject to the constraint that when a sequence is selected from $V \cup \tau$ the selection is probabilistic with uniform probability of selecting any possible sequence. It is sufficient to show that the state set, $A(\tau)$, is the only trapping set and that there is a finite probability of reaching a state within it at any point in the state trajectory.

$A(\tau)$ is trapping since by equation [2.16] for any state $s\ \varepsilon\ A(\tau)$, $\sigma(s,\tau)\ \varepsilon\ W(\tau)$ so that from equation [2.15] $\pi(s,\ \tau)=p^+$ and hence from condition (b)(ii) the next task will be and from equation [2.16] the state will remain within $A(\tau)$.

No other state sets outside $A(\tau)$ can be trapping since for states outside $W(\tau)$ there is a finite probability that a sequence $v\ \varepsilon\ V$ will be used for input that will take the state within $P(\tau)$ and that a sequence containing only τ will be used thereafter taking the state into $A(\tau)$. Whereas for states within $W(\tau) - A(\tau)$, a sequence consisting only of τ cannot be maintained since the state must eventually leave $W(\tau)$ (by definition of $A(\tau)$) and there is then a finite probability of not selecting τ.

3.3.2 Second Training Theorem

The first training theorem is clearly a weak result. However, it is worth noting that the training sequence had to be carefully constructed to avoid the problems caused by $W(\tau)$ and $P(\tau)$ not coinciding. It is also of interest that an essential random element was necessarily introduced into the training strategy to allow for the lack of knowledge of the 'correct' training sequence for a given state. The training strategy is simple and does not involve identification of the state of the adaption-automaton or detailed knowledge of its structure.

Stronger results on training can only result from stronger constraints, that is, greater knowledge of the structure of the adaption automaton.

There are clearly infinitely many possible sets of constraints to ensure the existence of convergent training strategies with additional desirable properites such as rapid convergence. Only one further case will be considered since it corresponds to the type of training problem considered in Chapters 4, 5 and 6. The assumption will be made that a sequence of tasks is known such that performance of one task leads to improved performance of the next task in the sequence. The training strategy is then to give the tasks in either forward or reverse sequence according to the performance.

Theorem 3.2 Given a finite-state adaption-automaton and associated sets as defined in Section 2.4.1, a specified task $\tau \in T$, and a finite sequence of tasks -

$$t_o , t_1 , t_2 , t_3 , \ldots , t_N$$

satisfying the following conditions -

(i) $t_N = \tau$

(ii) $\forall n : 0 < n \leq N , A(t_{n-1}) \subset P(t_n)$

(iii) $P(t_o) \supset \Sigma \cap S$

where Σ is the state semigroup generated by applying any sequence of t_o through t_N from an initial set of states $S_o \subset S$, then there is a feedback training strategy (based on the performance feedback of equation [4.3] only) such that the probability of the output of the adaption-automaton being satisfactory tends to unity.

Proof The proof is again constructive - consider the strategy in which -

(a) the initial task given is t_o

(b) if the previous task was t_n and the output of the automaton was p^+, the next task is chosen from t_n and t_{n+1} (t_N if t_{n+1} does not exist) with finite probability of chosing each.

(c) if the previous task was t_n and the output of the automaton was p^- the next task is chosen from t_n and t_{n-1} (t_o if t_{n-1} does not exist) with finite probability of chosing each.

As shown in the proof of the first training theorem, the condition in which the task τ is being given when the state of the automaton is within $A(\tau)$ is trapping, since the output will always be p^+ and by condition (b) above no change will be made in the task given.

Suppose that the task t_{n-1} is being given and the state of the automaton is within $P(t_{n-1})$, then there is a finite probability that t_n will continue to be given until the state comes within $A(t_{n-1})$ and that the task will then be changed to t_n. Since, by condition (ii) of the theorem, $A(t_{n-1}) \subset P(t_n)$, there is thus a finite probability of changing from the condition in which t_{n-1} is being given within $P(t_{n-1})$ to the state in which t_n is being given within $P(t_n)$. Hence there is a finite probability of reaching the condition in which $t_N = \tau$ is being given when the state is within $A(\tau)$.

Suppose that t_n is being given and the state of the automaton is outside $P(t_n)$, then, by the definition of $A(t_n)$, there is a finite probability of t_n continuing to be given until the state is outside $W(t_n)$ and hence the output of the automaton is p^-. Hence by condition (c) above, there is a finite probability of the task being changed to t_{n-1}. Then either the state will be within $P(t_{n-1})$ in which case the preceeding paragraph applies, or if not the argument of the present paragraph applies. Thus eventually a condition must arise in which the state of the automaton is within the potentially adaptive region for the task being given, or the task must come to be t_o but by condition (iii), $P(t_o) \supset \Sigma \cap S$, and the previous paragraph applies.

Thus the only trapping condition is τ being given within $A(\tau)$ with output always satisfactory, and there is a finite probability of reaching this from any other state.

3.3.3 Extension to a Lattice of Tasks

An immediate corollary to Theorem [3-2] is to consider a semi-lattice (Clifford and Preston (1961) p.24) of tasks, L, ordered by a relationship, $<$, such that τ is an upper bound, condition (ii) of Theorem [3-2] is satisfied by any pair of directly connected tasks in the lattice, and there is a set of tasks, L_o, each of which satisfies condition (iii). That is if:-

(i) $\forall t \in L \; , \; t < \tau$

(ii) $\forall t,t' \in L : t < t'$, but there is no $t^* \in L :$
$t < t^* < t', \; A(t) \subset P(t')$

(iii) $\forall t \in L_o \subset L, \; P(t) \supset S$

then it is clear that any directly connected chain of tasks starting within L_o and terminating at τ satisfies the conditions of Theorem [3-2]. Hence the result of Theorem [3-2] can be generalized to strategies based on probabilistic movement along connected chains in a semi-lattice of tasks, with movement towards τ when the output of the automaton is satisfactory, and movement towards L_o when it is not.

This corollary is important in the context of the situation chosen for studies of human learning, and is discussed further in Section 4.4.2 in relationship to the movement of the stability boundary of an operator learning a second-order tracking task. However, the result is also of interest in its abstract form in that it gives more meaning to the concept of relationships between tasks discussed in Section 2.3.10.

It is also of interest to consider the extension in more qualitative terms. One may picture the task given as tending to travel upwards through the lattice as learning progresses and the automaton's output becomes satisfactory for tasks of increasing difficulty, and downwards through the lattice if the automaton's output becomes unsatisfactory and it is suspected that the task is too difficult. If the path length from entering P(t) to entering A(t) is long, then a high probability of changing the task will tend to take the task up to too high a level on the one hand, and down to a level for which the automaton is already adapted on the other. Hence the level of difficulty will tend to oscillate with increasing rapidity as the probability of changes goes towards one. Equally, however, if the probability of change is very low, the transition to a more difficult task will be made later than is necessary but the automaton will be well adapted to the current task and hence within the region of potential adaption for the next - there will be little fluctuation of the task but steady progress towards τ .

It is possible to equate the probability of changing the task with the 'loop-gain' of the feedback training by equating p^+ to +1 units of performance, p^- to -1 units of performance, calling the difficulty of t_n , $\delta = n/N$, and the probability of changing the task, p. Then we have from (a) and (b) of Theorem [3-2] -

expected change in δ = p times current performance.

Hence, the argument of the last paragraph may be interpreted as showing that too high a loop-gain will lead to slow learning through oscillating task difficulty, whereas too low a level will lead to slow but steady learning and increase of task difficulty. This is the type of dynamic behaviour which would be expected of a simple servo system, and it is interesting to note that it is possible to derive it from an abstract and fairly general feedback trainer without considerations of continuity, linearity, etc. It also leads to the expectation, taken up again in Section 4.5, that an actual feedback trainer although clearly highly nonlinear may be reasonably analysed as a simple servo system.

3.4 Epistemological Constraints upon the Adaption Automaton

The problem of identifying the structure of the adaption-automaton of a learning system in complete detail has been circumvented to some extent by the two training theorems which demonstrate that training is possible without detailed knowledge of the adaption-automaton provided certain constraints on it exist. The weakness of the necessary constraints is demonstrated by Theorem 1, whilst Theorem 2 indicates a more practically useful level of constraint.

It is clearly of interest to consider whether there is any way other than by direct observation of the learning system in which the contraints can be inferred. In particular, starting from the assumption that the learning system can solve a certain class of problems and hence has a particular approach to, or hypotheses about, its environment, is it possible to infer probable (weak) constraints upon its adaption automaton. One particular epistemological problem which seems a universal source of difficulty in learning is the interaction between the acquisition of knowledge about environment and the degree of control exerted over it. In recent years this has been analysed system-theoretically as the "dual-control problem".

3.4.1 The 'Dual Control Problem'

In the development of adaptive control theory, one of the early steps in simplifying the overall problem was to split it into two parts - that of determining the characteristics of the system to be controlled, or system identification, and that of determining the best controller, relative to some performance criterion, for a known controlled element, or controller optimization. This separation has enabled powerful theories of identification and optimization to be developed separately from one another. However, in any real adaptive control situation, such a separation is impossible and if, from the total activity, 'identification' and 'optimization' are separated out, then the interaction between them makes nonsense of any theories holding for each individually; for example, even in the control of a simple linear system, Sworder (1966) has shown that the optimum identification technique coupled to the optimal control algorithm does not lead to the optimum controller. In more complex systems, it is clear that there may be a conflict between the two requirements that in order to control a system the controller must have information about its relevant characteristics, and in order to obtain information about system characteristics relevant to control the controller must cause the system to operate in the relevant region, that is, control it.

This conflict has been extensively analysed by Feldbaum (1960, 1960*, 1961, 1961*, 1963) who calls it 'the dual control problem', the dual aspects being 'investigating' and 'directing'. He states, 'in dual control system, there is a conflict between the two sides of the controlling process, the investigational and the directional. An efficient control can only be effected by a well-timed action on the object. A delayed action weakens the control process. But the control can only be effective when the properties of the object are sufficiently well known; one needs, however, more time to become familiar with them'. In his theoretical studies, Feldbaum considers the dual control problem from a statistical decision theoretic point of view under restricted conditions, and derives some algorithms for optimal control relative to an overall performance criterion which combine both identification and optimization.

Both Sworder and Feldbaum are more concerned with optimizing an overall performance criterion using an integrated strategy, rather than determining the effects of combining independent identification and optimization algorithms into a single learning control algorithm. In general, it is clear that overall optimization is not possible, and there will always be a conflict between strategies designed to learn about an environment, and strategies designed to control it; for example, even at the level of the simple maze, initial exploration may show that a path which approaches close to the goal is blocked - at some time later it may become unblocked and offer a very much more rapid route, but the possible advantages to be gained by knowing this must be balanced against the loss in time taken by continuous exploration.

3.4.2 The Sub-Environment Phenomenom

The funamental structure of all forms of adaptive controller is a two-level hierachy in which the lower level implements one of a class of control policies, whilst the upper level selects the class to be implemented, Figure 3-3. This definition emphasizes the relativity of the term 'adaptive' from a structural point of view, since any particular controller may be split up in many ways according to the definition of a class of possible control policies; a similar relativity in the behavioural definition of 'adaptive' has been noted in Section 2.3.11 due to the variety of possible ways in which an interaction may be split up into 'tasks'. The relativity of adaptive control structures emphasizes the progressive nature or our understanding of them - as control science progresses, we may expect the adaptive controllers of one era to become the control policies of the next.

To select a control policy approriate to the environment and its goal in controlling it, the upper level of an adaptive controller requires information about the nature of the environment and its interaction with it. There are two distinct classes of information relevant to the selection of a control policy: the nature of the environment itself, that is, identification of the input/state/output relationship for the environment; and performance measures for various possible control policies operating upon the environment. Figure 3-3 shows both these types of information being fed to the upper level of the adaptive controller, and clearly either source is capable of forming the basis for the selection of a control policy: if the controller is able to identify the environment completely and correctly, and has access to a mapping from environments to optimum policies, then it can implement the best policy available; equally, if the controller is able to establish performance measures for all possible control policies by implementing each in turn, then it can, finally, select the best available.

It is when identification and performance evaluation are not exhaustive, and are combined with the necessity to exert some control over the environment, that difficulties in learning arise. These occur because any given control policy will generate some sub-environment, that is, it will restrict the states and state-transitions of the environment to some sub-set of the total possible behaviour. This effect is analysed in Appendix A3.2.4, where it is noted that the behaviour of the sub-environment

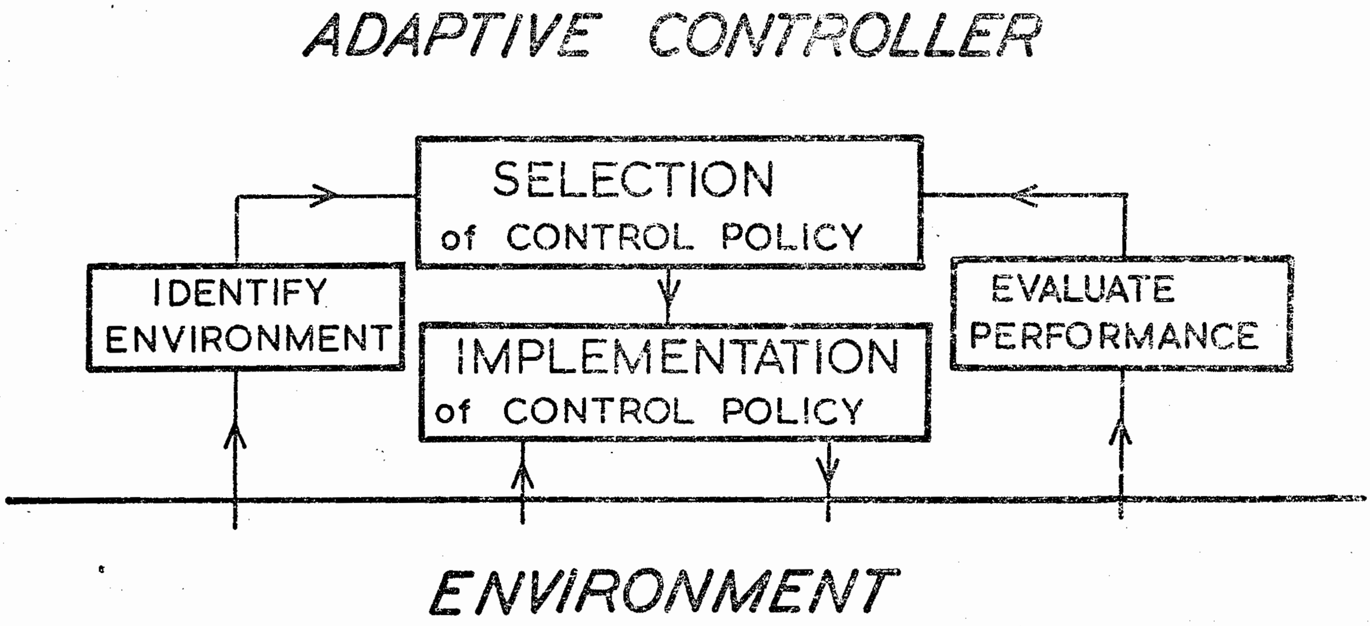


Figure 3-3 Adaptive control structure

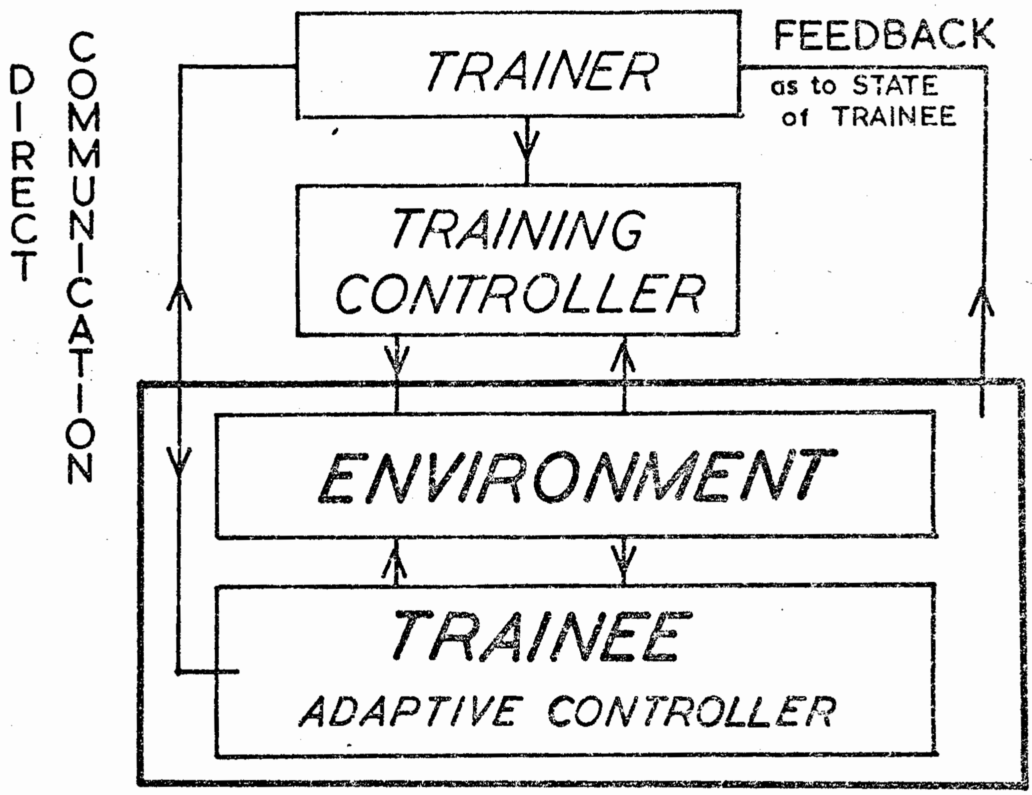


Figure 3-4 Feedback training system

generated by a control policy defines a system structure which varies with the control policy. The sub-environment generated by an initial control policy may be entirely different from that generated by an optimum or satisfactory control policy, and learning in the initial sub-environment may then be irrelevant or even deleterious to performance in the desired sub-environment. Alternatively, and especially if the adaptive controller adopts a deliberate search policy, the initial environment may be so extensive that learning about it would take an unacceptably long time.

The sub-environment phenomenom will have different effects on controllers using identification, and controllers using performance evaluation, in order to vary their control policy. In identification, the measured parameters of the environment will generally not characterize it completely, but only determine some particular properties. In normal circumstances, in the desired or expected sub-environment, many other properties will be correlated with these and may be deduced from them. If the initial control policy generates an abnormal, or unexpected, sub-environment, then the measured parameters may carry no inference about other characteristics of the environment, and the control policy selected through them might be invalid. If this policy also generated an abnormal sub-environment then mal-adaption would continue, so that a normal sub-environment and an optimal control policy would never be attained.

If a controller adapting through performance-evaluation measured the performance of every possible control policy, then the sub-environment phenomenom would cause no difficulty. In practice, however, the environment cannot be assumed to be stationary over periods long enough to do this, and the time taken would be unacceptable. Hence, controllers using performance as a guide to adaption assume that there is some topology on their control policies such that, given the evaluations of a number of different policies, they are able to select a new policy 'near' to the best and 'away' from the worst; that is, they modify their policies incrementally in a 'hill-climbing' mode. Incremental changes in policy will generally cause incremental changes in the sub-environment, and these changes may be such that the total environment fragments into a set of disconnected sub-environments. If the initial sub-environment is not connected to the desired one then the system will be trapped and the controller will never attain optimal control; this phenomenom is shown particularly clearly by the 'evolutionary' controllers of Bremermann referenced in Appendix 1.

The fragmentation of the environment resulting from difficulties in learning may be treated theoretically in terms of the concepts developed in Section A3.4, by examining the 0-minimal ideals of the combined controller/environment in terms of the state-semigroup of the environment. If the 0-minimal ideal in which the system eventually resides contains state-sequences outside the free semigroup generated by the states in which the encironment shows the desired behaviour, which the adaptive controller is attempting to enforce, then the controller is clearly unsuccessful.

3.4.3 Overcoming the Sub-Environment Phenomenom Through Training

The obvious training strategy to alleviate the difficulties caused by the sub-environment phenomenom is to force the initial sub-environment of the controller to be the desired sub-encironment. This is clearly only possible if there are additional inputs to the environment, and, possibly, additional outputs from which to determine how these inputs should be driven. The additional loop between these inputs and outputs necessary to produce the desired sub-environment may be represented, as shown in Figure 3-4, as the addition of a training controller to the environment. The concept of a training controller may be applied in a variety of learning situations - in teaching Euclidean geometry, the training controller might draw in a suitable construction to enable a problem to be solved by elementary procedures; in conversation, the training controller might repeat parts of a statement so that material relevant to the comprehension of later phrases is not missed; in manoeuvring a simulated vehicle, the training controller might apply auxiliary feedback to maintain the overall control loop marginally stable.

Since the trainee in Figure 3-4 is assumed to be adaptive, the training controller need exert less and less control to maintain the desired environment as time passes. If its control policy remains the same, then it is probable that the actual sub-environment will become only a small part of the desired one, which may have just as deleterious effects on the learning of the trainee as the generation of a sub-environment disconnected from the desired one; it is also necessary for the trainee to learn to control the environment independently of the co-operation of the training controller. Hence, there may be a training controller which selects a suitable training controller either as a function of time (open-loop training, Section 3.2.2), or according to information about the state of the trainee (feedback training, Section 3.2.4).

In particular, a sequence of tasks such that performance of a task is a certain region of the state space of the adaption automaton leads to a suitable sub-environment for learning relevant to the next task in the sequence will lead to conditions satisfying the requirements of the second training theorem (Section 3.3). Thus consideration of the relationship between controlling a system and learning about it without reference to any particular learning system may lead to probable constraints upon the adaption automaton relevant learning systems. Such an approach clearly cannot demonstrate that any particular system can actually learn a task, but many lead to optimal conditions for learning should system be capable of doing so.

3.4.4 Training and Communication

It may be noted that the structure of the training system shown in Figure 3-4 is an exact image of the structure of the trainee, regarded as an adaptive controller shown in Figure 3-3. The upper level of the training system adjusts the control policies of the lower level so as to maximize the effectiveness with which the upper level of the trainee varies the control policies of its lower level. Thus, the trainer communicates with the adaptive level of the trainee through a very complex system in which two controllers interact with a common environment. With the human controller, and with recent learning machines, direct verbal communication may be possible between the two higher levels, so that the trainer may short-cut the complex communication channel through the environment and prime the trainee with information relevant to its control problem.

The nature of verbal communication is not sufficiently well understood, especially in its effects on perceptual-motor skills, to enable a thorough analysis of the use of a direct channel of communication between trainer and trainee. In the context of identification and performance evaluation, however, it is clear that it may be possible for the trainer to pass information about the true nature of the environment or the optimality of various control policies to the trainee, and hence eliminate the sources of difficulty in the dual control problem. Discussion of possible forms of instruction for human operators is given in Section 5.1.8, and experiments with learning machines are described in Section 6.3.4.

The advantages of direct communication are great, but the conditions under which it is possible are very stringent - the trainer must not only have the information about the nature of the environment, or the optimality of control policies, available, but must also be able to communicate this as a form assimilable by the trainee. In practice, the trainer's knowledge about the environment will be incomplete, and his channel of communication with the trainee will be imperfect. He will be able to prime the trainee to some extent, and select training controllers in a reasonable way - the optimum training system will combine verbal instruction with feedback training, and the experiments reported in Chapter 5 investigate the interaction between these two techniques.

3.5 From Theory to Practice

This section concludes the description of work on the abstract, axiomatic study of learning and training, and is a natural point at which to review the implications and utility of the theory. The overall logic of the theoretical development is -

(i) Critically analyse the meanings of the terms 'learning', 'adaptive' and their derivatives (Section 2.1).

(ii) Give purely behavioural explicata for the concepts underlying these terms (Section 2.2) - this introduces the fundamental concept of a 'task'.

(iii) Use these as an axiom set for a calculus of adaptive behaviour (Section 2.3) - this introduces 'potential', 'compatible' and 'joint' adaption.

(iv) Derive a mathematical structure in which to express concisely the definitions of adaptive behaviour (Appendix 3) - this introduces the concept of an observable, non-determinate automaton, cybernetically equivalent to a system defined by its behaviour.

(v) Express the previous discussion of adaptive behaviour in terms of the equivalent 'adaption-automaton' (Section 2.4) - this enables learning behaviour to be described as a trajectory in the state-space of the adaption-automaton.

(vi) Analyse training as a probelm of control in the state-space of the adaption-automaton (Section 3.2) - this enables three modes of training to be distinguished, 'fixed', 'open-loop' and 'feedback'.

(vii) Note the fundamental impossibility of identifying the detailed structure of the adaption-automaton in most real systems, and investigate the feasibility of training when only weak overall properties of the automaton are available (Section 3.3) - this leads to two theorems on training, one showing that the very weak conditions necessary for training to be possible are also sufficient to derive an actual trainer, and the other demonstrating the improved results possible if more information is available about the adaption-automaton.

(viii) Consider other sources of information indicating the probable structure of the adaption-automaton apart from observation of the learning behaviour (Section 3.4) - this introduces the 'dual-control' problem and the associated 'sub-environment' phenomenom, and their epistemological influence.

The most important practical outcome of this chain of reasoning is that it gives a unified theoretical foundation to the two extreme approaches to training - the 'stimulus-response' approach on the one hand, typified by most 'programmed-learning' material (MacDonald Ross 1969), and the 'learning-environment' approach on the other, typified by the 'adaptive trainer' (Pask 1960) and certain 'computer-assisted-instruction' programs (Wexler 1970). The 'stimulus-response' approach, in which it is intended that each item given to a student be a function of his specific responses to previous items, may be seen as an attempt at optimum control in the state-space of the adaption-automaton. Its application suffers from the same problems that have beset control engineers attemtping to apply 'optimum control' theory: that far more detailed information about the controlled system (adaption-automaton of trainee) is required than is ever realistically available; and that the control strategy itself is very complex, so that both its design and implementation are major problems.

The 'learning-environment' approach, in which a situation is created and maintained which is expected to be conducive to learning, may be seen as an attempt to take advantage of the expected region of local stability in the state-space of the adaption-automaton (potential adaption in which learning takes place without specific 'training') and extend this to give global stability throughout the state-space. This is the approach taken in the design of linear feedback servomechanisms which are able to stabilize a wide variety of plants without being optimal for any one, but equally without requiring detailed information about the plant parameters. In many cases in control engineering (for example, Fuller 1967) the simple linear servo may be shown to have a performance only neglibly worse than a very much more complex and specific optimal control suitable for only one of the plants to be controlled.

In practice neither one approach nor the other may be taken as the best mode of training - in general, stabilization through a controlled environment, provides the means to extend the trainee's capability to learn without the trainer having to pay specific attention to particular attributes of the trainee. However, it is unreasonable to suppose that it will be possible to extend the capability of all possible trainees (that is, ensure global stability throughout the state-space of all adaption-automata), and for the residual trainees a specific training program designed through observation of their individual behaviour and problems will be necessary - in practice, the observation will show up specific defects of some trainees, e.g. dislexia, lack of spatial ability, etc. What is important in practice is that the approach taken is that appropriate to the training situation - in particular that a 'stimulus-response' approach is not attempted on a broad front, but used rather to overcome specific, well-defined problems.

In the following chapters the generality of discussion is greatly reduced, and a specific feedback training system for a perceptual-motor skill is examined. The training situation exemplifies the conditions necessary for Theorem 3-2 to be applied, and the reasons for this may be traced to a sub-environment phenomenom caused by instability in performing the skilled task. Clearly, with an axiomatic approach to the theory of training, experimental validation is not concerned with whether the results are 'correct', but whether the theory may be applied to any real situations. The experimental results of the following chapters indicate this is so in at least one practical situation, and hence make it more plausible that it is so in others.

CHAPTER 4 A FEEDBACK TRAINER FOR A TRACKING SKILL

4.1 Introduction

The theoretical developments of Chapters 2 and 3 provide a rationale for the application of feedback training techniques to aid the learning of an adaptive system. The theory is, however, to a large extent neutral in its application to real situations - it classifies adaptive behaviour but does not imply that the behaviour will occur in any particular situation. The possibility of maintaining the desired sub-environment is argued to be a suitable basis for feedback training, but it cannot be proved in general that the "training" will give rise to improved learning. Thus, empirical tests of the applicability of the theory to real situations are required, and, in particular, the viability and utility of feedback training techniques are in question. Hudson's (1964) results (Section A4.5.2) suggest that a feedback trainer may be useful, even if his particular form of automated feedback trainer did not prove to be viable, and Kelly (1967) has reported success with a different form of automated system, but has applied it largely to testing and not to training.

Since previous studies had not demonstrated the utility of feedback training, and indeed doubts had been expressed about this (Leonard 1962), the first objective of the experimental studies to be described was to determine a situation in which feedback training would give definitely improved learning compared with alternative techniques. The theoretical analysis predicts that such situations should exist, even with moderately complex adaptive systems, and demonstrating their existence for the human controller is a necessary step in the study of human learning behaviour and its control through training. Thus the questions at issue were whether automated feedback training was in itself viable, for example, in terms of the trainer's stability, and, secondly, whether feedback training is useful in any circumstances, rather than is it applicable to a particular training problem - the null hypothesis was that feedback training of the human operator never gives improved learning, in some sense, over the best open-loop or fixed training.

The detailed design of a training situation in which feedback training might be expected to have definite advantages is discussed in Chapter 5, where the problems in evaluating human learning are analysed and the experiments on training human operators are described. The present chapter is concerned with the viability of a particular form of feedback trainer, its behaviour and stability, and also with its applications to the testing of the ability to perform perceptual-motor skills.

4.2 Choice of Skilled Task

The choice of a skilled task for the experimental studies was largely arbitrary, the only formal constraint being that the skill should be one which could be learnt by a normal operator to a reasonably stable level of performance, without artifacts of fatigue, task-induced stress, and so on. For the experiments on training, it was also necessary that appreciable learning be possible, so that differences between initial and final performances would be measurable and an indication of the relative merits of different training techniques. It was also desirable that learning take place over a reasonable time-period, say thirty minutes to three hours, for experimental convenience, and that initial performance and ease of learning show little spread through the experimental population, so that statistically significant results might be obtained from a reasonably small sample of operators. The satisfaction of these training constraints is discussed in Chapter 5.

Informal influences on the selection of a skill were that it should be related to those studied by other workers, so that the general literature could be drawn upon in interpreting the results, and that the skill should be related to a practical situation where the training techniques might be applied. In their experiments with feedback testing or training, Chernikoff (1962), Hudson (1964), Jex et al (1966) and Kelly (1967) have used compensatory tracking tasks with continuous manual control and visual indication of error, and the psychological literature on compensatory tracking is vast, as is the corresponding control-engineering literature on single input, single output, regulatory controllers.

A compensatory tracking task, which is of great practical importance and involves training to a high level of skill, is the regulation of the attitude of an aircraft through the use of elevator control. The strategy used by a pilot in controlling the attitude, and the effect upon this strategy of changes in the longitudinal dynamics of the aircraft have been extensively investigated by many workers (Appendix 4), largely in the aircraft industry, who have also monitored other variables of interest, such as the pilot's opinion of the simulated craft (Hall 1963). Hence, the regulation of the short-period motion in the longitudinal dynamics of an aircraft was taken as the basis for a model tracking skill.

4.2.1 Dynamics of the Tracking Task

The longitudinal dynamics of an aircraft (Kolk 1961, Blakelock 1965) are those between the elevator control (the pilot's joystick) and the attitude, or pitch, of the aircraft. Because a constant angle of the elevators creates a turning moment to change the attitude, there is a pure integration between the joystick position and the attitude. Because of the interplay of aerodynamic forces, there are also two second-order, oscillatory transfer-functions in cascade between the joystick and pitch. One of these, causing the 'phugoid' motion of the aircraft, is of very long period (several minutes), and is generally neglected in studies of the pilot's control policy. The other, short-period, dynamics have a natural frequency in the region of 0.3 Hz, and are most relevant to the 'feel' of the longitudinal dynamics.

In his studies, described in Section A4.2.1, Hall (1963) took the longitudinal dynamics to have the form given in Equation A4.1. In the present studies, the lead term in the numerator, (1 + 0.6s), has been omitted and the overall transfer function has been taken to be:-

$$G(s) \quad = \quad L/s(w_n^2 + 2kw_ns + s^2) \qquad [4.1]$$

The term w_n^2 has been multiplied through the equation (so that the new value of 'L' is w_n^2 times the previous value), because k and w_n are taken as variables to be changed in training. If w_n is varied using Hall's form of the equation, then the concomitant change in gain gives a system which feels reasonable at one value of w_n, sluggish at lower values, and over-sensitive at higher values. It was found in initial informal experiments that variation of w_n over a wide range, for a constant value of L, gave an acceptable system to control using the dynamics of Equation 4.1. The range of variation of the two parameters was, setting $F_n = w_n/2\pi$:

$$0 \leq k \leq 1 \qquad [4.2]$$

$$0 \leq F_n \leq 0.8 \text{ Hz} \qquad [4.3]$$

A block diagram of the compensatory tracking task with these dynamics is shown in Figure 4-1: three integrators in cascade prove the overall, third-order dynamics between the manual control and the visual display; negative feedback from the output of the second integrator to the inputs of both first and second integrators sets up the oscillatory second-order transfer-function, followed by a pure integration; a disturbing signal is fed in at the same point as the operator's input. Oscillatory dynamics have been little studied in

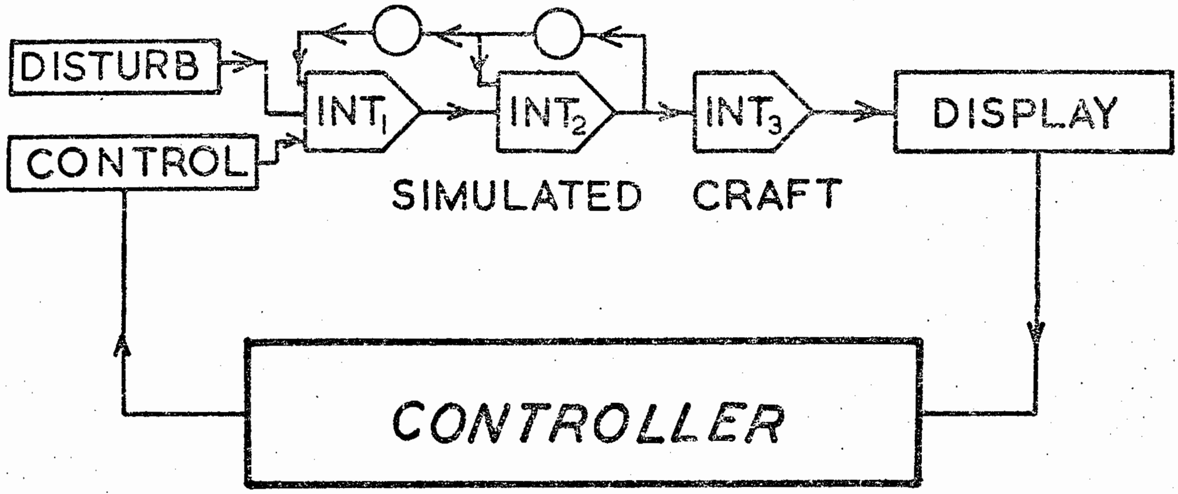


Figure 4-1 Third-Order Compensatory Tracking Task

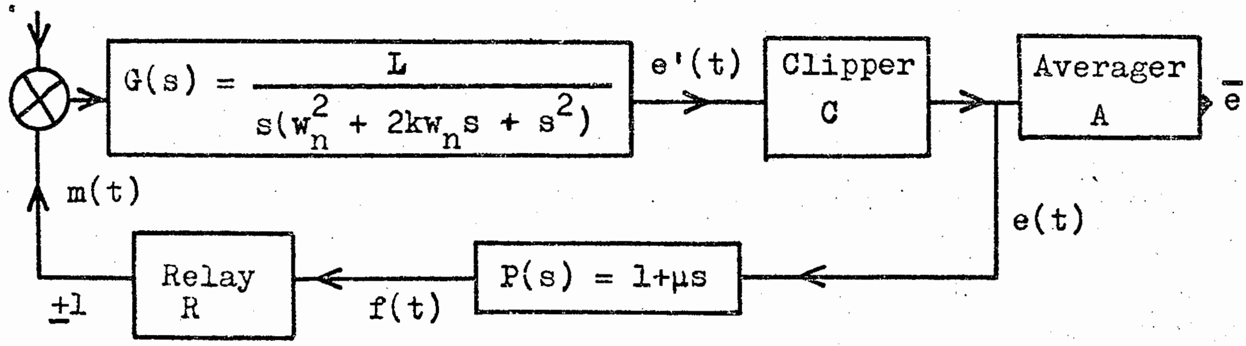


Figure 4-2 Relay Controller for Tracking Task

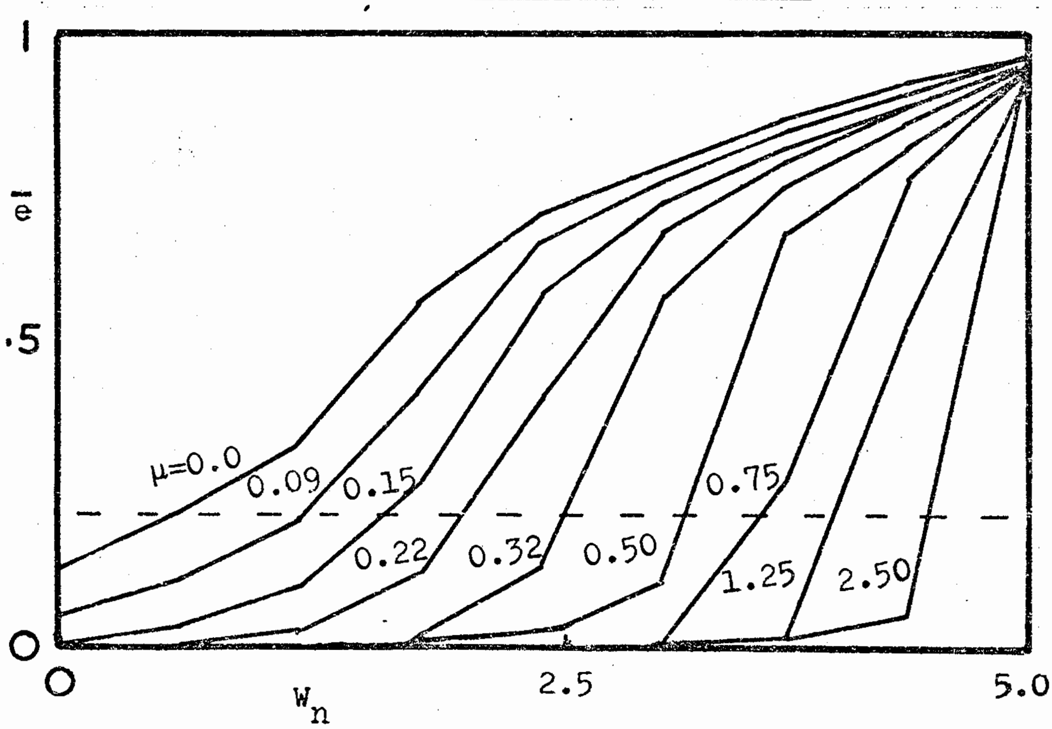


Figure 4-3 Mean Error as a Function of Natural Frequency

psychological experiments on the acquisition of skill, but, apart from their practical importance, they have characteristics, in particular a non-monotonic response, which create distinct and interesting difficulties in learning to compensate for them. Furthermore, the cascaded exponential lags, so often used in psychological studies, are a special case of this transfer-function with the damping-ratio, k, set to unity. Thus, the dynamics offer more scope for study than the lags utilized by Chernikoff (1962) and Kelley (1967); Hudson (1964) included oscillatory dynamics in his study.

4.2.2 Parameters of Difficulty for the Tracking Task

The implementation of a feedback trainer for the particular third-order tracking task selected clearly involves using some information about the learning behaviour of the human operator for this form of tracking skill. In previous chapters it has been emphasized that this information is available from diverse sources:- from observation of actual adaptive behaviour to determine the adaption-automaton; from a knowledge of the purpose of the controller and the desired sub-environment for learning; and from a knowledge of the behaviour of other controllers with similar objectives to those of the human operator. In practice, all these sources contribute some partial information about suitable structures for a feedback trainer, and the actual trainer is a synthesis from all three.

The advantage of a tracking task in designing a trainer is that the performance of the trainee may be continuously assessed in terms of the error which is displayed to him. Hence, it is possible to consider the use of a simple performance-feedback trainer, whose sole information about the trainee is derived from measurements of his performance. The justification of performance-feedback training in terms of maintaining the desired sub-environment is discussed in Section 4.4. In this section the dependence of performance on parameters of the tracking task is analysed, and it is convenient to suppose, informally, that these parameters affect the difficulty of the task for the operator.

From Figure A4-2(ii) which is Hall's plot of the contours of constant mean tracking error in the natural-frequency/damping-ratio plane, it may be seen that the operator's performance decreases monotonically with decrease in natural-frequency. Figure A4-2(iv), showing the pilot's opinion of the tracking task, also indicates that the difficulty increases with decreasing k and F_n (or w_n), for F_n less than about 0.8 Hz. In fact, decreasing k causes the system to become more oscillatory, and decreasing w_n leads to longer time lags, so that these results are are to be expected; the disparity between the contours of Figures A4-2 (ii) and (iv) is probably because too high a frequency of oscillation is visually annoying when the system is under-damped.

A third parameter affecting the difficulty of the tracking task is the amplitude and nature of the disturbance at the input. Clearly, the greater the amplitude of disturbance the greater will be the error, and generally the greater the high-frequency content of the disturbance the greater the error. In the initial informal experiments many forms of disturbance were used, including Gaussian noise, sine waves and square waves of different periods. Because the disturbance passes through the complete controlled system, which acts as a low-pass filter, much of its high-frequency content is smoothed out and there was little difference in effect between these various forms of disturbance. Throughout the formal experiments a square-wave of twenty seconds period was used as a disturbance - a repetetive waveform was used to enhance the possibilities for learning in the training experiments, and it was filtered through the controlled element to avoid the visual fatigue associated with a rapidly-changing display.

Although it may be shown that the error due to a disturbance in a linear control system is proportional to the amplitude of the disturbance, the exact effects of variations in the natural-frequency and damping-ratio of the controlled element are less readily determined. In control engineering, it is known that a regulator for the transfer-function described has a more 'difficult' task as F_n and k are decreased, in the sense that the acceleration and velocity terms in the controller transfer function have to increase in relative magnitude to maintain the same stability margin. In the following section the quantitative relationship between the tracking task parameters and controller performance is analysed in detail for a simple relay controller, in order to provide data for a theoretical analysis of the behaviour and stability of a performance-feedback trainer.

4.3 Performance of a Relay Controller for the Tracking Task

In choosing a form of automatic controller for a theoretical analysis of the tracking task, and for later trials of the feedback trainer, a controller with similar ouput characteristics to the human operator was desired. In Appendix 4 and from Figure A4-3 in particular, it is shown that the human operator's output for high-order controlled elements is discontinuous and far more similar to that of a switching, or 'relay', controller, than that of a linear servomechanism. In the theoretical and experimental analysis of the feedback trainer also, it was found that the behaviour of the trainer with a linear servo as 'operator' was very different from that with a human operator, whilst that with a relay controller was similar. Hence, the theoretical analysis is based on the behaviour of a relay controller performing the tracking task.

4.3.1 Theoretical Derivation of Mean Error of Relay Controller

A 'bang-bang' or relay controller (Gibson 1963 p.342) is the simplest possible form of regulator for a single-input, single-output system. The controller's output takes one of two values, $\pm$ M say, and the sign of the output is the opposite of that of the error, so that the controller makes the minimum decision necessary to apply some form of negative feedback. In systems containing lag, the simple relay is an inadequate controller and a predictive, or lead, element must be placed before the relay in order to compensate for the lag and cause it to switch before the error passes through zero. A block diagram of a relay controller coupled to the tracking task of Figure 4-1 is shown in Figure 4-2, consisting of plant dynamics, clipping element, lead network, relay and summer: the error, $e(t)$, is fed through the lead network, $P(s)$, to give an output, $f(t)$, which drives the relay element, R; the relay drives a binary signal, $m(t)$, into the controlled element, $G(s)$, and there is added in the disturbance, $a(t)$; since the loop may go unstable and the output of $G(s)$, $e'(t)$, become large, a clipping element C, is shown before the error signal, $e(t)$; an averaging network, A, forms the time-average of $e(t)$, $\bar{e}$, in order to measure the effectiveness of the controller.

The lead network $P(s)$, was taken to have the simple first-order form:-

$$P(s) = 1 + \mu s \tag{4.4}$$

and the relay itself obeys the equation:-

$$m(t) = -\mathrm{sgn}(f(t)) \tag{4.5}$$

If the disturbing input, $a(t)$, is taken to be zero, and the affect of the clipping element, C, on the loop behaviour is neglected (it may clearly be neglected if $\mu=0$), then the behaviour of the control system may be analysed using the describing function technique (Gibson 1963 Section 9.3). A limit signal oscillation is set up in the loop such that the input to the relay has the form:-

$$f(t) = f\, e^{jwt} \tag{4.6}$$

and satisfies the basic equation of describing function analysis:-

$$P(jw)G(jw) = 1/K_{eq} \tag{4.7}$$

where K_{eq} is the equivalent gain of the relay. This is given by:-

$$K_{eq} = 4M/\pi f \tag{4.8}$$

so that the 'gain' of the relay is inversely proportional to the amplitude of the signal at its input.

Substituting in Equation 4.7 from Equation 4.1 and 4.8, we obtain:-

$$4ML(1+j\mu w)/\pi f = -jw(w_n^2 + 2jkww_n - w^2) \tag{4.9}$$

and, separating real and imaginary components:-

$$w^2 = w_n^2 / (1 - 2k\mu w_n) \tag{4.10}$$

$$f = 2ML(1-2k\mu w_n)/\pi k w_n^3 \tag{4.11}$$

Continuing to neglect the clipping element, C, we have that:-

$$e = f / (1 + \mu^2 w^2)^{1/2} \tag{4.12}$$

and, substituting in Equation 8.12 from Equations 8.10 and 8.11:-

$$e = 2ML(1-2k\mu w_n)^{3/2} / (\pi k w_n^3 (1-2k\mu w_n+\mu^2 w_n^2)^{1/2}) \tag{4.13}$$

In obtaining the mean error, $\bar{e}$, from e, it is convenient to take into account the effect of clipping in C, which is assumed to limit the signal in amplitude to $\pm c$, so that:-

$$\bar{e} = \begin{cases} 2e/\pi & 0 \leq e \leq c \\ 2\left[e-(e^2-c^2)^{1/2}+c(\cos^{-1}(c/e))\right]/\pi & e > c \end{cases} \tag{4.14}$$

Equations 4.13 and 4.14 together enable the mean error to be calculated for a range of values of the plant and controller parameters.

The relationship between $\bar{e}$, k, w_n, and μ is not clear from the analytic form of the equations, but certain characteristics may be established from special cases. Firstly, the dependence of $\bar{e}$ on e may be made clearer by writing Equation 4.14 in implicit form:-

$$\text{let} \quad e = c \; (\mathrm{cosec}(\theta)) \qquad e > c \qquad [4.15]$$

$$\text{then} \quad \bar{e} = c \; (1 - 2(\theta - \tan(\theta/2))/\pi) \quad e > c \qquad [4.16]$$

When e is large, θ is small and $e \approx c/\theta$, so that:-

$$\bar{e} \approx c \; (1 - c/\pi e) \qquad e >> c \qquad [4.17]$$

and $\bar{e}$ is asymptotic to c. A table of values relating $\bar{e}/c$ to e/c is given in Table 4-1 - the relationship is linear for $e \leq c$ and then $\bar{e}/c$ increases less rapidly than e/c.

e/c	$\bar{e}$/c	e/c	$\bar{e}$/c
0	0	1·540	0·785
x	0·637x	1·701	0·807
1	0·637	1·914	0·830
1·052	0·663	2·202	0·853
1·122	0·690	2·613	0·877
1·173	0·707	3·152	0·901
1·236	0·724	4·284	0·925
1·315	0·743	6·393	0·950
1·414	0·764	12·745	0·975
1·540	0·785	∞	1

Table 4-1 Input Amplitude and Mean Output for a Limiter

Secondly, for the particular case when there is no velocity feed-back and μ=0, from 4.13 we have:-

$$k w_n^3 = 2ML/\pi e \qquad [4.18]$$

Hence, the lines of constant mean error in the natural-frequency/damping-ratio plane are such that k is inversely proportional to w_n cubed. Also, subject to limiting, the mean error is inversely proportional to w_n^3 cubed also. From Equation 4.13, it is apparent that this cubic relationship dominates even when μ is non-zero, although other terms in k and w_n then appear. In order to establish the relationship between the mean error, $\bar{e}$, and parameters of difficulty, such as w_n, experimental trials of the system shown in Figure 4-2 were carried out, and the results compared with the theoretical results predicted from Equation 4.13 and 4.14; these are reported in the following section.

4.3.2 Experimental Results on the Behaviour of a Relay Controller

The system shown in Figure 4-2 was set up on an analogue computer using chopper-stabilized amplifiers, and an integrator was used to measure the mean error, $\bar{e}$, over a four minute interval. The values of the parameters used in the experiments were:-

Disturbance	-	$a(t)$	=	0
Relay Output	-	M	=	0.0216
Open-loop Gain	-	L	=	1,000
Limiting Level	-	c	=	0.95
Damping Ratio	-	k	=	0.5
Natural Frequency	-	$0 \leq w_n \leq 5$ radians/second		
Velocity Feedback	-	$0 \leq \mu \leq 2.5$		

Table 4-2 shows the experimentally determined values of $\bar{e}$ for various values of μ and w_n, whilst Table 4-3 shows the theoretical values from Equations 4.13 and 4.14.

$w_n =$ \ $\mu =$	0.000	0.087	0.150	0.225	0.325	0.500	0.750	1.250	2.500
0.000	.966	.964	.957	.952	.947	.940	.934	.920	.950
0.625	.926	.901	.879	.883	.857	.815	.759	.542	.059
1.250	.860	.838	.818	.790	.759	.674	.272	.013	.000
1.875	.785	.753	.728	.687	.569	.109	.005	.000	.000
2.500	.706	.665	.584	.401	.134	.015	.000	.000	.000
3.125	.564	.418	.272	.116	.025	.000	.000	.000	.000
3.750	.329	.202	.107	.028	.000	.000	.000	.000	.000
4.375	.213	.114	.043	.006	.000	.000	.000	.000	.000
5.000	.127	.051	.004	.000	.000	.000	.000	.000	.000

Table 4-2 Experimentally Determined Values of Mean Error

$w_n =$ \ $\mu =$	0.000	0.087	0.150	0.225	0.325	0.500	0.750	1.250	2.500
0.000	.950	.950	.950	.950	.950	.950	.950	.950	.950
0.625	.947	.947	.947	.947	.947	.947	.944	.927	.000
1.250	.929	.927	.924	.920	.911	.872	.144	.000	.000
1.875	.880	.866	.848	.811	.686	.042	.000	.000	.000
2.500	.782	.724	.627	.373	.099	.000	.000	.000	.000
3.125	.574	.397	.256	.104	.000	.000	.000	.000	.000
3.750	.332	.207	.110	.022	.000	.000	.000	.000	.000
4.375	.209	.116	.047	.000	.000	.000	.000	.000	.000
5.000	.140	.068	.019	.000	.000	.000	.000	.000	.000

Table 4-3 Theoretically Derived Values of Mean Error

It is apparent from the tables that the agreement between theoretical and experimental values of the mean error is very close for the lower values of $\bar{e}$ at which limiting does not occur ($\bar{e} < 0.6$), and any discrepancy may be ascribed to the experimental inaccurancies in establishing w_n (of order 2.5 per cent). However, for higher values of $\bar{e}$ at which limiting plays a major role, the theoretical values rise more rapidly than the experimental ones. This may be explained in terms of the approximation made by neglecting the effects of limiting on the loop behaviour, and only calculating its effects on the measured mean error. The limiting clearly reduces the error signal circulating in the control loop rather than just the measured signal, and this gives rise to the disparity between experiment and theory. However, in the heavily limited region, the controller has effectively lost control of the loop and the behaviour is not of major importance in the theory of training.

Figure 4-3 shows the experimental values of $\bar{e}$ plotted as a function of w_n for different values of μ, making apparent the variation of performance with difficulty. The curves are sigmoidal limiting at high error amplitudes approximately according to Equation 4.14, and being asymptotic to zero according to the cubic relationship of Equation 4.13. It will be noted that the sensitivity of $\bar{e}$ to changes in w_n, defined by the slopes of the curves, varies as a function w_n, but that maximum slope occurs at much the same value of $\bar{e}$ in each case. This is the major characteristic which makes a simple feedback training system possible, and makes it attractive to use a constant-error criterion for testing a controller.

4.4 Feedback training Strategy

The three parameters of difficulty for the controlled element shown in Figure 4-1, selected for variation in training, are the undamped natural frequency, w_n, the damping-ratio, k, and the amplitude of the disturbance, a. There is an alternative viewpoint from which these parameters may be examined which relates the variation of difficulty to that of a 'training controller', discussed in Chapter 3. If the required control skill is taken to be that of regulating a pure third-order system, consisting of three pure integrators in cascade with the dynamics $- L/s^3$, subject to a disturbing signal, then a suitable training controller might place negative feedback loops around the integrators and inject a disturbance-cancelling input. If the negative feedback loops are such as to feed the velocity of the output back to the acceleration of rate of change of acceleration, then the overall transfer function will be that of the longitudinal 'aircraft' dynamics described in Section 4.2.1. In Figure 4-4, the controlled element of Figure 4-1 has been re-drawn to separate out the training controller from the pure third-order system on which the trainee is ultimately required to operate.

The feedback trainer has to give the naive trainee an easy task and then gradually increase its difficulty as the trainee's skill, as measured by his performance, improves. This was achieved in practice by driving the parameters of the training controller from the output of an integrator, such that one extreme of the output gave the easiest task, whilst the other extreme gave the most difficult task. The modulus of the error at the output of the controlled element was fed to the input of the integrator minus a tolerated level of the mean error, e_o, in such a sense that if the mean error is above tolerance then the difficulty of the tracking task is reduced, whilst if it is below tolerance then the difficulty is increased. The overall effect is clearly such that if there is a stable value of the integrator output then the mean error is equal to the tolerated level. Hence, in Hudson's terms (1964), the absolute difficulty of the tracking task is varied by the training controller to maintain its difficulty for the trainee constant.

There are two alternative interpretations of the training strategy in terms of the results of Chapter 3 which relate the practical trainer to the theoretical studies. These lead to alternative justifications of the training strategy which are discussed in the following sections.

4.4.1 The Training Strategy as Maintenance of a Desired Sub-Environment

The third-order controlled element of Figure 4-4 is a linear system with three state-variables, the position, velocity and acceleration of the output. The desired sub-environment of a regulatory controller is a region about zero in this state-space. Provided this region does not impinge on the boundaries of the state-space (the position, velocity and acceleration are each bounded in magnitude in any physical realization of the transfer-function of the controlled element), the system will behave within it in a linear manner. The desired sub-environment will be of finite size because of the disturbance which, even if perfectly predicted, cannot be cancelled instantaneously. The maximum value of the disturbance in all the experiments was in fact chosen so that a skilled operator could maintain the system in its linear region.

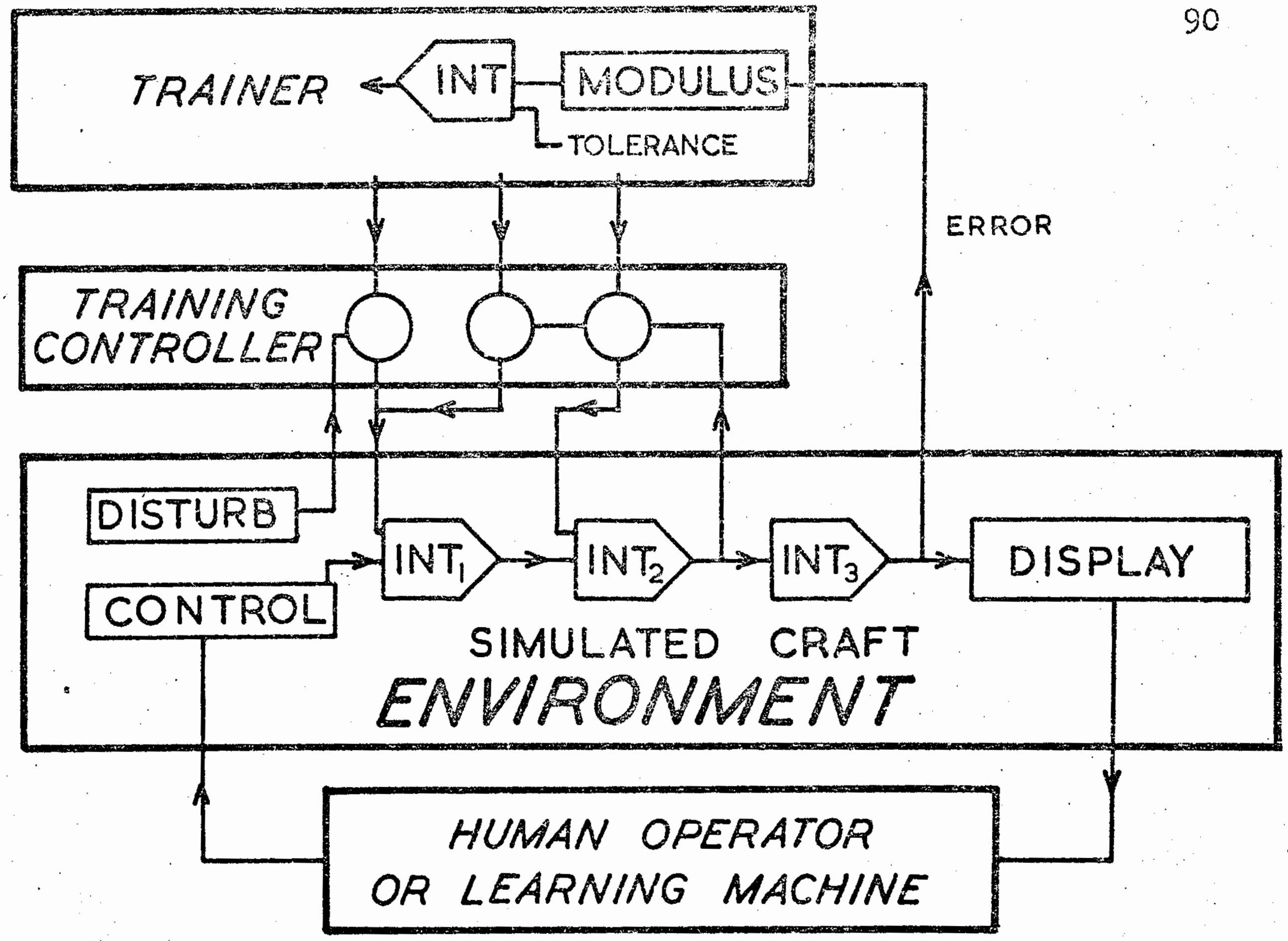


Figure 4-4 Feedback Trainer for a Tracking Task

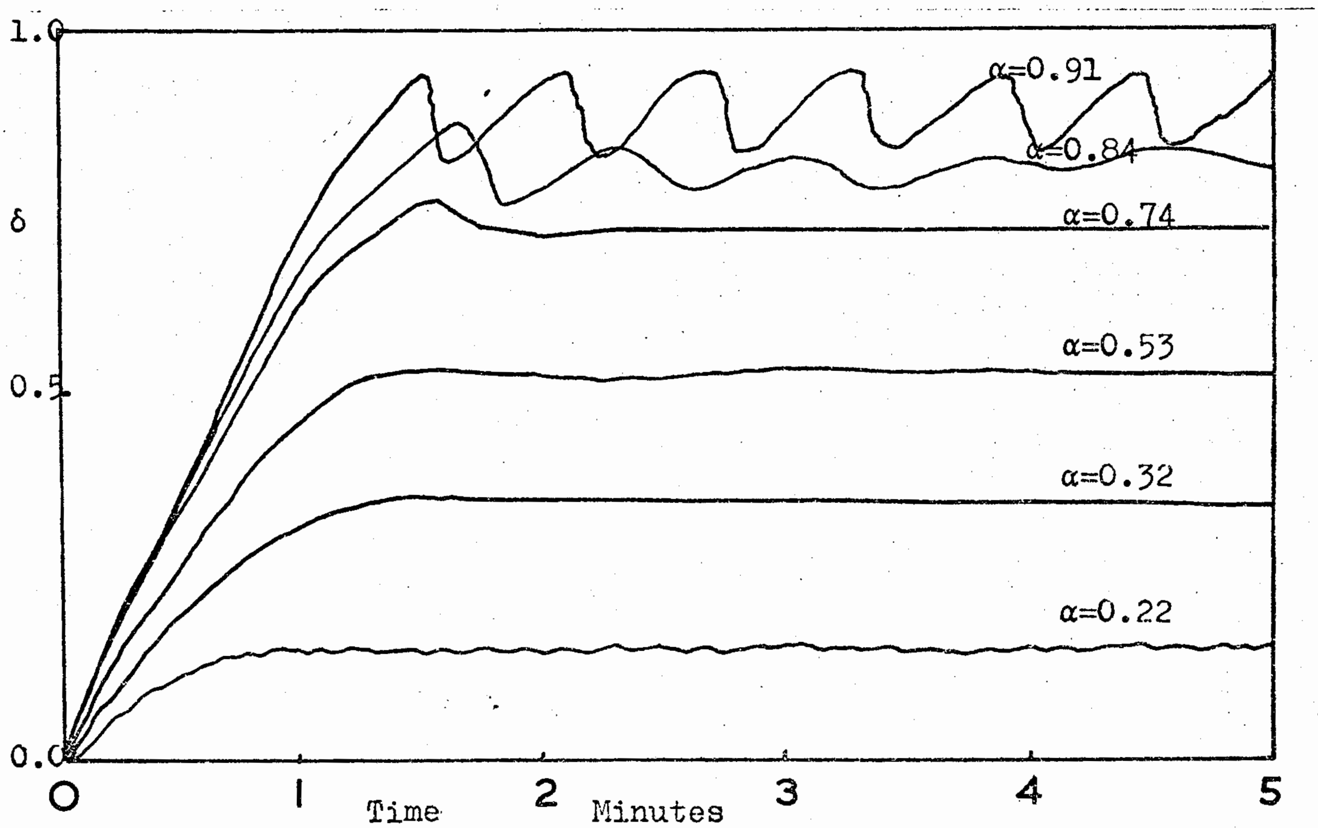


Figure 4-5 Variation of δ with Time for Different Values of α

The control policy of a naive operator attempting to control the third-order system gives rise to an unstable loop, and the state-trajectory of the system tends to follow the boundaries of the state-space. Thus, the initial sub-environment may lie entirely outside the desired sub-environment and will correspond to a nonlinear, rather than a linear, system. A suitable training controller will be one which attempts to maintain a linear sub-environment by cancelling the disturbance and stabilizing the control-loop; this is the effect of training controller shown in Figure 4-4.

Since the desired sub-environment is a region about zero in the state-space of the controlled element, it is possible for the trainer to detect by direct measurement whether or not this is being maintained. Under the experimental conditions the bounds on the error itself were very much more stringent than those on its rate or acceleration, and hence the value of the output of the controlled element was a sufficient indication of the effective sub-environment. A tolerated magnitude of error was fixed to define the boundary of the desired sub-environment, and the strategy of the trainer was such as to increase the difficulty of the task when the error was within tolerance and decrease it otherwise. This was achieved by the integrator in the training loop, shown in Figure 4-4 and described in Section 4.4. However, this may now be seen as acting continuously to maintain a sub-environment, rather than as a device for keeping the mean error constant; in this particular situation the two viewpoints are equivalent, but generalization to other situations follows from the sub-environment rather than the error-based approach.

4.4.2 The Training Strategy and the Second Training Theorem

The sub-environment interpretation of the training strategy given to the previous section leads to a further analysis of the trainer if it is noted that the desired sub-environment is obtained with plant parameters that lie on the stable side of the controller's stability boundary in the natural-frequency/damping-ratio plane (where a stability boundary is a line of constant mean-error in the plane). Equation 4.18 indicates that the mean error is a very rapidly increasing function of the natural frequency, w_n, so that the stability boundary is well-defined with respect to w_n, and Figure A4-2(ii) shows boundaries for different mean-errors for human operators. Since the feedback trainer is attempting to keep the mean-error constant, it may also be seen as attempting to keep the trainee on, or within, its current stability boundary.

Learning may then be seen as a movement of the stability boundary towards lower values of k and w_n. This interpretation is of particular interest because the trainee is effectively being described in terms of the set of tasks for which it is satisfactory. Informally one may say that the stability boundary is expected to move outwards provided the current plant parameters are on the stable side of it, and the movement will be most rapid if the parameters are near to the boundary - the rationale for this being the sub-environment argument of the previous section. However, this may now be re-phrased in terms of the adaption-automaton of the trainee, with the sub-environment phenomenom being the epistemological basis for constraints on the adaption automaton as discussed in Section 3.4.

Consider the terms of the Second Training Theorem (Section 3.3.2) and its extension to a lattice of tasks (Section 3.3.3). Let a task be defined as a fixed period of interaction with a plant of certain values of natural-frequency, w_n, and damping-ratio, k. Consider the order relationship on the two-parameter family of tasks $t(w_n,k)$ such that -

$$t(w_n,k) < t(w_n',k') \Leftrightarrow \text{ either } w_n > w_n' \text{ or } k > k' \text{ or both} \qquad [4.19]$$

that is, from Equation 4.13 or 4.18, one task is higher than another in the order if the mean error for a relay controller (and, from Hall's results in A4.2.1, for a human operator) is greater for that task than for the other. Hence the order on the tasks corresponds to the order of the mean error and stability boundaries.

It is now possible to re-phrase the concept of learning as movement of the stability boundary - because, for at least some range of values of w_n and k, we expect the stability boundary for a task, if it is originally near the task, to move away from it and hence encompass other tasks with higher w_n or k, we have -

given the ordering of tasks of Equation 4.19, there exists a range of tasks from t_a to t_b, such that -

$$\forall t : t_a \leq t \leq t_b \ \exists t' : t < t' , A(t) \subset P(t')$$

that is, performing the task t (within the region P(t)) causes the trainee to become potentially adaptive to a task, t'. This interpretation gives a form of condition (ii) of Theorem 3.2 - condition (i) may be satisfied is $t_a \leq \tau \leq t_b$ - whilst condition (iii) effectively requires that there is some task (value of w_n and k) in the range t_a through t_b for which the trainee is able to exert stable control at any stage of training.

The main differences between the training strategy used in the proof of Theorem 3.2 and the feedback trainer described in this Chapter is that the actual trainer is continuous rather than discrete and has no inbuilt random behaviour. In practice these differences are small because the human operator himself injects an effective random component, the 'remnant' (A4.2.1), and also because since the trainer is integrating the modulus of an oscillating error signal it is actually changing the difficulty in a 'discrete-plus-jitter' type of mode.

Thus the feedback trainer used in the experimental studies may be regarded as a teaching-machine varying the 'difficulty' of a task according to the performance of the trainee; as an attempt to maintain the optimum 'sub-environment' for learning the required task; or as a realization of a training strategy based on fairly general constraints upon the adaption-automaton of the trainee. In the current training situation all three interpretations are clearly closely related - however, each offers a different basis for generalization in the interpretation of the results obtained.

4.4.3 Implementation of Feedback Trainer

The feedback trainer of Figure 4-4 was realized on an analog computer, using chopper-stabilized operational amplifiers and 1 per cent accuracy components. The time constant of integration in the training loop, and the tolerated mean error were both adjustable, and reasonable values of these variables were established during the initial informal trials. These parameters, and the effects of changing them, are analysed in more detail in the following section on the trainer's stability.

The output of the feedback training integrator could be coupled to any combination of the three servos adjusting the parameters of difficulty of the task. In practice, one or two of the servos were locked in fixed positions and the remainder were coupled to the integrator. The space of all possible training environments, defining 'tasks', is three-dimensional, since the difficulty increases as:-

(i) The disturbance is increased from zero to its maximum value.

(ii) The undamped natural frequency is decreased from its maximum value to zero.

(iii) The damping ratio is decreased from its maximum value to zero.

The trajectories of the training environment through this three-dimensional space were reduced, by locking or co-varying the servos, to single dimensional paths along lines either parallel to one of the axes or diagonal to a pair. In the informal experiments on testing only the natural-frequency varied (for a range of fixed values of the damping ratio and a single value of the disturbance), and in the experiments on training the natural frequency was locked and the amplitude of the disturbance and the damping ratio co-varied. These particular restrictions had the advantage of making the theoretical analysis simpler, but were otherwise arbitrary choices.

This particular form of feedback trainer is similar in its strategy for variation of the task difficulty to the feedback testing systems of Jex, McDonnel and Phatak (1966) and Kelley (1967) (the same strategy was suggested by Hudson (1964) in his recommendations for future work).

4.5 Stability and Dynamics of the Automated Feedback Trainer

The expected behaviour of the feedback trainer is that it will maintain the desired sub-environment by variation of the parameters of the training controller, or, more precisely, that it will adjust the difficulty of the task to cause the mean error to come to a certain level and maintain it at that level. With a non-adaptive controller, the only stable value of difficulty will obviously be uniquely determined by the ability of the controller to regulate the control system. It is not obvious, however, that the feedback training loop is stable, and indeed it may be shown that with certain forms of controller instability may occur. Analysis of the loop stability is complicated by the number of feedback loops operative and the nonlinearities in both human operator and automatic trainer, but a simple analysis may be based on linearization of the outer, parameter-adjusting, loop.

However, in Section 3.3.3 it has been shown that even at a highly abstract level a feedback trainer of the type under discussion may be expected to show behaviour similar to that of a linear servomechanism, and hence an analysis of the trainer basdd on linearization may be useful. In the following section a stability analysis of the feedback trainer based on linearization of the training loop is described, and the loop dynamics are derived for a relay controller acting as operator. In later sections the analysis is confirmed by experimental studies of human and automatic controllers.

4.5.1 Derivation of the Dynamics of the Training Loop

Consider first the variation of the mean error modulus with change of difficulty, that is, the natural-frequency, damping-ratio, or disturbance in the main loop. For an operator with a fixed control policy, at zero disturbance, there will be both an amplitude-dependence and a time-dependence in this variation. If his control policy is nonlinear so that a limit-cycle forms, then, under the experimental conditions, the mean error modulus, $\bar{e}$, increases monotonically in amplitude for decreasing natural-frequency and damping-ratio; the theoretical results of Section 4.3.1, and the graphs of Figure 4-3, illustrate this dependence for a nonlinear, 'relay' controller. The limit cycle takes time to build up and decay as the task difficulty changes, and this time dependence may be approximated by an exponential lag with a time contant of the same order as the period of the limit cycle. If the control policy is linear, however, there is no stable limit cycle, and the error modulus rises exponentially in time to its maximum possible value on one side of the stability boundary, and decays exponentially to zero on the other.

The relation to be expressed approximately in linear terms is that the mean error modulus and its rate of change are together linearly dependent on the difficulty of the task for the operator. Since the error modulus cannot be less than zero, for the linearization it must be expressed as a deviation from some positive value, and this is conveniently taken as the tolerated level, e_o. It is clear that the error must increase with the difficulty of the task and decrease with the operator's ability, but of these only the task difficulty is independently measurable and it is convenient to relate the operator's ability to this. Let the task difficulty increase monotonically with the increase of some parameter, δ , and let the operator's ability be defined in related units as α, such that when $\alpha=\delta$ the mean error modulus, $\bar{e}$, is at the tolerated level, e_o.

The behaviour of the mean error modulus may now be approximated by the equation:-

$$f(\bar{e} - e_o) + g^2 s\bar{e} \doteqdot \delta - \alpha \qquad [4.20]$$

where s is the time differentiation operator. The constant, f, will be large relative to g for switching mode controllers, whilst f will be a function of the disturbance and b large for linear controllers. It is clear from Equation 4.10 through 4.14, that f and g are functions of $\bar{e}$, δ and α - however, for small deviations from a possible stable point, Equation 4.20 will be valid.

The relationship between task difficulty and the mean modulus error, realized by the integrator in the training loop, is:-

$$s\delta \quad = \quad -h^2(\bar{e} - e_o) \qquad [4.21]$$

where $1/h^2$ is the time constant of the integrator. Combining Equations 4.20 and 4.21, we obtain:-

$$\delta \ + \ (f/h^2)\ s\delta \ + \ (g^2/h^2)\ s^2\delta \quad = \quad \alpha \qquad [4.22]$$

This is the overall equation for the training loop dynamics, and it may be seen that δ follows α through a second-order transfer-function with undamped natural frequency of h/g radians/second, and a damping ratio of f/2hg.

If there is no true limit cycle and f is zero then so is the damping ratio and the training loop becomes oscillatory. It was found experimentally that this did occur when a linear controller was used as the 'operator', and the difficulty oscillated widely. However, this has no practical effect since the human operator's control policy is sufficiently nonlinear to cause the value of f to dominate over that of g^2. When this is so, and g^2 can be neglected, Equation 4.22 reduces to:-

$$\delta \ + \ (f/h^2)\ s\delta \quad = \quad \alpha \qquad [4.23]$$

so that again δ follows α, but this time through a simple exponential lag of time-constant, f/h^2.

These equations give the transient behaviour of δ in response to changes in α; but do not allow for the error signal itself having an oscillatory form. The effect of this on the steady-state value of δ may be approximated by assuming that, under steady-state conditions, the error signal has the form:

$$e \quad = \quad (1 + \sin(wt))e_o \qquad [4.24]$$

that is, an oscillatory signal, always positive, with a mean equal to e_o and a frequency equal to that of the oscillations in the lower control loop (given by Equation 4.10 for the relay controller). Equations 4.21 and 4.22 then give the steady-state solution for δ as:-

$$\delta \ = \ \alpha \ + \ (h^2/w)e_o\cos(wt) \qquad [4.25]$$

Equations 4.22 and 4.25 indicate that, if the time-constant of the training loop integrator is sufficiently long so that h is small, then the difficulty adjustment is well-damped and little of the oscillation in the tracking task control loop appears in the difficulty variation.

The linearization of the training loop is a very strong approximation and requires empirical checks of its validity. In the following section a quantitative check is given using the analysis of the relay controller described in Section 4.3.1, whilst in Section 4.5.3 qualitative checks of the predicted loop behaviour with human operators are described.

4.5.2 Behaviour of the Training Loop with a Relay Controller

In the experiments on the behaviour of the training loop with non-adaptive controllers, only w_n was varied, with k at some fixed value, and the amplitude of the disturbance, a, either zero (for relay controller) or at a small fixed value (for human studies). It is convenient to have the difficulty, δ, vary from zero (easiest) to unity (most difficult), and set:-

$$\delta = 1 - w_n/5 \qquad [4.26]$$

The tolerated level of mean error modulus, e_o, was set at ·175, because this was found to give a demanding but comfortable control task for human operators when the difficulty reached its steady state. This gives an error oscillation whose amplitude is well below the limiting value, and hence the mean error modulus may be determined from Equation 4.13; it is convenient to write this in the form:-

$$\bar{e} = (4ML/\pi^2 k)(1-2k\mu w_n)^{3/2}/(w_n^3(1-2k\mu w_n+\mu^2 w_n^2)^{1/2}) \qquad [4.27]$$

Substituting $\bar{e} = e_o$ in this equation enables the value of w_n, and hence δ, to be derived for which $\alpha = \delta$.

The constant, f, may be seen from Equation 4.20 to be the steady-state rate of change of δ with $\bar{e}$ for $\bar{e} = e$, and may be derived in a convenient form by logarithmic differentiation of Equation 4.27:-

$$f = 1/(5e(4/w_n + 3k\mu/(1-2k\mu w_n) - (1-k\mu w_n)/(w_n(1-2k\mu w_n + \mu^2 w_n^2)))) \qquad [4.28]$$

If Equation 4.20 were truly linear, g^2/f would be the time constant from the moduls error to change in response to changes in δ. It is a difficult term to estimate, however, because it is so dependent on the mode of behaviour of the nonlinear relay servomechanism. When δ, and hence w_n, is changed, the relay servo tends to enter a 'chatter' mode (Gibson 1963 p.445) in which the trajectory in the $(e,\dot{e})$ phase-plane spirals around the origin with a natural frequency of $w_n(1-k^2)^{1/2}$, oscillating about this trajectory with a natural frequency of w. The period of one spiral was taken to be a reasonable value at which to set

g^2/f to provide some quantitative comparison of theoretical and measured loop dynamics. Hence:-

$$g^2 \approx 2\pi f/(w_n(1-k^2)^{1/2}) \qquad [4.29]$$

The constant, h^2, was set at the value, $h^2=0.057$, and was such that, when e = 0, to rise from zero to maximum difficulty took about 100 seconds.

μ	α (theor)	α (expt)	f	g	damping train-loop
0.000	0.072	0.08	1.77	1.65	2.24
0.087	0.215	0.22	1.17	1.46	1.67
0.150	0.318	0.32	0.852	1.34	1.33
0.225	0.425	0.42	0.558	1.18	0.988
0.325	0.531	0.53	0.333	1.011	0.691
0.500	0.653	0.64	0.150	0.788	0.398
0.750	0.752	0.74	0.055	0.592	0.194
1.250	0.844	0.84	0.0162	0.386	0.0878
2.500	0.921	0.91	0.00216	0.198	0.0209

Table 4-4 Training Loop Dynamics

Values of the training loop dynamic parameters derived from the theoretical equations are given in Table 4-4, together with the experimentally measured values of α , the 'ability' of the relay controller in terms of the difficulty of the task. It may be seen that there is very close agreement between the theoretical and measured values of α , showing that the describing function analysis of the relay control loop is adequate, and that the training loop sets up the correct steady-state conditions. The degree of agreement between the theoretically-derived dynamics of the training loop and the measured results may be determined from Figure 4-5. This shows the variation of δ with time in the experimental system for different values of α. The range of damping ratios predicted by the theory is similar to that found experimentally, and a damping ratio of 0.7 for $\alpha = 0.5$ ties in closely with the measured value.

A more detailed examination of the relationship between measured and predicted dynamics, however, shows up major discrepancies. For the low damping ratios in the training loop, the period of oscillation should be $2\pi g/h \approx 26g$ seconds, and this leads to theoretical values which are very much lower than those measured. From the form of the oscillations in the graphs for low damping ratios, it is clear that the time constant in increasing δ is very much less than in decreasing it, and that Equation 4.20 is a very coarse approximation. The rise time of the graphs for various levels of α also cannot be predicted from the steady-state, and the actual rises are rate-limited by the limiting of the error signal.

These results represent the limit to which a quasi-linear analysis of the behaviour of the feedback trainer may usefully be taken - it accounts for the steady-state behaviour and important characteristics of the dynamic behaviour, but does not accurately predict the detailed dynamics. In the following section the implication that the training loop will be stable and responsive for relay-like controllers, as is the human operator in this situation, is confirmed by experimental studies with human controllers.

4.5.3 Behaviour of the Training Loop with Human Controllers

In the experiments with human operators, the control was a rolling-ball joystick of diameter 9 inches, which gave an output of 0.55 units for 1 degree rotation and a maximum output of $\pm$ 3.8 units. The error was displayed on a 5 inch diameter oscilloscope as a horizontal deviation from a central vertical line with a sensitivity of 0.38 units for 1 inch of movement. The sense of the control was such that a movement to the left sent the spot to the left, and the joystick itself was centralized by light springs. At 0.5 inches on either side of the central display marker were two other vertical markers, and the operator was instructed to move the control so as to keep the spot on the oscilloscope within the outer markers. The experiments took place in a soundproof room, 9'x9', dimly but comfortably lit and free of experimental apparatus except for some tables, a typists' chair for the operator, and the oscilloscope mounted 3 feet from the ground about 3' from the chair.

The parameter of difficulty varied by the feedback trainer was the undamped natural frequency of the third-order transfer function, w_n, according to Equation 4.26 as for the experiments with relay controllers. The damping ratio, k, was set at one of a range of fixed values, for example 0.25, 0.75, 1.0, so that the minimum natural-frequency at which the controller was stable for a given damping ratio was obtained. It was found desirable to have some small disturbance injected in the loop, since otherwide the human operator's tended to adopt a control mode in which they brought the position and velocity of the error almost to zero, centred the joystick, and then ceased to make control movements until the error had become large again. This conditional-control mode enabled the difficulty to increase whilst the controller was ineffective until the tracking loop was potentially unstable, but, since it had effectively been opened, there was no effect on the error. When the loop was closed again, however, oscillations rapidly built up, forcing the difficulty to decrease. A small disturbance prevented the operator from adopting this mode of control, and an amplitude of 0.0033 units was found to be adequate. This was added to the output of the joystick, so that the total input was -

input to tracking system = joystick ouput + 0.0033sgn(sin(t/10))

Figure 4-6 shows the variation of difficulty as a function of time for various human operators under various conditions. Graph A is that of an operator new to the task with the integrator constant in the training loop set to $h^2 = 0.057$ - it can be seen that the difficulty rises slowly and irregularly to an asymptotic value, but oscillates somewhat about this. Graph B is the second trial for the same operator, and it may be seen that the rise to the asymptote is faster - when this final range of values was approached, the value of h^2 was decreased by a factor of four to $h^2 = 0.0143$, and this smooths out the final value of difficulty. This procedure of changing the integrator time-constant in the training loop to obtain a fast rise to the asymtote but then a smooth reading of it was adopted in all the experiments on the use of the feedback trainer to test human perceptual-motor skills; with RAF pilots, who found the tracking task with rolling-ball and oscilloscope simple and natural, it was found that between one and two minutes of tracking were adequate for a near-final value to be reached. Graphs C and D were generated by highly-skilled flying instructors, and it may be seen that the ultimate level of difficulty is rapidly reached and closely maintained.

It was found that the tracking task was very fatiguing, and that between four and ten minutes operation was all that could be reasonably demanded, even from pilots and well-practiced operators. After the first trial, no appreciable learning was noticeable, although there was a clear separation between individuals in ability. No very long series of fifty or more spaced trials was carried out, however, since learning over long periods was not of interest in the context of the present experimental studies, and this might have shown definite evidence of learning; Hudson (1964), using similar task dynamics, gave ten hours

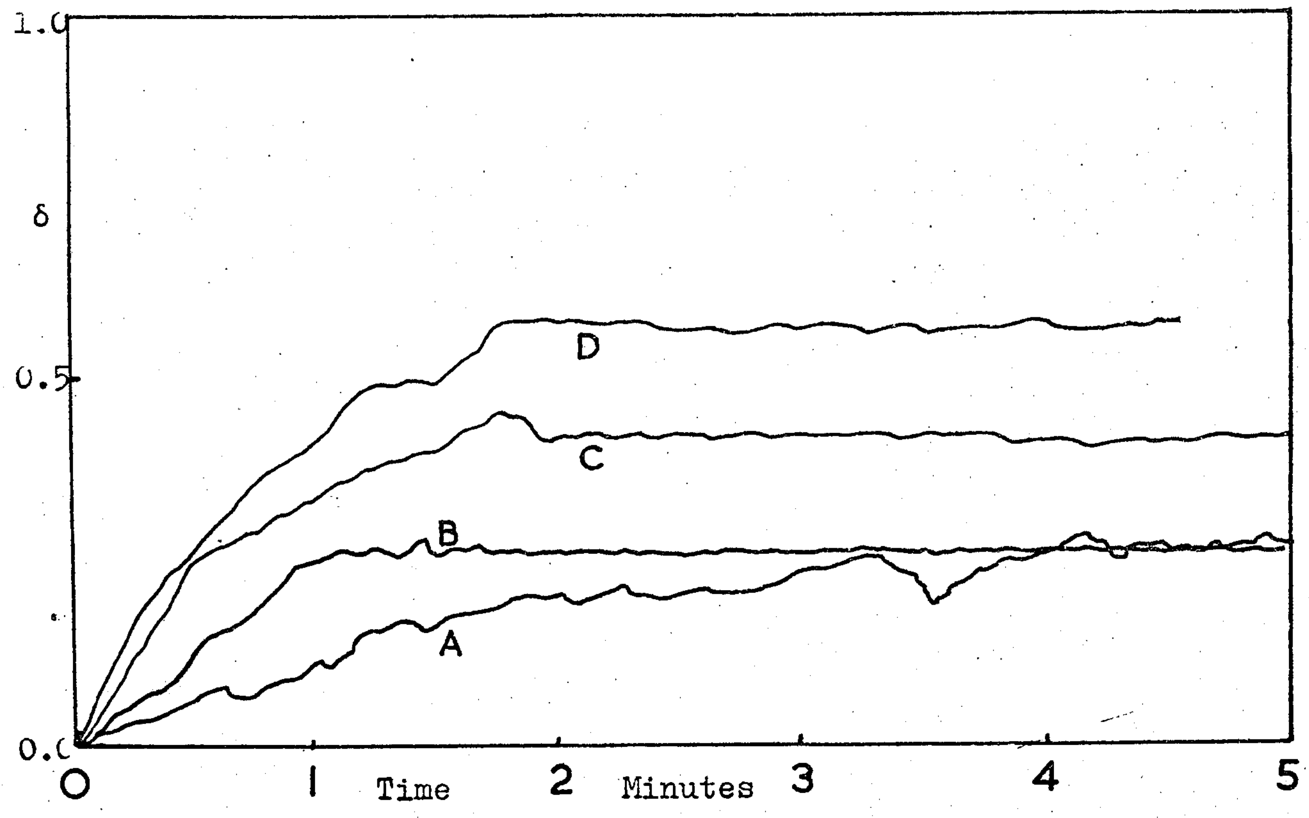


Figure 4-6 Variation of Difficulty with Time for Human Controllers

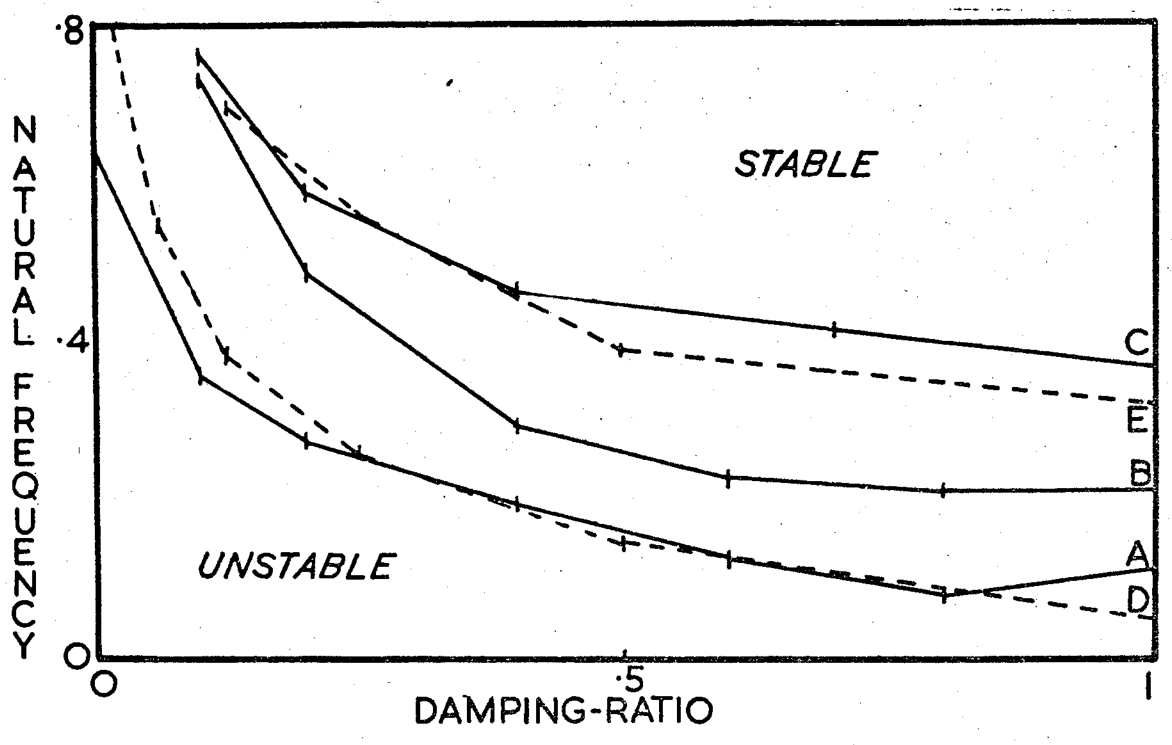


Figure 4-7 Stability Boundaries of Human and Automatic Controllers

training in total to each of his subjects. The relevance of these findings on fatigue and learning to the study of the utility of a feedback trainer is discussed in Chapter 5.

4.6 Use of Feedback 'Trainer' for Testing

It has been suggested by a number of workers, particularly Kelley (1967, and Prosin 1968), that an important application of feedback 'training' systems is not to training itself, but rather to the accurate measurement of perceptual-motor skills. By measuring, for a given operator, the difficulty at constant error, rather than the error at constant difficulty, a performance-feedback system can greatly increase the sensitivity of tests for evaluating an operator's capabilities. As Poulton (1965) has noted, tests at constant difficulty lack discrimination at the upper and lower end of the range of abilities, if the level of performance extends into regions where it is physically limited. This is clear from Figure 4-3, showing mean-error for various controllers as a function of difficulty (variation of undamped natural frequency). A test at constant difficulty corresponds to a vertical line in this figure, and it may be seen that such a line, at any value of difficulty, effectively dichotomizes the controllers into those whose mean error is low, and those for whom it is high; the controller's capability is given one of two values instead of being set out on a continuum. A test at constant error corresponds to a horizontal line in Figure 4-3, and that for $\bar{e} = 0.175$, the value used in the experiments, is shown as a dashed line; it can be seen that this intercepts the curves for different controllers at approximately equal increments of difficulty, and discriminates well between their capabilities.

No validation studies of the feedback trainer described has been carried out in order to test its utility in measuring some aspect of perceptual-motor ability. The 72 RAF pilots who took part in the training experiments described in Chapter 5 were tested, as described in Section 4.5.3, at three values of the damping ratio. Each test lasted five minutes and they were given in the order - k=0.50,0.25,1.00, to combat the effects of possible learning. For purposes of experiments on differences due to training, the population tested has been selected for their homogeneity, established through RAF selection procedures, and hence were unsuitable for the validation of tests of individual differences. However, the correlation coefficients between the values of α measured for each value of k are an indication of the replicability of this type of test; these are given in table 4-5, and it can be seen that they indicate a high degree of replication, given the homogeneous nature of the population; α itself ranged from 0.00 to 0.53.

k=	0.25	0.50	1.00
0.25	1.00	0.63	0.74
0.50	0.63	1.00	0.70
1.00	0.74	0.70	1.00

Table 4-5 Test Correlations

The measurement of difficulty, in terms of w_n, for constant error at various values of k, enables the stability boundary of a controller to be plotted out automatically. Figure 4-7 shows measured boundaries for three human operators (A,B,C), obtained in this way, and those of two relay controllers (D,E). These may be compared with contours of constant mean error obtained by Hall (1963), and shown in Figure A4-2(ii).

4.7 Conclusions

The theoretical and experimental studies of this chapter demonstrate that the particular automated feedback trainer developed, based on a third-order tracking task and a constant mean-error feedback criterion, is a viable system, free of artifacts such as might be caused by its instability. The close agreement between experimental and theoretical results with relay controllers shows that the equipment itself, used in the experiments described in Chapter 5, is reliable and capable of high-accuracy measurements. The theoretical foundations developed indicate directions for the extension of the trainer to other skills, and the type of problem that will be encountered.

The experimental situations described in this chapter have been such that the controller cannot, or does not, learn the control skill. In the following two chapters, experiments are described in which both human operators and adaptive controllers learn the skill under a variety of training regimes, in order to evaluate the utility of feedback training.

CHAPTER 5 EXPERIMENTAL EVALUATION OF FEEDBACK TRAINING

5.1 Considerations in Experimental Design

Previous chapters have laid the theoretical foundations for the study of training, and have lead to the development of the particular form of feedback trainer for a tracking skill described in Chapter 4 . This has been shown, both by theoretical analysis and through experimental trials, to be a viable system, free of artifacts such as might be caused by instability. The 'difficulty' of the tracking task follows the 'ability' of th operator in a stable manner and in a reasonable time, so that the 'trainer' may certainly be used to test the ability to perform the tracking skill. It remains to be shown that its concomitant maintenance of the desired sub-environment has the expected effect of maximizing the rate of learning of the skill, and, hence, that the trainer is a useful device.

There are many experimental and methodological problems in the comparative study of various training techniques, and conclusions drawn from experimental studies which neglect these problems may be completely invalid. In the following section these problems are outlined briefly together with the approach taken in the present studies to overcome them; results of some informal experiments to estimate the magnitude of certain problems are also outlined.

5.1.1 The Nature of 'Good' Performance

To determine whether one training technique is better than another, it is necessary to have some measure of the goodness of the end-product, that is, the trained human operator. Usually training is considered to be required for some reasonably well-defined task, and the performance of the operator on this task is evaluated by transferring him to it after training. However, questions arise as to what inferences may be made from this about his performance in the range of task situations he is likely to meet in practice; whether his skill is robust and a reasonable standard of performance can be maintained under stress; whether the performance can be maintained for a period of time; and whether he retains the skill after a period of time without use. Obviously, in practice, one does not want to train so specifically that slight changes in the task cause great deterioration in performance, and neither should some degree of 'stress' cause such a deterioration. Equally, one requires the standard of performance to be maintained over a reasonable interval, and expects the skill to be remembered even though it has not been performed for a while.

These various approaches to the evaluation of training are to some extent, independent, and may be treated in the context of the taxonomy for adaptive behaviour developed in Chapter 2. The operator should become adapted to the task, so that his performance not only attains a high standard but remains stably so. He should be jointly adapted to all variants of the task, of interest, including those corresponding to differing degrees of stress. He should be compatibly adapted to the main task with respect to all the 'tasks' which may fill the intervals between performance of the main task, including periods of 'inactivity'. The criteria for evaluation may be rigorously defined in these terms, but the problem remains of determining whether the criteria are satisfied from observation of behaviour in appropriate experimental circumstances.

In the present studies neither the maintenance of performance over extended periods of time, nor the retention of the skill, were measured. For various reasons, the performance on each of several different tasks was evaluated at the end of the training phase, so that over-specificity of training could be evaluated; by the nature of the task, it was unlikely to occur.

5.1.2 Push-button Controls and Fatigue

It has been noted in Sections A4.3.5 and 4.5.3. that the use of the rolling ball joystick in a continuous tracking task produces complaints of fatigue after a few minutes, whereas the use of discrete push-button controls produces no complaints of fatigue even after extended periods of tracking. A decrement in the performance of a continuous tracking task after as little as one minute has been noted by other workers (Ornstein 1963), and improved performance with push-button controls has been explained as an effect of the reduced computational loading on the operator (Young and Meiry 1965). A further important advantage of the push-buttons in the present context is that they offered the possibility of withdrawing completely from any interaction with the system. With a joystick control, such withdrawal is liable to cause large errors, unless the control is very light and accurately self-centering at zero output, whereas the push-buttons give zero output immediately when they are not depressed.

The effects of fatigue are a minor nuisance in studies of human control strategies where short tracking runs may be used, but in the study of learning they present a major problem. Even breaking the tracking sessions up into short intervals is not realistic way of overcoming fatigue, since continuous control tasks causing acute fatigue are not met, for obvious reasons, in the systems which a human operator is normally required to control. In order to estimate the magnitude of the difference between types of control and determine whether a push-button system would be reasonable, an initial informal experiment was carried out comparing the time for which operators were prepared to track, in a fairly free situation, with each form of control.

The continuous joystick was that described in Section 8.5.3. The push-button controls consisted of two microswitches with half inch flat buttons, mounted in the arms of a typists' chair at such a position that they were comfortable for all operators. The output of either push-button was a pulse of 10 milliseconds width, $\pm$ 0.12 units in amplitude, which may be regarded as an impulse of 0.0012 units times the Dirac delta function. Either control fed into the third-order system described in Section 4.2.1, and the experimental environment, controls, and so on, were as described in Section 4.5.3.

Visitors to the laboratory and other subjects were given the opportunity to try out the tracking system and to track as long as they wished; the interval of voluntary tracking was noted. These experiments were informal in that the tracking task varied from person to person, and the majority of operators tracked under uncontrolled conditions. Also, some performed both tasks, othersonly one - This effect was 'balanced' by always using a further operator under than dondition. The discrepancy between the times is, however, so great that it is considered worthy of note; no more formal study was made because the effect is not central to the objectives of the present work.

Table 5-1 shows the duration, in minutes, for which operators tracked voluntarily with the rolling ball control and with the push-buttons. The results are given in chronological order, with the operators and conditions arbitrarily labelled so that the extent of balance in the experimental 'design', and possible auxiliary effects, may be seen. The mean time with the push-buttons is 20.0 minutes, compared with 6.1 minutes for the joystick; this difference is significant at the p = 0.001 level using the two-sided Mann-Witney U-test (Siegel 1956). This result bears out the verbal comments of operators, and even if one ascribes it, for example, to an increased interest in the push-buttons, it indicates that sustained tracking with the push-buttons was more acceptable to the operators. Hence, the push-button controls were used throughout the experiments on training.

Operator	Condition	Joystick	P.B.	Operator	Condition	Joystick	P.B.
A	α	3.5		H	γ	6.1	
B	α	2.7		I	α		8.5
C	α	10.2		J	α		29.1
D	α	2.1		K	β	21.0	
D	α		14.4	L	δ		33.0
E	β	5.2		L	δ	5.1	
E	β		10.7	M	δ	1.7	
F	α		24.3	M	δ		10.4
G	β		15.4	N	β		18.3
G	β	2.0		P	δ	8.0	
H	γ		11.7	P	δ		44.0

Tabel 5-1 Voluntary Tracking Times with Different Controls

5.1.3 Use of Complex Controls to Give Scope for Learning

In the evaluation of different training techniques, if the skill to be learnt is either such that little learning is possible in the time of the experiment, or learning is a function of time rather than environmental conditions, then clearly the effects of different training techniques will be indistinguishable. In a practical situation this implies that training is irrelevant, but, since the objective of the present study was to demonstrate a difference between training techniques, it was considered desirable to develop a task which gave adequate scope for learning.

It was also desirable that the task could be learned by a naive operator to a level of performance approaching that which was ultimately possible in a reasonable time - 30 minutes say, both to avoid problems in obtaining trainees, and to avoid artifacts due to differential rates of initial and final learning under different training regimes. If, in comparing two training techniques, the performance under one regime is uniformly better than that under another, then there is no problem in determining which is best. If however, the relative performances interchange their relationship at some time in the training period, then this might be missed if the experimental period is too short. This phenomenom might be expected in open-loop training at low and high levels of difficulty, where the low level of difficulty might give rapid initial acquisition, but be inadequate for later learning.

Previous workers on feedback training have used simple tracking skills as tasks for which training is required (Chernikoff1962, Hudson 1964), similar to that used in the experiments described in Chapter 4. However, the skills involved in performing the task have very little structure and only involve the acquisition of a certain control policy, so that the scope for learning is limited. Since there are few sub-skills and there are no strong interactions between them which makes satisfactory performance of one necessary to the learning of another, then, as discussed by Pask (1965), it is unlikely that feedback training will give very great advantages.

The requirement for tasks with interactions and scope for learning is not entirely methodological - in reality, tasks involving the performance of a single skill at a very high level of performance are very rare. For example, the task of flying an aircraft is difficult, not because any individual tracking task has anywhere near the difficulty of those commonly used in the laboratory, but because a large number of different activities have to be integrated together, and poor performance of one creates an undesirable sub-environment for learning another - an aircraft diving after a stall is not a suitable environment for learning the finer points of rudder control. In the majority of real-life perceptual-motor skills, such as flying, driving and typing, the skill to be acquired is a complex of many minor sub-skills, and the problem of learning is to integrate them into a cohesive whole.

There are many possibilities for tracking tasks involving interacting sub-skills to be set up in the laboratory. For example, the cross-couplings between the various axes in an aircraft might be simulated in a two-dimensional tracking task, with differing dynamics in the two axes and strong cross-couplings between them. In the present study, for purposes of simplicity and ease of interpretation of the results, it was considered desirable to use a single-dimensional tracking task, and an interaction was introduced by use of unnatural controls, the push-buttons which reverse their sense at each push, described in Section A4.3.5. The function of these controls is difficult to determine when the system is not under control, but, equally, the system is virtually impossible to bring under control until the function of the push-buttons has been, at least partially determined.

5.1.4 Choice of Feedback, and Comparative, Training Situations

Given the use of the reversing push-buttons as controls, the parameters of task difficulty to be adjusted by the training loop were chosen to be those, which from theoretical considerations and initial informal experiments, most varied the difficulty of learning to use the controls. Since a prime requirement for performance feedback when using the push-buttons is to be able to note the direction of motion induced in the CRT spot, both damping-ratio and amplitude of disturbance were expected to have major effects on speed of learning - when the operator presses a button, an unexpected zero-crossing of the disturbance may cause the spot to move in the 'wrong' direction, and, equally, a low damping-ratio leads to oscillations which make it difficult to determine the net direction of motion. Hence, in the dynamics of Equation 4.1, the undamped natural frequency was set at the mid-range of the values used previously - w_n = 2.5 radians/second; the damping ratio was varied from k=0 to k=0.5, and the amplitude of the disturbance was varied from 0.0033 units to zero, as the parameter of difficulty, δ, varied from unity to zero.

When the difficulty was near zero, so that the disturbance was low, it was possible for the error to become zero and remain near zero whilst the operator took no control action - the difficulty would then rise slowly until the disturbance became appreciable, and then fall back to zero. To remove this artifact, a small constant term was added to the output of the push-buttons and the disturbance to form the total input to the system. This had an amplitude of 0.00033 units, and was sufficient to cause the spot to drift over the right-hand side of the screen when the difficulty was zero and there was no control input. If the output of the push-button controls is written as 0.0012u, where u is plus or minus the Dirac delta function, then the overall equation for the loop dynamics is :-

$$s(s^2 + 2.5(1-\delta)s + 6.25)e = 1.2u + 0.33 + 3.3\,\mathrm{sgn}(\sin(\pi t/10))\delta \qquad [5.1]$$

It may be seen from the coefficients of the push-button input, drift and disturbance, that the operator only has to push the buttons three times a second to neutralize the disturbance, whilst one push every four seconds is sufficient to overcome the drift term. It was found that the error tolerance used in the training loop of the system described in Chapter 4 was unreasonably stringent when tracking with push-buttons, and that a mean error tolerance of 0.34 units gave a comfortable, and acceptable, level of performance when the training loop was in a steady-state, The time-constant of the training integrator was set at the higher of the two values used in testing since this gave the best compromise between speed of following the operator's learning curve, and superimposed noise. The equation of the training loop is then:-

$$s\delta \quad = \quad 0.0143(0.34 - \mathrm{MOD}(e)) \qquad [5.2]$$

where MOD(e) is the amplitude of the error. Since this is bounded below by zero, and above by 0.95 units, the minimum time for δ to fall from unity to zero is about two minutes, and to rise from zero to unity is about four minutes.

The choice of alternative training situations for comparison with feedback training is clearly very great - if one particular value of difficulty is taken to define the task for which training is required, then fixed training at that same level is one obvious possibility - training at some other value of difficulty and then transferring to a test at the required level is the simplest form of open-loop training - a time-varying trajectory of difficulty is a more general open-loop training sequence. Clearly, only a limited number of alternative training techniques could be evaluated, and with the limited information gained in the initial informal experiments on the relative merits of different training techniques it was decided to use training at a constant level of difficulty as the open-loop technique for comparison with feedback training.

In order to provide an adequate evaluation of the operators' capabilities after training it was necessary to test their performance at several levels of difficulty (Section A4.4.1), and it was convenient to choose these also as the levels for open-loop training, since the same experimental results could then be used as a basis for the evaluation of fixed training and of open-loop training at higher, or lower, levels of difficulty than the required task. Three levels of difficulty were selected as a result both of the initial experiments with human operators, and of the computer-simulation experiments with learning machines described in Chapter 6. These levels were δ = 0.25, 0.50, 0.70, and their relative levels can be appreciated from approximate descriptions of the appearance of the system to the operator:-

δ=0.25 (L-low difficulty) Very easy for the skilled operator, and spot can be kept within plus or minus five per cent of CRT centre. With naive operators, the spot moves lackadaisically, traversing the full width of the screen slowly and regularly.

δ=0.50 (H-high difficulty) More demanding for the skilled operator, but it is still within his capability to keep the spot within plus or minus twenty five per cent of CRT centre, never letting it reach the edge of the screen. With a naive operator, the spot moves rapidly from one side of the screen to the other, and remains for a while at each edge.

δ=0.70 (V-very high difficulty) Approaching the limit of the highly skilled operator's control - he finds it difficult to prevent the spot reaching the edge of the screen occasionally. With a naive operator, the spot races about, both the system oscillation and the disturbance affecting its movement.

In the initial experiments it was clear that learning was virtually impossible at the very high difficulty level, and that training at δ=0.50 was the highest which would give useful results. The difference in situations between the low and high difficulty conditions was so great, however, that both were considered of interest. Hence, three separate training regimes were established:

(i) High Difficulty - H - the 16 operators trained under this condition had the level of difficulty set at δ=0.5 (H) throughout the training period. From the informal experiments, it was predicted that this group would show little learning and perform badly at all test levels of difficulty.

(ii) Low Difficulty - L - the 24 operators trained under this condition had the level of difficulty set at δ=0.25 (L) throughout the training period. It was predicted that some members of this group would learn to a high standard, but that others would not.

(iii) Feedback - F - the 32 operators trained under this condition started with δ=0 and had the feedback training loop operative throughout the training period. It was predicted that all members of this group would learn to a high standard.

The numbers trained in each group were chosen to maximize the information from training conditions of most interest, and to provide an adequate separation between the groups expected to be most similar.

The integrated error under each of the three test conditions, δ=0.50, δ=0.25, δ=0.70, was measured at the end of the training period. It was possible to regard any of these three levels as that for which training was required, and hence it was possible to compare fixed training (2 combinations), open-loop training at a higher level of difficulty (1 combination) open-loop training at a lower level of difficulty (2 useful combinations), and feedback training (3 combinations).

5.1.5 Effects of Individual Differences

The obvious way to evaluate the relative merits of different training techniques is to take an individual, train him under one regime and measure his performance on the required task, and then erase his learning, train him under another regime and again measure his performance. Unfortunately, as discussed previously, the adaption-automaton of the human operator is generally irreversible, and learning cannot be 'erased'. Hence, it is not possible to compare the effects of different training regimes on an individual, and, indeed, the same operator, before and after training, will probably show far larger differences in behaviour than are apparent between different operators before training.

Thus, it is necessary to compare the effect of different training regimes on populations of operators rather than individuals, and to take a large enough sample to ensure that the probability of assigning a disproportionate number of individuals of one type to one training condition is very low. The size of the group required to give a certain sensitivity to differences in the effects of training regimes reduces as the overall population becomes homogeneous, containing individuals similar in their characteristics and abilities. In the present study, RAF pilots at an advanced stage of training and selection formed the experimental population. They were a middle-stream group, who had passed through all the selection procedures testing general flying and navigation skills and personal qualities, but had failed to graduate to the more demanding aircraft. Thus the population was inherently homogeneous, and also very well documented, so that any effects of individual differences could be examined.

The use of non-human operators such as learning machines enables the identical individual to be trained under two or more different regimes. Clearly, experiments with learning machines cannot replace those with human operators in a study whose objective is the evaluation of training techniques for human beings. However, since the arguments of Chapter 3 suggest that the effects of different training techniques are, to some extent, independent of the nature of the trainee, studies with learning machines, although they cannot eliminate those with human operators, are an aid both to establishing experimental conditions, and to interpolating between results obtained for human operators. The experiments with humans described in this Chapter have also been carried out with learning machines, and the results are described in Chapter 6.

5.1.6 The Induction of 'Stress'

Some measure of the effect of 'stress' on performance was considered desirable in order to determine the robustness of the acquired skill to the performance of non-related activities. 'Stress' is a term covering a variety of phenomena (Section A4.4.2), and, in the present studies, it was taken to mean merely the potential cause of a deterioration in the operator's performance of the main task, not induced by the performance of other physical or mental skills. It was noted in the initial informal experiments that telling the operator that his performance was being tested caused a different approach to the task, for example, a different posture, a look of concentration, deeper breathing, and general indications of anxiety. It was, therefore, assumed that knowledge of the occurrence of performance evaluation was in itself stressful, and might be detrimental to performance. This assumption was brone out by the comments of the RAF pilots in the main trials, who were very concerned to know when they were under test, and commented on the 'fairness' of the stated test procedures.

In order to induce this particular form of stress in a uniform and controlled manner, the operators were given explicit instructions stating when the performance evaluation would take place; it was, in fact, continuous, but this is irrelevant to the stress induced. The overall experimental technique was to allow the operator a 'learning' phase in which he was unaware that his performance was being monitored, and then to give him instructions stating that he was to be tested, and evaluate his performance again. Any difference in performance immediately before and after the instructions is clearly due to activities in the intervening period, none of which was performing the task and one of which was assimilating the stress-inducing instructions. Although this procedure was designed to measure the effects of a form of 'stress', this is clearly not necessarily related to any other form of stress- however, the results of the operational procedure described are an indication of the robustness of the acquired skill.

5.1.7 Verbalization and Instructions

Verbalization and associated thought processes play an important part in the learning of skilled tasks, even though the final control policy may be essentially non-verbalizable. Even when tracking with the conventional joystick control, many operators gave a running commentary on what they were doing, why they were doing it, and especially about what the spot of light on the CRT screen 'intended' to do next. This verbalization was even more apparent with the reversing push-buttons where there is clearly a cognitive, or problem-solving, element in determining the relationship between control actions and their effects on the display. It is probable that such a component necessarily plays a part in any 'structured skill' (Pask 1965), since the inter-relationship between sub-skills is a higher -order function, or 'meta-language' (Pask 1965*), that the relationships between variables within a single skill.

In order to evaluate the effects and importance of verbalization, informal experiments were carried out with a variety of operators using the reversing push-buttons and adaptive trainer described in Section 5.1.4. They were asked to comment on their control strategy as they attempted to learn the tracking task, and their comments were noted. In these experiments, a large-screen (12 inch) oscilloscope was used as a display, and the levels of difficulty attained are not comparable with those in the formal experiments described later. Unless otherwise noted, the operators in the formal experiments were not told anything about the push-buttons, but were asked to use them to hold the spot on the oscilloscope in a region centred on a marked mid-line. The following is a brief description of some of the results which most influenced the main experimental design.

<u>Operator A</u> (Electronic Engineer) Curve A_1 of Figure 5-1 shows the variation of difficulty with time for an operator who learnt the skill rapidly, and to a high level. He remarked after the experiment, 'After about five minutes I suddently managed to stop pushing the opposite switch when the one I pressed was wrong'; a sudden change in ability at this time is apparent from the curve A_1. A_2 is a trajectory for the same operator several hours later, with no intervening practice. It may be seen that it rises rapidly to the previous maximum level, showing retention of the skill.

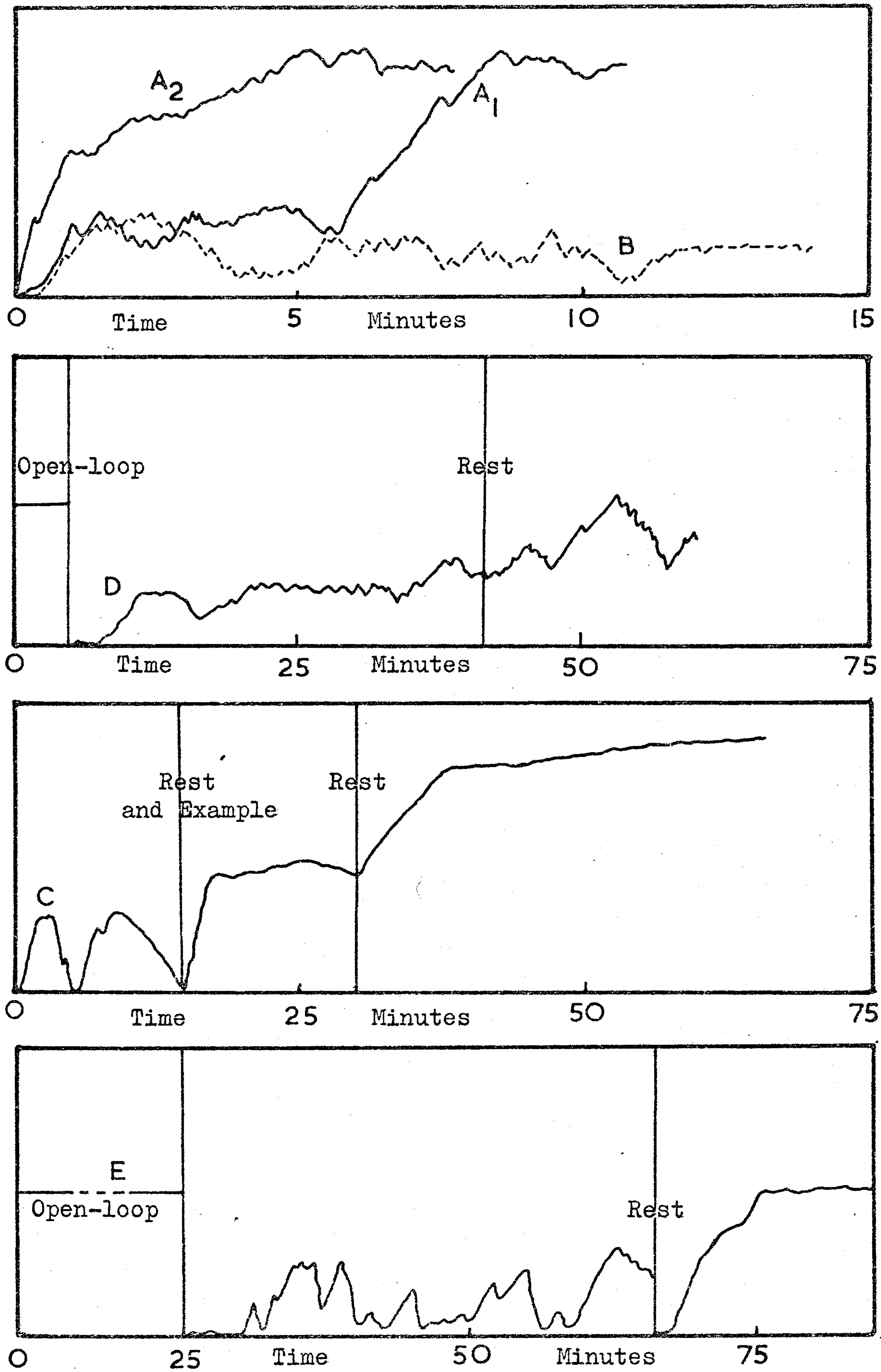


Figure 5-1 Feedback Training of Human Operators

Operator B (Psychologist) Curve B of Figure 5-1 is most noticeable for the lack of any real learning, although it shows a definite rise in the first few minutes. After the experiment, the operator remarked that he had first formulated the hypothesis that pressing one button gave alternate steps, whilst pressing one after another gave steps in the same direction: this is correct, and explains the initial rise. However, he said that this hypothesis had proved to be incorrect, and he had not been able to determine how the push-buttons operated. From further discussion, it became apparent that he was unaware that there was a disturbance that moved the spot independently of the controls. This is zero initially, and only becomes sufficient to overcome the draft when δ is about 0.2. When the disturbance became sufficient to reverse the expected direction of motion when he pushed the controls, this 'refuted' his hypothesis.

Operator C (Mathematician) Curve C of Figure 5-1 provides one of the most fascinating insights into the cognitive aspects of learning the skill. The operator immediately pushed both push-buttons at a very high rate and very wildly. After some 5 minutes he graduated to a strategy in which he pushed the buttons rapidly and at 'random' until the spot was in the centre region, and then left them alone. This approach produces some degree of control, and took him up to $\delta=0.35$, thence slowly declining; the change in the smoothing of the trajectory after seven minutes is due to an increase in the time-constant of the training loop. After fifteen minutes, the operator rested and stated that he had the impression that the push-buttons did one thing when the spot was in one place on the screen, and another when it was in a different place. He had tried various hypotheses as to the nature of this positional relationship, but none had proved correct.

The operator then asked the experimenter what the push-buttons did, and he replied, 'watch me tracking', and gave the buttons ten pushes, keeping the spot in the centre. The operator immediately took over the buttons, and within three minutes had attained a difficulty level of $\delta=0.5$. After a further fifteen minute tracking session, during which the level $\delta>0.5$ was maintained, the operator rested for five minutes and then tracked for a further thirty-six minutes. The level of difficulty was set at δ=0.5 at the start of this final run, and an immediate rise to δ=0.77 is apparent. After eleven minutes (total time on graph forty-one minutes), a rise to δ=0.87 may be seen - on seeing the charts afterwards, the operator remarked that this was when he realized that the 'pendulum' had to be moved in the mean (that is, the centre of oscillation of the spot was what he was controlling), and that the energy of oscillation was increased by pulsing at the end of a swing, and decreased by pulsing at the other. At the end of this very long training period, the operator was still keen to continue tracking.

Operator D (RAF Trainer) As shown in curve D of Figure 5-1, this operator was trained at the level, δ=0.5, for six minutes initially. At the end of this period he was asked what the buttons do, and replied that the right-hand button stops the spot when coming from the right and the left-hand button stops the spot when it is coming from the left, but neither has any effect in the centre of the screen. The operator then tracked for thirty seven minutes with the feedback trainer operative, attaining a maximum level of δ=0.4. During the rest period it became clear that he was now aware of the manner of operation of the reversing push-buttons, but he still talked of the capability of stopping the spot rather than being able to return it to the centre. After a further fifteen minutes tracking, during which the level of difficulty rose to a maximum of δ=0.62, he said that it was possible to control the spot very easily when it was in the centre, but the best strategy when it went to the edge was to leave it until it returned; this policy probably accounts for the jagged nature of his difficulty trajectory.

Operator E (Electronic Engineer) This operator trained for twenty-five minutes at δ=0.5, as shown in curve E of Figure 5-1, and at the end of this period had no idea at all of how the push-buttons operated, professing complete ignorance. Without any information about the push-buttons, he tracked for a further forty-two minutes under feedback training conditions, taking seven minutes even to move the spot into the centre region, and never reaching a level of difficulty above δ=0.32. He was still unable to state what the push-buttons did, and this was explained to him by the experimenter, after which, in a further fifteen minutes tracking, he performed the task reasonably well and attained a level of difficulty, δ=0.5.

Discussion From these informal experiments, and many others of the same nature, it was clear that the third-order tracking task with reversing push-button controls had the required features for use in the experimental comparison of feedback and other training techniques. Operators were able to track for periods of twenty-five minutes or more without experience, or complaint, of fatigue. All operators started with a total inability to perform what was a new and strange task for everyone, but it was possible to attain a high level of ability after as short a period as five minutes of feedback training, but more typically after twenty minutes. Training for extended periods at high levels of difficulty produced little or no learning - from the results with operators D and E, it might even produce negative transfer, possibly because of the type of behaviour shown by operator B - and yet these levels of difficulty were readily attainable under feedback conditions.

Most interestingly, from these initial experiments it was clear that verbal instructions could exert a strong influence over the learning of an operator, and that verbalization was a major effect in the learning. It was decided to investigate possible interactions between the effects on learning of the form of instructions given and the training technique used, by giving two different forms of instruction, one of which gave no information about how the push-buttons worked, and the other of which gave a great deal if information; these are described in the following section. It was also decided to attempt to evaluate the operator's knowledge of the task and his degree of verbalization by giving appropriate questionnaires after training; these are described in Section 5.1.9.

5.1.8 Forms of Instruction

There appear to be three basic forms of instruction which might be used to help the operator: firstly, those which describe to him the nature of the system he is to control - that is, if you do this then this will happen; secondly, those which advise him on a suitable control policy - that is, if this happens then do this; and thirdly, those which inform him of sub-goals to be attained - if you are able to achieve this then it will be useful in performing the task. It is assumed, of course, that instructions as to the main objectives and the 'rules of the game' are always necessary. It was not possible to include all these variations, and the first type was chosen because the form of the instructions is simple and obvious, being purely descriptive of the system and not requiring any knowledge of human control strategies.

Hence, only two forms of instruction were used in the experiment: the weak instructions, telling the operator nothing about the task he is to perform, except the performance criterion and the controls to be used; and the strong instructions, telling him, in addition, the nature of the coding of the push-button controls. The instructions were given to the operator on one side of a foolscap page at the start of the experiment, and he was asked to read them throughly. The actual form of the instructions was as follows:

R.A.F., ******* Medical Psychology 1966

Introduction In order to investigate training techniques for various skills it is necessary to use both a range of subjects, from those professionally involved in similar skills to those who may never have attempted them before, and also a range of skills, some of which must be novel for all subjects. The tasks to be performed will be presented to different subjects in different ways as part of the investigation. The particular skills you will be asked to perform all involve keeping a spot of light in the centre region of a display, using either a rolling ball joystick or a pair of push-buttons.

This is the background to this study. For the results to be valid we have to rely on your co-operation both in performing the tasks as well as possible, and in answering questions about them.

TASK I The spot of light on the display moves from side to side only, and your task to to maintain it in the centre of the display (marked by the centre black line), not deviating outside the black lines on either side of the centre line. If the spot comes to the edge of the screen it will not disappear, but should rest there so that you can see it.

(The following paragraph was used only in the weak instructions)

The red push-buttons on the arms of your chair are to be used as controls. You may find their effect puzzling at first, but part of your task is to learn what they do and this is not very complicated.

(The following paragraph was used only in the strong instructions)

The red push-buttons on the arms of your chair are to be used as controls. Depressing either push-button imparts an impulsive movement to the spot of light. At any instant one of the push-buttons is capable of knocking the spot to the left, and the other one is capable of knocking it to the right. Neither button consistently gives a left or right impulse, however, but instead they alternate in their effects each time you press one. The effects of the push-buttons may be puzzling at first, but part of your task is to learn how to use them.

(The remaining three paragraphs terminated both)

If it is not possible to maintain the spot of light always within the black markers then you should try and control it so that its average position is in the centre - that is, so that the spot deviates equally to right and left without any tendency to be more one way than the other.

The red light will come on to indicate that an experiment is in progress. If at any time whilst the light is on you wish to stop tracking please inform the operator (who can hear you through the intercom). He will lock the apparatus until you are ready to continue and there will be no need to repeat the earlier stage.

Please read through again if you wish.

-*-*-*-*-*-*-*-*-*-*-*-*-

Each of the three main experimental groups, H,L and F (Section 5.1.4), was subdivided into two groups with weak or strong instructions , labelled w and s respectively. Thus there were six experimental groups altogether, with eight operators in each of the groups Hw and Hs, twelve in each of the groups Lw and Ls, and sixteen operators in each of the groups Fw and Fs.

5.1.9 Form of Questionnaires

For all operators, the training period was divided into two sessions of twenty-five minutes, after each of which the operator was required to fill in a questionnaire. The prime objective of these was to provide some measure of the individual operator's attitude to the experimental situation, and some measure of his verbal reaction to the control problem. Auxiliary objectives were to require the operator to read the instructions again at the end of the first training period, and to give an interesting but relaxing task in between experimental trials. The questionnaires were based on the types of comment, and topics of interest, which had become apparent during the informal training sessions outlined in Section 5.1.7, and, wherever possible, obtained data in a quantitative rather than qualitative form. No time limit was placed on the filling in of questionnaires, and this varied widely between operators.

Figures 5-2 and 5-3 show the questionnaire which was administered at the end of the first 25 minute training session. The first question asks for the experimental instructions to be read through again, and requests comments on them to ensure that this is done. When these instructions were first read, the operator had no experience of the tracking task, and was unable to gain any whilst reading them since the equipment was inactive. Hence, it was felt that the instructions might have little effect, and that it was desirable to give them to the operator again; in the present experimental design, it is not possible to separate out the effect of instructions before, and after, the initial training session, but this is a very interesting possibility for future experiments.

The second question in Figure 5-2 requests an estimate of the initial training period - it was felt that this might reflect the level of stress, or involvement, of the operator in the task. In further questions, the operator is required to respond by marking a position on a line, ten centimetres in length, the two ends of which correspond to different extremal responses. All operators marked the lines without query and without apparent difficulty, so that this form of answer appears to be acceptable. The various questions seek to evaluate interest in the task, its apparent difficulty, the operator's estimate of his present performance, and of his potential future performance.

The first question in Figure 5-3 attempts to evaluate the operator's estimate of his own ultimate potential, and of the possibility of ever performing the task according to the instructions. It is interesting to note at this point, that this question was badly filled in, with many operators omitting one figure. This contrasts with the 10cm lines, which were completed by all 72 operators, showing the advantage of requiring questions to be answered in this way. The final questions on Figure 5-3 attempt to evaluate the extent to which the operator is able to solve the control problem in verbal form.

Figure 5-4 shows the questionnaire administered after the second training period. The first three questions attempt to discover the degree

This questionaire is to help us in evaluating the training situation. Please add any comments necessary if the given answers to questions are not adequate.

Please read through the instructions given to you at the start of this task. Were they adequate or should further instructions have been given ?

ANY COMMENTS:

Please estimate how many minutes you have been performing the task ....MINS.

Was the task itself interesting or boring ?- please indicate by marking an appropriate position on this line between the two extremes:

Very Boring ———————————————— Very Interesting

ANY COMMENTS:

Was the task itself too difficult or too easy to learn and perform ?

Far too Difficult ————————|———————— Far too Easy

Just Right

ANY COMMENTS:

How well do you feel you were performing the task finally ?

Complete Failure ———————————————— Perfectly

At the end of another practice run of the same length how well do you estimate you could perform the task ?

Complete Failure ———————————————— Perfectly

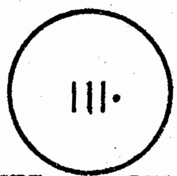
Figure 5-2 First Questionnaire - Part I

How many further practice runs would you need to perform the task - adequately ? .... perfectly ? ....

ANY COMMENTS:

What effect do the push-buttons have on the display ?

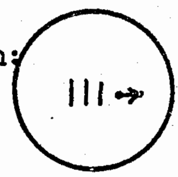
The spot is stationary as shown:



which push-button would you press ? LEFT RIGHT

ANY COMMENTS:

Having pushed the button the spot moves as shown:



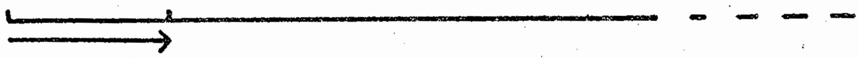
how would you press the push-buttons to bring it back to the centre ?

ANY FINAL COMMENTS:

Figure 5-3 First Questionnaire - Part II

You have now received two training runs on this task. Were they too long or too short? LONG JUST RIGHT SHORT

Please mark the length of time which would be most suitable for these practice runs:



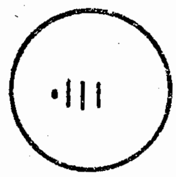
Present length

Is the experimental set-up itself in any way uncomfortable or fatiguing, and if so what improvements might be made ?

How many further training runs would you require to perform the task - adequately ? ...... perfectly ? ....

Have you any further comments about the effects of the push-buttons on the display ?

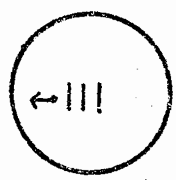
The spot is stationary as shown:



which push-button would you press ? LEFT RIGHT

ANY COMMENTS:

When you push the button the spot moves as shown:



how would you press the push-buttons to bring it back to the centre ?

ANY FINAL COMMENTS:

Figure 5-4 Second Questionnaire

of stress, or discomfort to the operator resulting from performance of the task. The remaining questions are similar to those on the first questionnaire, and investigate estimation of ultimate performance and knowledge of the control policy.

5.1.10 Summary of Experimental Design

Seventy-two RAF pilots were trained in the push-button tracking task described in Section 5.1.4, using the reversing push-button controls described in Section A4.3.5 and the display and experimental environment described in Section 4.5.3. Three forms of training regime were used, as described in Section 5.1.4, and two forms of instructions, as described in Section 5.1.8, giving six experimental groups whose constitution is summarized in Table 5-2.

Training Regime	Instructions	
	w - Weak uninformative	s - Strong informative
H - High Difficulty $\delta=0.5$ (H)	Hw - 8	Hs - 8
L - Low Difficulty $\delta=0.25$ (L)	Lw - 12	Ls - 12
F - Feedback δ variable to maintain performance constant	Fw - 16	Fs - 16

Table 5-2 Numbers of Operators in Experimental Groups

All operators had the same schedule of training, testing, answering questionnaires, and so on, and this is summarized in Table 5-3.

Activity	Duration	Description
Read Instructions	Variable (7-18 min.)	Instructions reproduced in Section 5.1.8 - weak and strong forms.
Train_1	20 min.	Track under one of three conditions, H,L,F; the feedback group started with $\delta=0.00$.
Test_1 $\delta=0.5$(H)	5 min.	Continue tracking without interruptio[n] and without knowledge of change, but with $\delta=0.5$, and performance measured.
Questionnaire_1	Variable (18-53 min.)	Fill in Questionnaire shown in Figure[s] 5-2 and 5-3; this involves reading instructions again.

$Train_2$		20 min.	Track under same conditions as $Train_1$; the feedback group started with δ at the same level as it was at end of $Train_1$.
$Test_2$	δ=0.5(H)	5 min.	As for $Test_1$.
$Questionnaire_2$		Variable (12-31 min.)	Fill in Questionnaire of Figure 5-4; informed that there will be three tests, each of five minutes.
$Test_3$	δ=0.5(H)	5 min.	Measure performance over last 4 minutes of test; this is at same level of difficulty as previous, unannounced test.
$Test_4$	δ=0.25(L)	5 min.	Same as $Test_3$ but at level of difficulty as that at which L group trained.
$Test_5$	δ=0.70(V)	5 min.	Same as $Test_3$ but at higher level of difficulty than that met by any operators, except a few of F group, during training.

Table 5-3 Experimental Schedule

5.2 Experimental Results

An experiment of this size and nature generates a great deal of data which may have within it not only the answers to the questions originally posed, but also indications of new phenomena, possibly more important than those which it was intended to examine. Hence it is desirable to present the results in as detailed a form as possible, without losing track of overall trends in a mass of data. This has been attempted by giving the raw data in numerical form in Appendix 5, displaying it in graphical form in this chapter, and giving statistics of the data in the appendix - the major effects are then clearly visible, and any peculiarities of distribution may be seen, whilst their magnitude may be checked either from the (parametric) statistics given, or by manipulation of the raw data.

5.2.1 Learning Behaviour

The interaction between the feedback trainer and the trainee is itself of interest for the 32 operators under the Feedback condition, and a complete set of trajectories of difficulty against time for these operators is given in Appendix 5. For convenience, these results are plotted in pairs, but the pairing is based on clarity of presentation only and does not reflect any properties of the data. Graphs 1 through 16 are those of the Fs group, and 17 through 32 are those of the Fw group;

these results can be linked to the raw data through the last column of this data.

A wide variety of possible behaviour is apparent in these graphs, from the steady climb to a high level of 13, through the rapid climb to a plateau of 7, the two-plateau characteristics of 8, the oscillatory uncertainty of learning of 30, and the rise and fall of 18 and 22. Some effects can be related to the experimental situation - the central discontinuities in 4 and 8 are related to the pause given for filling in the questionnaire, and, possibly, to the re-reading of the instructions - but many of the differences between the curves, particularly in relative smoothness and timing of rises and falls, appear to be data of a significant nature apparent only in the trajectories, and suggest that profile-matching could be applied to obtain more information from feedback 'trainers' used as tests.

5.2.2 Performance on the Tests

Figures 5-5 and 5-6 show bar graphs in which the performance of each of the 72 operators on a particular test is shown as a horizontal line at the appropriate ordinate. The bars are grouped in six columns corresponding to the various training conditions, and these charts illustrate the performance differences between groups induced by different training conditions. The significance of these differences may be determined from Table A5-2 of Appendix 5, which gives the mean and variance for each group on each test, together with t-statistics and variance ratios for the comparisons of means and distributions of the various groups; values which attain a one per cent level of significance are bracketed for ease of interpretation.

Figure 5-5(a) shows that, on a test at the high level of difficulty (H: δ=0.5) session, the Hw, Hs, and Lw groups show a uniformly low level of performance, whilst the Ls, Fw and Fs groups show a spread of performance ranging from the very low to very high; only the Fs group is significantly better than the first three groups, however. The spread of the two groups, Hw and Hs is significantly less than all the other groups, and this may be related to the sigmoidal nature of plots of performance against difficulty.

From Figure 5-5(b) it appears that, at the end of the second training session, these differences have been enhanced, and all three groups, Ls,

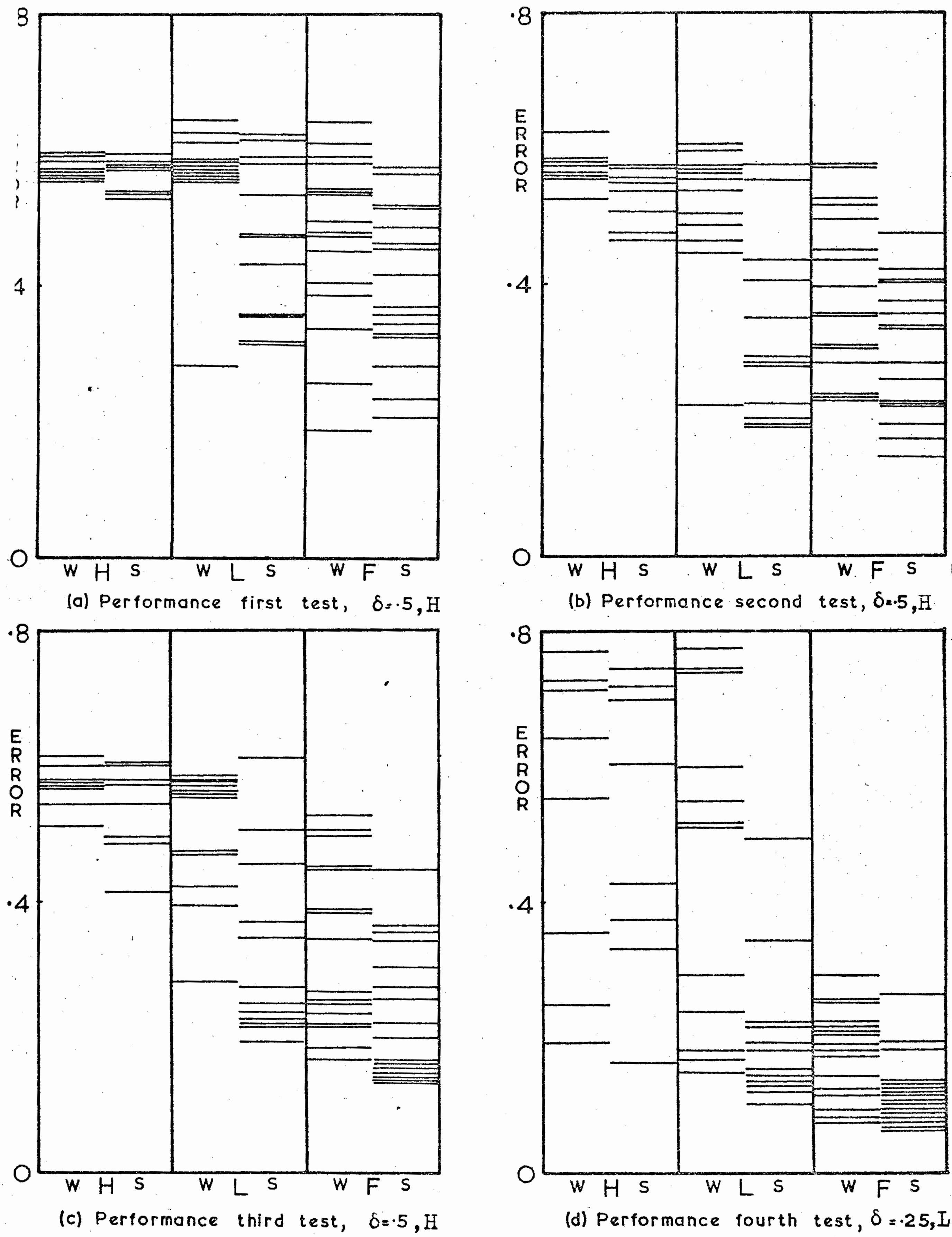


Figure 5-5 Performance on Tests

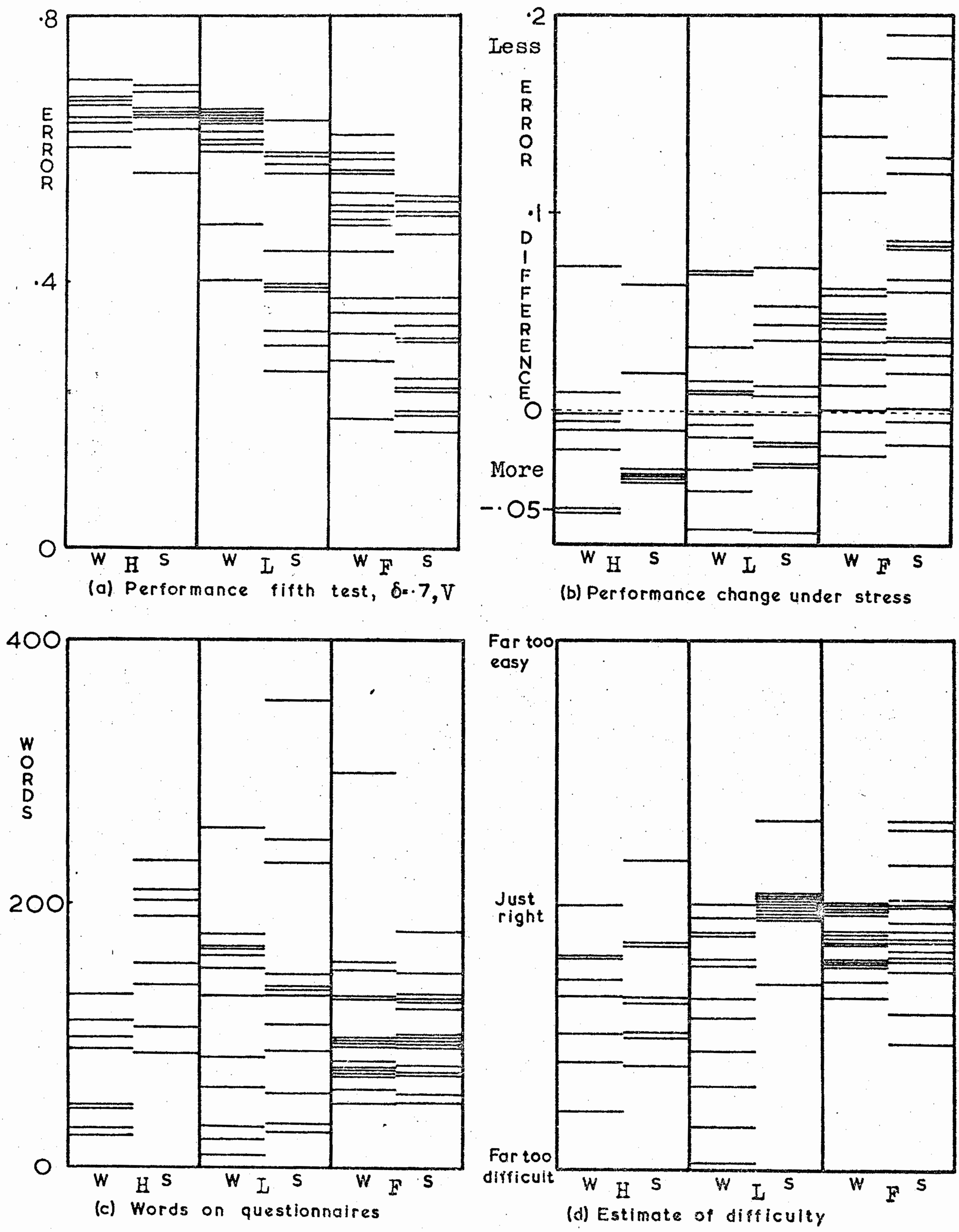


Figure 5-6 Performance on Tests and Questionnaire Results

Fw and Fs, are now better than Hw, Hs, and Lw, but there are no significant differences within each set of three groups. However, the graph is suggestive of an overall difference between groups with strong and groups with weak instructions. Figure 5-5(c), which shows the performance on a test at the same level of difficulty after the operators were informed of the test, has the same interpretation as Figure 5-5(b). However, the variable of greatest interest is the relationship between these figures and the effect of instructions on performance - this is shown by the plot of performance differences in Figure 5-6(b) and analysed in Section 5.2.4.

Figure 5-5(d) shows the performances on a test of lower difficulty (L: δ=0.25), and in this the Hw, Hs, and Lw groups again appear as not significantly different, the Ls and Fw groups are significantly better than these three, and the Fs group is significantly better than the other five. It is interesting to note the wide spread in learning of the Hw, Hs and Lw groups, particularly since the Lw group is being tested at the same level as that at which it trained. The high performance of some members of the H group on this test is due, in some part, to learning during the five-minute test period. The test results at a higher level of difficulty (V: δ=0.7), shown in Figure 5-6(a), demonstrate the spread in abilities which still exists in the better groups.

5.2.3 Effect of Instructions

The effect of giving informative (strong) instructions, containing a description of the operation of the complex controls, compared with that of giving uninformative (weak) instructions, was a pronounced improvement in performance, significant in all but the high-difficulty (H) group. The effect is by far the most pronounced in the group, L, trained at a low level of difficulty, in which there is a clear dichotomy of performance according to the instructions given. It is reasonable to suppose that, at this level of difficulty, a control policy sufficient to maintain the desired sub-environment could be set up and applied verbally - the operator had time to think. The effect is less apparent in the group trained at a high level of difficulty, H, who learnt uniformly badly, and the group trained under feedback conditions, F, who learnt uniformly well.

Another interesting feature of the effect of instructions is that it is more pronounced in the group undergoing feedback training, F, at the end of the second training session than at the end of the first. It had seemed reasonable to predict that the instructions would be of most benefit to the naive operator, and it is clear that the effect cannot be explained, for this group, by the sigmoidal nature of performance curves. It appears, however, from the comments of the operators, that many of them could not comprehend the instructions at first reading, whereas, after some experience in tracking, the instructions were very helpful. This may be partially due to poor instructions but is also an indication that an optimum interplay between direct communication and feedback training is required, and suggests that best results will be obtained with a system in which the instructions are under the control of the training system and can themselves be made contingent on performance feedback.

5.2.4 The Effect of Instruction-Induced 'Stress'

Figure 5-6(b) shows, for each operator, the error on the third test minus that on the second, and hence a positive 'error difference' corresponds to an improvement of performance. Since the third test is at the same level of difficulty as the second test (H: δ=0.5), and follows it after an interval with no practice at the tracking task, any error difference must be due to the effect of events in the intervening interval. During this interval the operator filled in the second questionnaire, and was then informed that his proficiency was to be tested. As discussed in Section 5.1.6, this information was expected to be stress-inducing, and hence, possibly, to cause a deterioration in the operator's performance. Alone, however, the interval of other activity might be expected to lead to an improvement in performance.

From Figure 5-6(b), it may be seen that the effect of the instructions varies widely over the three groups: out of the sixteen operators trained at a high level of difficulty, twelve show a deterioration in performance: the group trained at a low level of difficulty split equally into twelve who get worse and twelve who improve; out of the thirty-two operators trained under feedback conditions, only four show any deterioration, and the general performance improvement is very marked. From Table A5-2, only the improved performance of the F group over the H and L groups is significant at the 1% level; no effects of the main instructions are apparent.

As noted in Section 5.1.6, the effect of the information that the operator is to be tested is not necessarily one of 'stress' and it is in any event too superficial a conclusion to state that the performance of the operators trained under feedback conditions improved with 'stress', whilst that of operators trained under open-loop conditions deteriorated or remained unchanged. For example, the mean level of performance of each of the groups differs widely, and the nonlinearity of the performance scale magnifies changes at the mid-level of performance and minimizes the apparent extent of those at very high, or very low, levels. However, taking account of this effect only increases the contrast between the three groups, since the deterioration of the H group would be more pronounced, as would the improvement of the F group.

The most reasonable explanation of the overall effect of instruction-induced 'stress' is that the feedback group had spare capacity in test two, or had become fatigued through controlling at the high level of difficulty many of them had attained, and were able, after the instructions or a rest, to produce a higher standard of performance; the group trained at a high level of difficulty had learnt little and became highly stressed when asked to apply this learning; and the group trained at a low level of difficulty either show a mixture of both types of behaviour, or a random spread in performance. The circumstances of test two are anomalous for this last group, L, because it was probably apparent to them at the end of the training interval that the task had become more difficult. This might have induced them to use all their available capacity in the, supposedly unknown, test, and hence show no improvement in performance when informed of the test.

5.2.5 Responses to the Questionnaires

The marking of the ten-centimetre lines of the questionnaires was carried out by all operators without question or comment, whereas the response to questions requiring a written answer was poor, answers often being completely omitted. Because of the variety, both in quantity and nature, or the written responses, comparisons between the groups at a semantic level are not possible. However, the total number of words written by each operator on the questionnaires was evaluated to give an indication of the degree of verbalization, if not its nature. The time estimate was uniformly filled in, and this was recorded.

There is no significant difference between the groups in their estimates of the actual time of the training sessions, which is about five minutes less than the true time. However, the estimated optimal training time varies widely between the groups, especially in the degree of within-group agreement. The Hs group, trained at a high level of difficulty with informative instructions, request a rather shorter training session, and are the only group in which the optimal length is less than the estimated actual length; the high variances of the Hw and Lw groups are largely due to single individuals putting down very high values.

The interest in the tracking task which is indicated does not vary widely between the groups, although that of the Ls group is greatest and significantly more than that of the Lw and Hs groups. This uniformity of interest suggests that the differences in performance which were obtained were not a function of the relative motivations, or degree of boredom, of the groups under different conditions. Performance estimates again do not vary as widely between the groups as might be expected. Those of the Hw and Hs groups are lower, than the others, but by no means in proportion to actual performances; this reflects the 'adaption level' effect in performance evaluation, since no absolute standard is given to each operator.

The estimates of task difficulty show interesting differences between the groups, apparent in Figure 5-6(d) - as expected, the Hw, Hs and Lw groups, all of whom performed badly, find the task too difficult, but there is a remarkable consensus of opinion in the Ls group, emphasized by the availability on this particular scale of a centre point marked 'just right'. The total number of words written on the questionnaires also brings out an interesting difference between the groups, in that the Hs group wrote over twice as many words as the Hw group. It may be noted from Figure 5-6(c) that the Hs group has no individual writing less than about ninety words, which is very much higher than the minima of the other groups. This seems to reflect the unique status of the Hs group, who were told how to do the task and then found they could not in practice - a situation apparently creating much verbal behaviour.

The results obtained with the questionnaires are interesting and throw some light on the effects of the different training situations on the motivation, comfort and verbalization of the operators. Much more precise information could have been obtained if an automated questionnaire system with data-logging facilities for response times, such as that described by Gedye and Miller (1969), had been available. Such a system would also be valuable in enabling the instructions to be presented to a controlled level of comprehension.

5.2.6 Differences Between the Experimental Groups

The group, H, trained at a high level of difficulty (H: δ =0.5), show virtually zero learning compared with the other groups. At the end of the second training session, the sub-group, Hs, with informative instructions show better performance than Hw (significant at 5 per cent level). The level of difficulty, H: δ=0.5, is not in itself too high for successful learning and performance, however, since 65 per cent of the feedback group attained it, or much higher levels, during training. The Hs group, in particular, show interesting verbal behaviour, both in requesting significantly shorter training sessions, and writing significantly more on the questionnaire than the Hw group, presumably because they find the tracking task unexpectedly impossible, using the verbal instructions alone. In the easiest test, $Test_4$ (L: δ=0.25), the H group show a very wide spread of performance; those who did well showed appreciable learning during the test.

The group, L, trained at a low level of difficulty, L: δ=0.25, split clearly according to the instructions given - those with the weak, non-informative instructions do not show appreciably better performance than the group trained at a high level of difficulty, whereas those with strong, informative instructions show a spread in performance from very high to very low throughout the tests, but are comparable in performance to the group under feedback training. The Ls group stand out as expressing the greatest interest in the task and estimating that its difficulty was 'just right'.

The group, F, trained under feedback conditions in which δ was adjusted to maintain their mean error constant, again split according to the instructions given, but not in nearly so dramatic a manner as the L group. Both Fw and Fs groups learn to a high level of performance, and are significantly better than the Hw, Hs and Fw groups on all tests. The Fs group is significantly better than the Ls group on the fourth test (L: δ=0.25), which is particularly interesting since this is the level at which the Ls group trained. There is no significant difference between the Fw and Ls groups on any of the tests, and indeed the Ls group is slightly better in three out of the five. However, under instruction-induced stress, both the F groups show significantly better results than the Ls group, and, of course, all other groups.

5.3 Implications of Experimental Results

The interpretation of the results in terms of transfer from a training condition to an easier, or more difficult, test condition is very interesting. The Hs group shows little learning, and hence little transfer, to tasks either easier or more difficult, whereas the Ls group shows good transfer to more difficult tasks. In particular, the results of the fifth test,V: ς =0.7, show that training on an easier task leads to poor transfer, whereas training on a very much easier task leads to good transfer - no theory in terms of relative difficulty can account for this result. As noted in Section A4.5.1, Gibbs (1951) expresses his conclusions on transfer of training in terms of learning, 'carried on until the total possible skill is approached in both tasks'. This was not done in the present experiment, and it is possible that ultimately the H group might have learnt the task. However, it is clear that they would take very much longer to do so, and that no practical importance attaches to laws of training expressed in these terms unless predictions are also made about the rates of learning.

The utility of feedback training is best examined by considering separately the groups under w and s conditions of instruction. With the non-informative instructions, w, the interaction between learning how to control the system and learning how the system operates, the dual control problem (Chapter 3), is predominant and the sub-environment phenomenom may be expected to strongly influence learning. This is strongly borne out by the experimental results in that the Fw group, under feedback training, show overwhelmingly better performance at all test levels of difficulty, than either of the Hw and Lw groups under open-loop training. Thus, in a situation where the task is complex and poorly defined, and where interactions between performance and gaining knowledge may be expected, the experimental results clearly demonstrate the predicted advantages of feedback training.

With the informative instructions, s, the operator has the possibility of overcoming the sub-environment phenomenom by setting up a control policy 'verbally', and initially taking a cognitive approach to the perceptual-motor tracking skill. This will only be possible if the level of performance required of him is not too high. Comparison of the results for the Hs, Ls and Fs groups gain shows a significant advantage to feedback training, but now the L group is more similar to the F group than to the H group. This interaction between the effects of verbal instruction and the level of difficulty in training is, perhaps, the most important outcome of the present series of experiments. It not only provides experimental evidence of the meaningfulness and applicability of the approach to problems in learning advanced in Chapter 3, but is also relevant to the practical instruction situation in flying and driving, where verbal instruction and variation in task difficulty are closely combined.

The demonstration of an interaction between verbal instruction and the various modes of training, and its relationship to the sub-environment phenomenom, is , in particular, a vindication of the approach taken to the study of learning and training by Pask (1960,1961,1964, 1965,1965*), who has emphasized the importance of language in learning, and the linguistic nature of all processes in the learning hierarchy. The theoretical developments in the first half of this thesis indicate that feedback training will be most effective when there are complex interactions between the 'sub-skills' required for the learning of a particular task, and it is these interactions which are most amenable to description through language - thus, it is no coincidence that both feedback training and verbal instructions exert a profound influence on learning in the experimental situation chosen.

The experimental situation is itself of interest in that the use of reversing push-buttons to control a high-order system provides a task new to all operators, and which is learnt in about thirty minutes by operators under one training regime but not learnt at all by those under another. Although the task is clearly artificial, it involves an interaction between learning to use the controls and learning to control the system which is found, from one cause or another, in most skilled tasks for which training is required. Thus, the task provides an interesting and useful addition to the repertoire of laboratory situations for the investigation of human skills, their learning and training.

In the following chapter, experiments with adaptive-threshold-logic controllers, paralleling those with human operators, show that the results obtained are not unique to human learning, but are found with other forms of learning system, and, hence, are a function of the learning situation. In Chapter 7, possible extensions of the present experiments to other situations are discussed, together with the relevance of the results to practical training problems.

CHAPTER 6 : EXPERIMENTS WITH LEARNING MACHINES

6.1 Introduction

It has been noted several times that automatic adaptive controllers may be used as 'subjects' in experiments on learning and training, and that the results may not only illustrate fundamental phenomena of learning, but also be of direct relevance to the learning behaviour of human operators in similar situations. The advantages of using automatic controllers in this way, apart from the obvious ones of availability and experimental convenience, are that an ensemble of identical machines may be used to compare the effects of different training regimes, and that the reasons for particular behaviour shown by a machine may be investigated in detail by examination of the internal behaviour of the machine.

The choice of adaptive controllers is already wide and grows with the increasing number of machines being described in the literature.At one extreme are the linear controllers with parameters varied by cross-correlation (Donalson and Kishi 1965), whose behaviour is amenable to detailed theoretical analysis but which show only a limited repertoire of adaptive phenomena, and at the other extreme are multi-strategy, hierarchical learning systems, such as STeLLA (Andreae and Cashin 1969), whose behaviour defies prediction and shows a complex variety of adaptive reactions to the environment. Between these two extremes are pattern-classifiers and adaptive controllers based on adaptive theshold logic elements (ATLEs), whose basic structure is simple and amenable to analysis (Appendix 1), but whose behaviour can range over the full repertoire of adaptive phenomena described in previous chapters.

Two studies of computer-simulated learning systems are reported in this chapter: the first demonstrates the richness of behaviour possible for even a very simple adaptive system, and exemplifies the modes of adaption and phenomena of training discussed in Chapters 2 and 3; the second utilizes an ATLE controller as a range of subjects for the feed-back trainer discussed in Chapters 4 and 5, and compares the learning behaviour with that of human operators.

6.2 Adaptive Behaviour of an ATLE Pattern-Classifier

In the theoretical discussion much emphasis has been placed upon the inherent complexity or 'richness' of adaptive behaviour, and it is useful to chose as an experimental system with which to illustrate some modes of adaptive behaviour and training a very simple adaptive-threshold logic pattern-classifier (ATLE). The nature and behaviour of ATLEs is discussed in detail in Appendix 1, where it is shown (A1.3.3) that an ATLE with bounded weights does not necessarily converge to a solution of a pattern-classification problem, even when the Novikoff conditions are satisfied. Its convergence is dependent upon the initial values of the weights, that is its initial state, and, in terms of the discussion of Section A3.5, there is more than one 0-minimal ideal in which its state may ultimately reside. In Section 2.2.5, a pattern-classifier of this type has been used to examplify the concept of a task, and in this section the example is developed in more detail through experimental studies.

Using the notation of Section A1.3.3, consider the ATLE with five weights, W_i $1 \leq i \leq 5$, which are bounded in the range from -4 to +4, so that:

$$-4 \leq W_i \leq +4 \quad , \quad 1 \leq i \leq 5 \qquad [6.1]$$

and consider the set of stimulus vectors:

$$\begin{array}{lll}
A = (1, 1, 1, 1, -1) & & A' = (-1,-1,-1,-1, 1) \\
B = (1,-1,-1,-1, 1) & & B' = (-1, 1, 1, 1,-1) \\
C = (-1, 1,-1,-1, 1) & & C' = (1,-1, 1, 1,-1) \\
D = (-1,-1, 1,-1, 1) & & D' = (1, 1,-1, 1,-1) \\
E = (-1,-1,-1, 1, 1) & & E' = (1, 1, 1,-1,-1)
\end{array}$$

The left-hand set of vectors may be separated from the right-hand set by the weight vector, $W \equiv (1,1,1,1,3)$, since -

$$\begin{array}{l}
W.A = W.B = W.C = W.D = W.E = 1 > +1/2 \\
W.A' = W.B' = W.C' = W.D' = W.E' = -1 < -1/2 \qquad [6.2]
\end{array}$$

Consider now an ATLE pattern-classifier using the decision and adaptive procedures of Section A1.3. Because A=-A', B=-B', and so on, it is unnecessary to take the two sets of patterns separately, and training sequences may be regarded as made up of A,B,C,D and E, only. In terms of the definition of Section 2.2.5, let the sequence of patterns, $t \equiv (E,A,D,C,A,B)$, be a 'task' for which it is required to train the pattern-classifier. The effect of giving the classifier this task once, starting with a weight-vector (0,0,0,0,0), may be calculated as in Section A1.3.3:

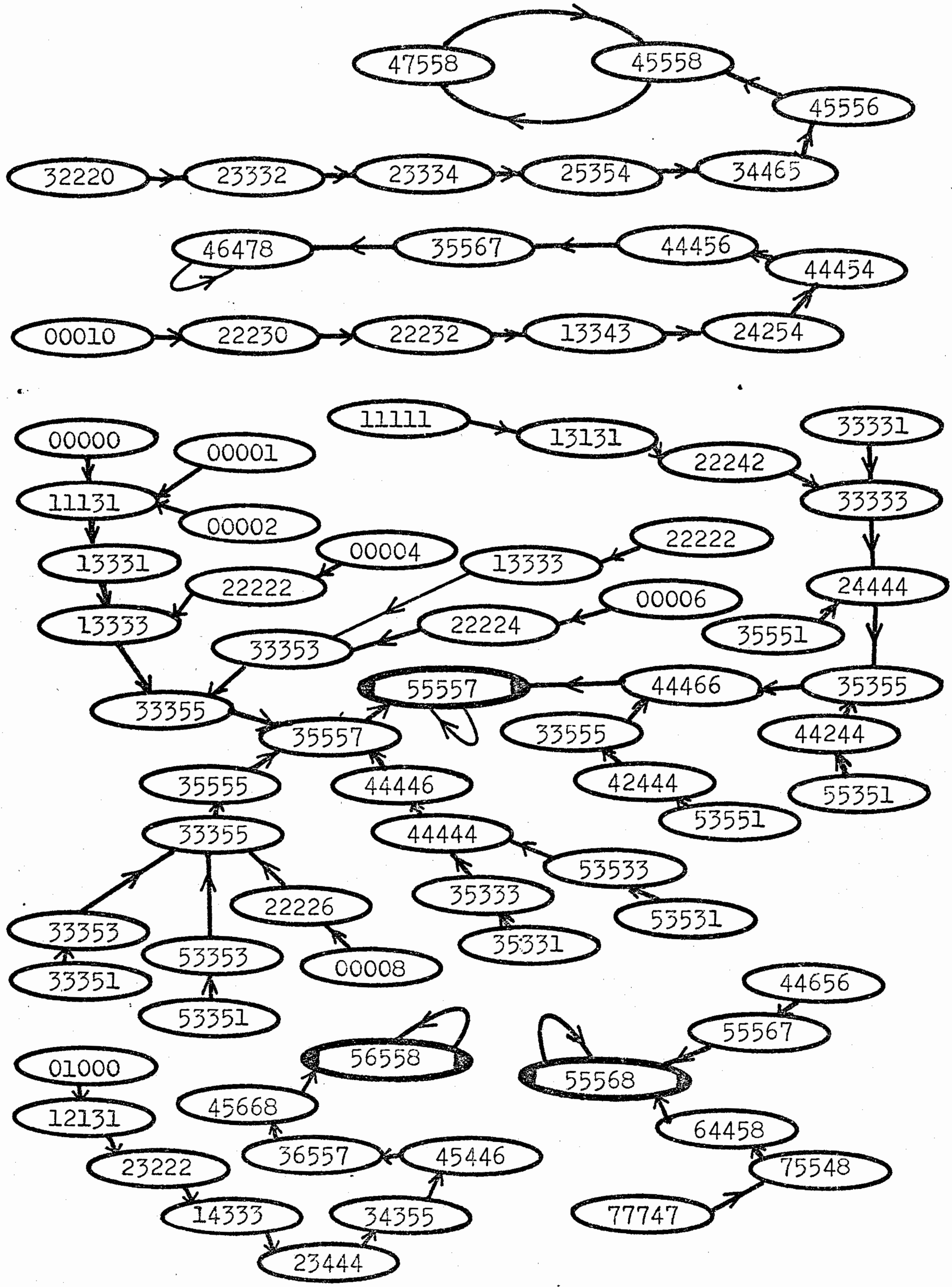
			W_1	W_2	W_3	W_4	W_5	Correct
$W_{Initial}$ =	W(0)	B	0	0	0	0	0	No
	W(1)	A	1	-1	-1	-1	1	No
	W(2)	C	2	0	0	0	0	No
	W(3)	D	1	1	-1	-1	1	No
	W(4)	A	0	0	0	-2	2	No
	W(5)	E	1	1	1	-1	1	No
W_{Final} =	W(6)		0	0	0	0	2	

Hence, giving the pattern-classifier the task, t, changes its state from the initial weight-vector, (0,0,0,0,0) to the final weight-vector, (0,0,0,0,2). This is still not a solution to the problem, but it may be shown that repeating the task another two times leads to convergence to the solution, (1,1,1,1,2). Hence, the pattern-classifier is potentially adaptive to the task, t, when its state is given by the weight-vector, (0,0,0,0,0). However, given a different initial weight-vector, such as (0,0,0,1,0), the pattern-classifier does not necessarily converge, and shows limit-cycle behaviour, as in the example of Section A1.3.3.

By plotting out the state-sequences of the pattern-classifier, given the task, t, in every possible state, its adaption-automaton may be completely identified, and experiments may be carried out on compatible adaption, open-loop training, and so on. However, even this very simple adaptive system has 9^5 = 59,049 states, which makes it a major computational problem to examine the structure of the complete adaption-automaton in practice, and only particular parts of the transition diagram can be mapped out. It is convenient in doing this to simplify the nomenclature for states of the automaton by adding 4 to each component of the weight-vector and writing the result as a string of digits - thus, (0,0,0,0,0)≡ 44444, and (-1, -2, 4,-4,0) ≡ 32804.

6.2.1 Adaption-Automaton of the Pattern-Classifier

Figure 6-1 shows some trajectories induced by the task, t, in the state-space of the adaption-automaton of the ATLE. There are five possible states in which the classifier has attained a solution to the problem, 55557, 55568, 55658, 56558, 65558, and the first, second and fourth of these are shown in the figure marked in heavy rings. By the nature of the error-correcting adaptive procedure, once these states



<u>Figure 6-1 State-Transitions of Adaption-Automaton</u>

are attained the weights are not changed, and these states are invariant under the task, t. The centre part of the figure shows part of the tree of states converging in to the final state, 55557. This is clearly an ideal of the state-semigroup of the automaton, and the monogenic sub-semigroup generated by the state, 55557, is the only 0-minimal ideal contained within it. The states 56558 and 55568, shown in the lower part of the figure, also generate 0-minimal ideals, and the centre and lower parts of the figure clearly contain states which are within the region of potential adaption of the automaton.

In the upper part of the figure, are shown state sequences which do not lead to solutions, such as that terminating in the state, 46478, which, although it is invariant and generates a 0-minimal ideal, is not a solution to the pattern-classifiecation problem. Because it is possible for weight-vectors which are not solutions to change under t, the interesting behaviour shown in the topmost part of the figure is possible- states 47558 and 45558 together form a cycle, and the automaton alternates between them.

6.2.2 Modes of Adaption

The transitions in state-space of Figure 6-1 are for the single task, t, only - by considering also the state-transitions introduced by other tasks, illustrations may be given of all the modes of adaption defined in Chapter 2 and 3. For example, consider the set of vectors obtained by interchinging the first and last components of the vectors, A, B, C, D, E, A', and so on, defined in Section 6.2 - let these be A_1, B_1, C_1, and so one, so that, for example, $A_1 \equiv (-1, 1, 1, 1, 1)$ and $E_1 \equiv (-1, 1, 1, -1, 1)$. The new sets of vectors can clearly be separated by the weight vector, $W_1 \equiv (3,1,1,1,1)$, since -

$$
\begin{aligned}
W_1.A_1 = W_1.B_1 = W_1.C_1 = W_1.D_1 = W_1.E_1 = 1 > +1/2 \\
W_1.A_1' = W_1.B_1' = W_1.C_1' = W_1.D_1' = W_1.E_1' + -1 < +1/2
\end{aligned}
\qquad 6.3
$$

The training sequence, s $\equiv(E_1,A_1,D_1,A_1,B_1)$, bears the same relationship to the new sets of vectors as the sequence, t, to the old ones, and hence the effect of performing s on the adaption-automaton of the pattern-classifier may be derived from Figure 6-1 by interchainging the first and last digits of each state label.

Figure 6-2 shows a fragment of the state-space of the automaton, generated by taking the sequence in Figure 6-1 commencing with 11111 and

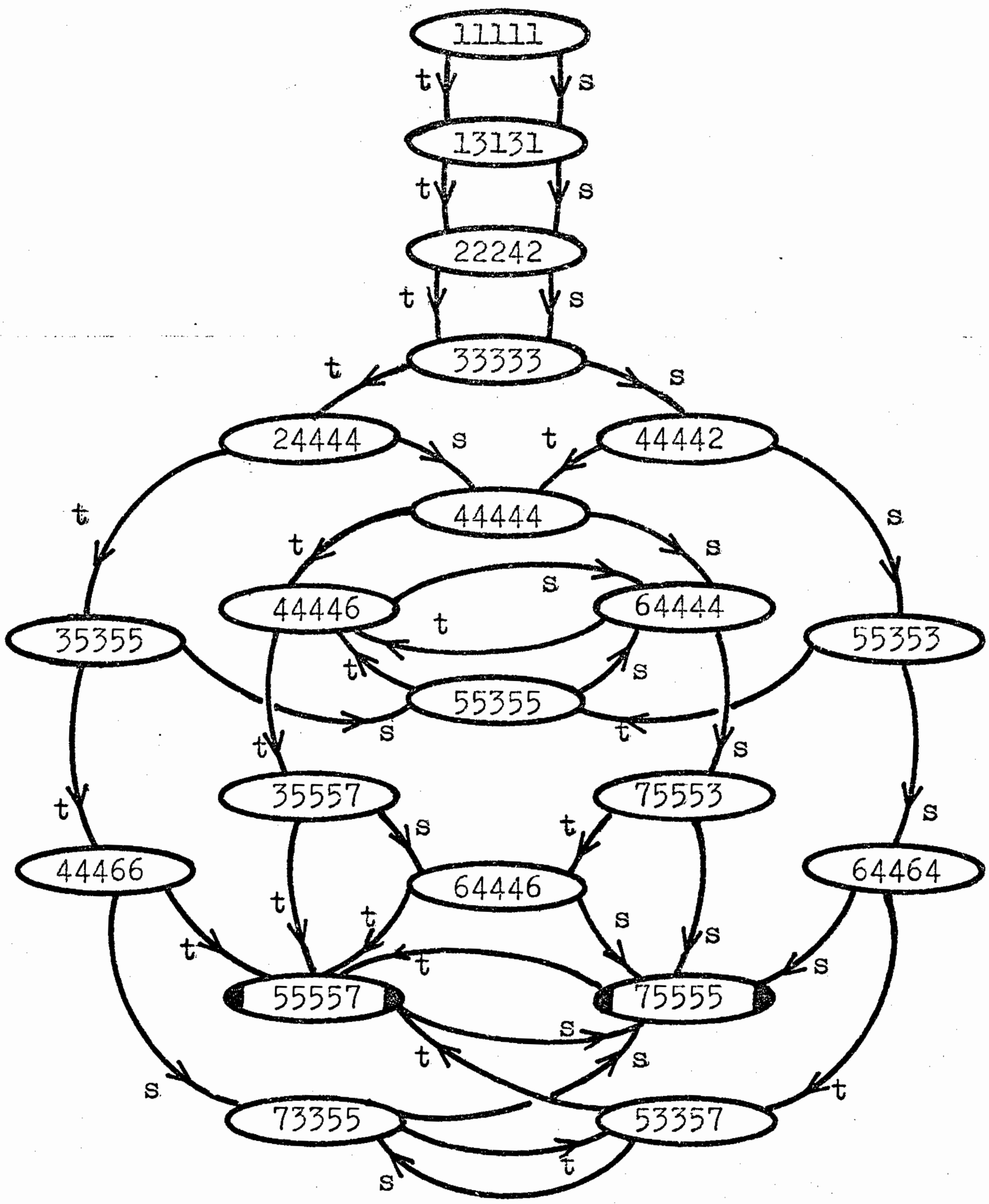


Figure 6 -2 Compatible Adaption for Two Tasks

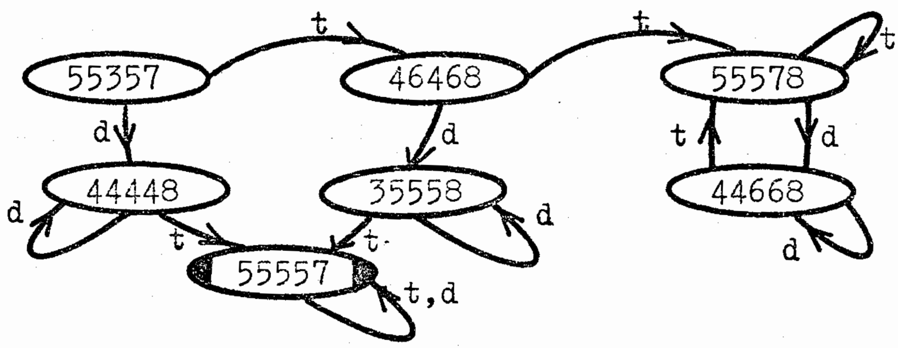


Figure 6 -3 Application of Training Sequence terminating in 55557, and drawing in also the transitions induced by the task, s. Initially the tasks both induce the same transitions, but then the state sequences diverge, and a task sequence of the form t^n leads to convergence, for that task, in the state, 55557, whereas a task sequence of the form s^n leads to convergence, for that task , in the state, 75555. The state diagram of Figure 6-2 has many interesting features, particularly the interchange relationships between 44446 and 64444, 73355 and 53357, and 55557 and 755555. The last two states are those corresponding to solutions, and there is a single-step transition from one solution to another, showing that the automaton is compatibly adapted to t with respect to s in 55557, and vice versa in 75555.

It is clear that all the states shown in Figure 6-2 are within the compatibly adaptive region for the set of tasks, (s,t), in that any sequence from the free semigroup generated by s and t leads to a state from which 55557 can be reached under the action of t^n, and 75555 can be reached under the action of s^n. The similarity in action of s and t is due to the selection of tasks both requiring positive weight vectors - if instead the exact opposite dichotomy to that required by t were selected, so that A, B, C, D and E were assigned to the negative class, then the solution weight vector would be $W_2 \equiv (-1, -1, -1, -1, -3) \equiv 33331$. It may be seen from Figure 6-1 that there is a sequence under t leading from 33331 to 75557, and vice versa, so that the automaton is compatibly adapted with respect to t in the state, 33331. In this case, however, the two tasks will clearly tend to induce state transitions in opposite directions.

The phenomenom of joint adaption may be investigated by considering the effect of a different sequence of the tasks constituting t - let r ≡ (E,D,C,B,A), so that whenever the automaton is adapted to t it is also adapted to r, and hence jointly adapted to the set of tasks, (r,t). Even though r is similar to t, however, and only has one stimulus less, learning with r takes many more task performances than learning with t - for example, starting from 11111 the sequence induced by r^n is:-

11111 → 21111 → 12202 → 21113 → 32222 → 30224 → 12224

→ 23315 → 23335 → 34226 → 54226 → 43337 → 44446 → 55337

→ 44448 → 55557 , a fifteen-step, rather than a seven-step, learning sequence.

6.2.3 Training

In Figure 6-2, s^n may be regarded as an open-loop training sequence for the pattern-classifier learning to perform the task, t. Since s never induces a transition directly into 55557, none of the states shown is within the conditional adaption set, A(t: u), where u=s say, defined in Section 6.2.3 as one of complete open-loop trainability. However, all the states shown are within P(t:u), the set of potential open-loop trainability. It is interesting to note that from some states s^n gives more rapid training than t^n - for example, from, 55353 s^2t is better than t^4. It is also interesting that, whilst s^n and t^n are both effective training sequences, $(ts)^n$ or $(st)^n$ are not, both leading to trapping in 44446 $\rightleftarrows$ 64444. One might say that training under one regime lays the foundations for further training under that regime, and that, although the ultimate goals are similar, the routes taken under the two regimes are different.

In Figure 6-2, the use of s^n as a training sequence for t is interesting but not important since the classifier is always potentially adaptive to t anyway. Figure 6-3 shows a sequence of transitions under t starting from the state 55357 which is outside the region of potential adaptivity to t, so that, given t^n the classifier does not converge but instead becomes trapped in the state 55578. The open-loop training sequence considered is that generated by giving the stimulus, D, alone, so that the task d $\equiv$ (D). From the first two states in the sequence, 55357 and 46468, d^n induces trajectories into the potentially adaptive region for t. From the final state, 55578, the trajectory induced by d^n terminates at 44668 which is not within the potentially adaptive region for t. One might say that training on the task sequence d^n is necessary for learning, but it must be given early in the learning of the main task, t^n, if it is to be effective.

6.2.4 Conclusions to be Drawn from Pattern-Classifier Experiments

It is clear that experiments with simple learning systems, such as ATLE classifiers may be used to illustrate the various phenomena of learning and training previously defined, but it is not clear whether such experiments have any significance in studying learning behaviour and establishing new 'laws' of learning. One obvious conclusion is that very simple systems, such as a device operating upon five 'weights' each taking nine values can have very complex behaviour, so complex in fact that it is impossible to study it in detail completely.

Secondly, approaches to the analysis of behaviour which seem plausible and reasonable for the human operator should also appear plausible and reasonable for simple learning systems. In Section 6.2.3, at the end of each paragraph, the results obtained have been re-stated in broader, behavioural terms. Each operational definition of a phenomenom of human behaviour may be applied to artificial learning systems, and each statement of a 'law' of behaviour in terms of these definitions can be evaluated in experiments with artificial systems, for example, the possible 'law' of relative transfer between easy and difficult tasks. Not only does such an evaluation check that the definitions are truly operational, but also it may lead to a greater understanding of the basis for the proposed 'law'.

6.3 Adaptive Behaviour of an ATLE Controller

It was desired to parallel the experiments on the utility of the feedback trainer for a tracking task, described in Chapters 4 and 5, with similar experiments using adaptive controllers as trainees rather than human operators. Such experiments were expected to aid in the design of experimental situations for the human subjects, to enable the most sensitive evaluation of the utility of feedback training, and to provide an interesting comparison between human and machine learning. An ATLE controller was chosen as the learning system since it could be simply and rapidly simulated on a digital computer.

The experiments with this controller were carried out during the design stage of the experimental system for human operators (Section 5.1) and provided the data on which the levels of difficulty in the informal design experiments were based. In the following sections the ATLE controller is described and the experimental results with it are analysed in relation to their influence on the main experimental design and in comparison with the results of the human operator studies.

6.3.1 Description of ATLE Controller

An ATLE controller with the structure shown in Figure A1-2 and analysed in Section A1.4 was designed to act as the trainee 'learning machine' for the feedback trainer shown in Figure 4-4; the equations of the particular trainer used are given in Section 5.1.4. The inputs and outputs of the ATLE controller were constrained to accuracies and information rates roughly equivalent to those of the human operator in the same situation. The position and velocity of the spot on the oscilloscope were coarsely quantized, encoded into a binary pattern, and sampled at 200 milli-second intervals. A positive or negative impulse, u (Equation 5.1), was given at the output of the ATLE controller 100 milliseconds after the corresponding binary input was received.

A fifteen-bit binary pattern, $Y_i = \pm 1 \quad 0 \leq i \leq 14$, was generated at the input of the ATLE by thresholding the position and velocity, e and e respectively, each at seven levels; the remaining bit was permanently set. The threshold levels were chosen to cover the ranges of position and velocity, nominally $\pm$ 1.0, with maximum discrimination in the region about zero; they were, $\pm$ 0.6, $\pm$ 0.35, $\pm$ 1, and 0.0. The sign of the impulse at the output of the ATLE was determined in the usual way -

$$\mathrm{sgn}(u) = \mathrm{sgn}(w_0Y_0 + W_1Y_1 + W_2Y_2 + \ldots + W_{14}Y_{14}) \qquad [6.4]$$

where W_i are the weights of the ATLE;in the particular case when the right hand side of the equation was zero, the sign of u was taken to be positive.

Performance feedback to the ATLE to adjust the weights and hence adapt the control policy was the source of much difficulty, as described in Section A1.4.1. Various trials were carried out with possible performance feedback strategies, such as, for example, averaging the error over an interval and applying positive or negative bootstrapping over that interval according to whether the mean error was less than, or greater than, the mean error over the preceeding interval. The majority of strategies investigated did not lead to adaption to a reasonable control policy. Out of the remainder, the following was selected as a reasonable and successful procedure.

At any sampling instant, n, the input pattern, $Y_i(n)$ and the output 100 milliseconds later, u(n), were stored, together with the error at that sampling instant, e(n). k sampling instants later, this data was examined and if the error modulus had decreased the decision giving rise to u(n) was taken to be successful and the weights adjusted accordingly, otherwise the decision was taken to be unsuccessful and the weights were adjusted in the opposite direction. Thus if -

$$\Phi(n) = \mathrm{sgn}(\mathrm{MOD}(e(n)) - \mathrm{MOD}(e(n+k))) \qquad [6.5]$$

then
$$W_i(n+k) = (1-\eta)W_i(n+k-1) + \eta Y_i(n)\,\mathrm{sgn}(u(n))(1-\beta + \beta\Phi(n)) \qquad [6.6]$$

which is a standard ATLE convergence procedure for continuous, bounded weights, in which $1/\eta$ is the approximate time-constant of convergence - set at about 3,000 sampling instants in the experiments, and β determines the relative effects of reward and punishment - $(2\beta - 1)$ is the ratio of the magnitude of the weight change made when a decision is unsuccessful to that made when it is successful.

k and β are parameters of the ATLE controller which were adjusted in the experiment to give a family of controllers with different 'personalities'. Although this particular form of controller has a problem-determined performance feedback loop, and is not such a general form of controller as was originally hoped for, this is irrelevant to the results of the exepriments, in that these are concerned with the relative effects of different training regimes on the learning of given controller.

6.3.2 Control Policies of the ATLE Controller

The control polcies implemented by the ATLE controller are best described in a figure giving the locations within the (quantized) position/velocity phase plane in which positive or negative impulses will be emitted. If position is located on the horizontal axis, and velocity along the vertical axis, with the normal senses, then each axis is divided up into eight regions, and there are sixty-four cells in the phase plane. Representing a positive output by an asterisk,*, and a negative output by a dash, ', a control policy in which the output is entirely position dependent is -

```
****''''
****''''
****''''
****''''
****''''
****''''
****''''
****''''
```

A control policy with predictive velocity feedback would have more dashes in the upper left hand quadrant, and more asterisks in the lower right hand quadrant.

Every one hundred sampling instant (20 seconds tracking) the computer-simulated ATLE controller printed out its control policy in the form shown, so that its progress in learning could be evaluated not only in terms of its performance, level of difficulty attained and stabilization of the weight values, but also in terms of the type of policy it was implementing. A number of these policies were taken as fixed, non-adaptive controllers, and the value of α, the maximum level of difficulty at which they were able to maintain the mean error at the tolerated level of 0.34 units, was measured using the feedback trainer as described for relay controllers in Section 4.5.2; the mean error for each fixed controller at the standard test levels of difficulty, δ=0.25(L), δ=0.50(H) and δ=0.70(V) (Section 5.1.4), was also measured.

Table 6-1 shows the control policies, values of α, and values of $\bar{e}$ for the three levels of difficulty, for eighteen different policies having varying degrees of velocity feedback. P-01 corresponds to positional control only and leads to very poor performance - α is virtually zero. A minimal amount of velocity feedback in P-03 causes an increase in α to 0.33. A wide variety of other control policies, P-05 through P-13, including the maximal velocity feedback of P-13, lead to α between 0.5 and 0.6, whilst the maximum value of α found for this type of controller was α = 0.65 for P-18.

To some extent the quadruples of (α,L,H,V) may be compared with the corresponding quadruples for human operators in Appendix 5, Table A5-1 (subjects 41-72 on α_2, test results 4, 3 and 5, respectively - noting that the decimal point before the figures in A5-1 has been omitted). However, for a given value of any of the quadruple, comparisons within the human or machine groups reveal a wide range of possible values for the other members of the quadruple, as does a cross-comparison between humans and machines. It is certainly not possible to infer what types of control strategy the human operators were using, and the ATLE controller was not intended to be a model of the human controller at the control policy level, but rather a possible model at the comparitive difficulty of learning level; this is discussed further in Section 6.3.5

6.3.3 Experiments with the ATLE Controller

It has been noted in Section 6.3.1 that the majority of performance feedback strategies investigated for the ATLE controller did not lead to the learning of a reasonable control policy. Some lead to definite mal-adaption, and others to virtually positional control with a slight varying velocity component. In the latter case it was found that under adaptive training / took up a value between 0.2 and 0.3 (compared with mal-adaption where α oscillated between 0.0 and 0.05). The gradient of task difficulty with α in the range 0:0 to 0.25 seemed to be so slight that a minimal level of accomplishment corresponded to α about 0.25.

By adjusting k and β (Section 6.3.1) a range of values was found for which the ATLE controller learned under adaptive conditions to substantially higher values of α. Again the learning appeared dichotomous

****!!!! ****!!!! ****!!!! ****!!!! ****!!!! ****!!!! ****!!!! ****!!!!	!!!!!!!! !!!!!!!! !!*!!!!! *****!!! *******! ******** ******** ********	***!!!!! ****!!!! ****!!!! ****!!!! ****!!!! ****!!!! ****!!!! ****!!!!	***!!!!! ***!!!!! ****!!!! ****!!!! *****!!! *****!!! ******!! ******!!	***!!!!! ***!!!!! ****!!!! *****!!! ******!! ******!! *******! *******!	**!!!!!! ***!!!!! ***!!!!! ****!!!! *****!!! *****!!! ******!! ******!!
P-01	P-02	P-03	P-04	P-05	P-06
α 0.024	α 0.258	α 0.331	α 0.460	α 0.529	α 0.548
L 0.442	L 0.424	L 0.325	L 0.173	L 0.231	L 0.143
H 0.620	H 0.291	H 0.570	H 0.266	H 0.316	H 0.288
V 0.821	V 0.582	V 0.807	V 0.775	V 0.766	V 0.772
!!!!!!!! !!!!!!!! !!!!!!!! ***!!!!! *****!!! ******** ******** ********	!!!!!!!! !!!!!!!! !!!!!!!! ****!!!! ******!! ******** ******** ********	!!!!!!!! !!!!!!!! **!!!!!! ****!!!! ******!! ******** ******** ********	!!!!!!!! !!!!!!!! !*!!!!!! ****!!!! ******!! ******** ******** ********	!!!!!!!! !!!!!!!! ***!!!!! ****!!!! ******!! ******** ******** ********	!!!!!!!! !!!!!!!! !!!!!!!! ****!!!! ******!! *******! ******** ********
P-07	P-08	P-09	P-10	P-11	P-12
α 0.562	α 0.564	α 0.569	α 0.579	α 0.579	α 0.590
L 0.117	L 0.168	L 0.115	L 0.119	L 0.119	L 0.183
H 0.315	H 0.187	H 0.238	H 0.326	H 0.325	H 0.233
V 0.512	V 0.567	V 0.500	V 0.518	V 0.540	V 0.537
!!!!!!!! !!!!!!!! !!!!!!!! ****!!!! *****!!! ******** ******** ********	!!!!!!!! **!!!!!! ***!!!!! ****!!!! *****!!! *******! ******** ********	!!!!!!!! !!!!!!!! ***!!!!! ****!!!! *****!!! ******** ******** ********	!!!!!!!! ***!!!!! ***!!!!! ****!!!! *****!!! ******!! ******** ********	!!!!!!!! **!!!!!! ***!!!!! ****!!!! *****!!! ******!! ******** ********	!!!!!!!! !!!!!!!! ***!!!!! ****!!!! *****!!! *******! ******** ********
P-13	P-14	P-15	P-16	P-17	P-18
α 0.595	α 0.612	α 0.621	α 0.629	α 0.629	α 0.646
L 0.173	L 0.099	L 0.103	L 0.120	L 0.120	L 0.096
H 0.252	H 0.278	H 0.224	H 0.261	H 0.260	H 0.127
V 0.540	V 0.489	V 0.409	V 0.490	V 0.469	V 0.446

Table 6-1 Control Policies and Performances of ATLE Controllers

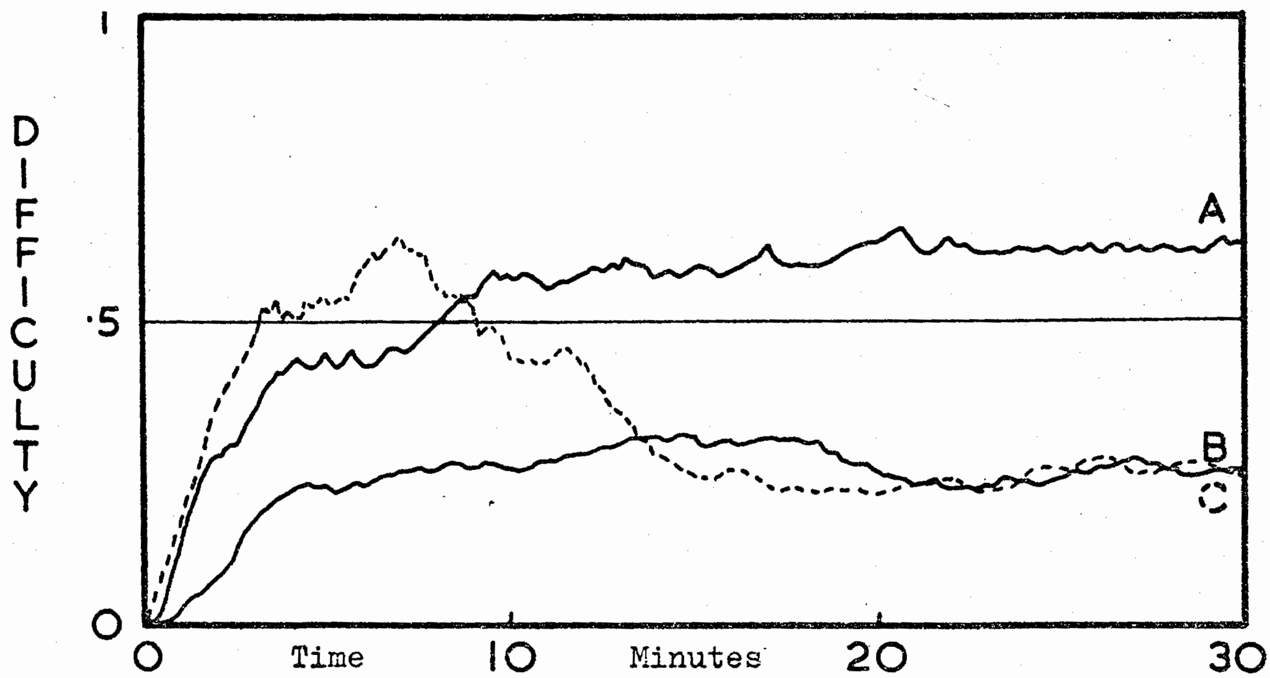


Figure 6-4 Feedback Training of ATLE Controllers

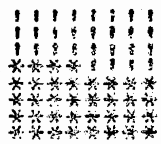
if a stable value of α was reached at the higher level it was in the range 0.55 to 0.65, and no controllers stabilized at levels of α between 0.3 and 0.5. The upper bound on α corresponded to a maximal velocity feedback policy such as P-18. It was found that this bound could be increased to about α=0.7 by increasing the number of levels at which the position and velocity were quantized for the ATLE. It could be increased substantially beyond this only by reducing the sampling interval from 200 milliseconds. However, the 7 levels of quantization on position and velocity and the 5 per second sampling rate has been chosen to be plausibly related to those of the human operator and were retained.

It was found that all the ATLE controllers which could learn to a high level of α had two stable final states of learning, one of which corresponded to a value of α about 0.25 and the other of which corresponded to a value of α about 0.6. Which of these two final states was attained was a function of the learning conditions. This was clearly what was required if the machines were to be used as indicators of the probable effect of different training strategies on the human operator - machines which never learnt or machines which always learnt were both useless for purposes of ascertaining the relative merits of training strategies.

As might be expected the learning behaviours of the ATLE controllers was considerably more stereotyped than that of the human operators shown in Appendix 5 (Figure A5-1). Figure 6-4 shows the three main forms of behaviour obtained with the feedback trainer plotted as task difficulty, δ (which from Equation 4.20 provides a lagging measure of α), against time. Machine A (k=4, β=0.625) learns rapidly to a high level, δ=0.63, and remains stably there implementing the control policy P-16; an extended experiment showed that this policy remains stable for at least another 30 minutes with no indication of any potential relapse. Machine B (k=4, β=0.6) rises to δ=0.3 but no further and finally stabilizes with δ=0.25 implementing a policy similar to P-03. Machine C shows a hybrid between the two behaviours, rapidly rising to δ=0.58 implementing P-15, but then gradually declining to δ=0.25 with a policy similar to P-03.

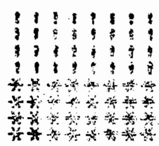
From these experiments and informal ones with human operators (Section 5.1.4) the levels of δ=0.25 and δ=0.5 were chosen as suitable for non-feedback, open-loop training at fixed difficulty. The lower level was chosen because it seemed relatively easy for controllers to attain a control policy at δ=0.25 which lead to good performance and hence the desired sub-environment. The higher level was chosen because it was possible for human and machine controllers to learn to that level under adaptive conditions and maintain a consistently good level of performance there. The two levels were well separated and provided quite distinct training and test conditions. For testing, a third level δ=0.70 was also used to provide a difficult task for even the best performers - however, it adds very little to the results of Chapter 5 and none of the learning systems investigated was capable of learning at this level of difficulty.

The ATLE controllers were tested under open-loop training conditions in which they learned at a fixed level of difficulty for the equivalent of 30 minutes and then were tested on the adaptive trainer. Machine A (k=4, β=0.625) was trained under open-loop conditions at fixed difficulties of δ=0.25 and δ=0.5 - in both cases it learnt quickly to a high level with a final policy -

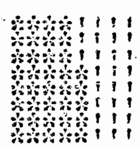


which is similar to P-08. However, training A on the feedback trainer starting at maximum level of difficulty caused it to stabilize with policy P-02 and a low level of performance. Another machine, D (k=5, β=0.6), had a virtually identical trajectory to A on the feedback trainer and attained a final level of difficulty, δ=0.63 with the control policy P-16. However, machine D was unable to learn to a high level under open-loop training conditions with δ=0.5, and stabilized at -

a policy similar to P-03. At δ=0.25, however, D learnt the policy P-08, which leads to a high level of performance. Further experiments with D showed that the level, δ=0.3, was a critical one for its learning, and at this level it learnt the velocity feedback policy -



which kept the error-rate low but exerted no positional control. It was clear that D must have two stable final policies under feedback conditions, one of which was P-16 - by trial and error the other policy was found to be that shown on the left, one similar to that learned under open-loop



training at δ=0.5. Machine C, which showed unstable learning under feedback conditions, also learnt to a high standard at δ=0.25, but not at δ=0.5; in this case there was only a single stable state of learning under feedback conditions, however, giving rise to the anomaly that, although the machine could maintain at δ=0.25 a stable policy which would enable it to perform well at δ=0.5, when the difficulty reached the higher level this policy deteriorated into one which would give satisfactory performance only at the lower level of difficulty.

There results with ATLE adaptive controllers are summarized in Section 6.3.5 where they are compared with those with human operators. For comparison purposes, however, some equivalent of the effect of instructions on human operators was desired and this is discussed in the following section.

6.3.4 Linguistic Interaction with ATLE Controllers

The major influence of the type of instructions given to human operators on their learning of the tracking task, and the interaction of this effect with that of the training regime, described in Chapter 5, made it desirable to investigate the possibility of such effects with the ATLE controllers. The main problem was clearly to decide what would constitute verbal instruction in their case, and how they might be expected to take note of it.

One approach to language which seems most fruitful in control situations is to consider the linguistic structure that must be set up to replace an existing, or hypothetical, physical link. For example, consider a controller which uses a 'fast model' of its environment to search for an optimum control sequence, such that the normal link between model and controller is broken and replaced by communication in a 'natural language'. The controller must be able to pose questions of the form, 'If the environment is in state * η* and I use control action * α *, what will the next state be.', and the model must be able to reply, 'The next state will be * μ*.'

This is language at a simple and apparently trivial level, but consider the similar, but more realistic, situation in which a controller exists, some form of model exists, and it is desired to have the one make use of the other. Assuming that some mechanism may be set up within the controller to enable it to use knowledge about the effects of control actions, a typical problem will be that the representation of the environment in the controller is very different from that in the model - for example, that it is unable to specify a state, * η*, but only an output, and that as an analogue variable rather than a digital pattern. Alternatively, the model may be incomplete and the controller must be able to use replies of the form, 'It may be * μ*, * π* or * λ*. ', or 'I do not know.' These examples illustrate the need for assumptions and typical vagueness of linguistic communication between unmatched, ill-assorted structures (such as human beings).

There is no obvious mechanism for causing the instructions given to human operators (Section 5.1.8) to affect the ATLE controllers. It would be adequate for the purposes of this chapter to assume that the instructions would cause the ATLE's to start with a useful control policy, such as P-03, and check whether learning takes place starting from this policy when it does not starting with the normal useless policy (which, from Section 6.3, is a policy consisting of all asterisks). However, some experiments were carried out with a simple mechanism for 'verbal' communication with the ATLE's, not identical to that with human operators (of the second type described in Section 5.1.8, rather than the first), which demonstrate that the concept of such communication can be made meaningful and operational.

The ATLE controller previously described was given the capability of accepting statements like, 'When the position of the spot on the oscilloscope is x and its velocity is v, then a sensible sign of control signal is c.', and using them to adjust its control policy accordingly. The controller 'imagines' the input-pattern it would receive resulting from x and v, considers that it has emitted the output c, and rewards itself for so doing.' This simple structure is readily extended to take account of non-quantitative specifications, 'When the spot is on the far left moving fast to the right...', and other qualifiers, 'It is very sensible....'. The overall effect of a message is to modify the ATLE controller's policy, or, initially, to prime it with a control policy. In the context of the experiments with human operators, it was of interest to discover whether this priming through instructions would enable a controller previously unable to learn a suitable policy to establish an initial sub-environment in which it could do so.

The weight changes of the ATLE are equivalent to adding in the stimulus vector if it should cause a positive output, and subtracting it if the output should be negative. Hence, given the instruction, 'If the spot is on the left, press the right-hand button', the ATLE would generate the stimulus vector - (-1, -1, -1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1), where the first seven components are positional information, assuming that e=0.1, the next seven components represent a lack of velocity information and the last component is always set. The output required is positive, and hence this stimulus vector becomes the weight vector. The corresponds to the control policy -

```
****!!!!
****!!!!
****!!!!
****!!!!
****!!!!
****!!!!
****!!!!
****!!!!
```

This is purely positional policy with no velocity feedback, as might be expected from the instructions, and is sensible, even though, from P-01 of Table 6-1, it is ineffective.

A variety of sets of initial instructions were experimented with, some of which gave rise to very powerful control policies - for example, the set: when the position is x and the velocity is v, the sign of the output should be c, for:-

(x,v,c) = (-0.4,0.3,-1), (-0.2,-0.3,1), (0.2,0.2,-1), (-0.5,-0.2,1) generates the weight vector, (0,0,-2,-2,-2,0,0,0,0,-4,-4,-4,0,0,0), which corresponds to the control policy -

```
!!!!!!!!
!!!!!!!!
!!!!!!!!
***!!!!!
*****!!!
********
********
********
```

This is P-07 of Table 6-1 and corresponds to good performance at δ=0.562, a very high level, and one adequate to ensure learning to a stable policy at this level of performance for machine D under open-loop training at δ=0.5, a level at which it was previously unable to learn.

Thus, instructions may be used with the ATLE controller to overcome their problems in learning. One outcome of the experiments on various sets of instructions, such as those in the last paragraph, was to demonstrate that the effect of instructions could not be determined in advance - there is no particular reason why the instructions given above should have such a good effect, and, indeed, on detailed examination they seem rather odd. In practice it was found that instructions could be used effectively by giving one, examining its effect on performance, and then selecting another - that is, 'telling' was ineffective (Lewis and Cook 1969), but instructions based on feedback as to their effect could be used to control behaviour.

6.3.5 Comparison of Human and ATLE Experiments

In comparing the experiments with human operators and ATLE controllers learning the tracking task of Chapters 4 and 5 it is important to make clear on what grounds the comparison is based, in particular at what level the ATLEs might be expected 'model' the human behaviour. The starting point for the study of learning controllers was -

given an arbitrarily chosen class of adaptive controllers with similar timing/accuracy constraints to the human operator (that is, in the sense of Section 2.2.1, the peripheral system of the human being was regarded as part of the environment), and the design problem of finding controllers in the class which would learn to perform well the tracking task given to human operators, do, the controllers found, show more than one stable state of adaption and is the relative effect of different training strategies on the state of adaption similar to that for human operators.

This position may be further clarified by considering some of the possibilities that might have arisen -

(i) No machines learn - bad choice of adaptive controllers - chose another or drop experiment.

(ii) Some machines learn and do so to much the same standard under all training conditions - bad choice of adaptive controller for modelling human learning behaviour - if at early stage of study might also have lead to change in tracking task or training conditions - at later stage, when human differential learning had been established, would have been of interest in showing that learning differences were not inherent in the task.

(iii) Machines learn at δ=0.25 (L) but not under feedback conditions, or, worse, learn at δ=0.5 (H) but not under feedback conditions - at an early stage this would have been taken as an indication of a bad feedback trainer- after the studies with human operators, it would be a difficult result for which to account.

None of these possibilities actually occurred, and the range of parameters covered in the experiments is such as to rule them out for the class of ATLE controllers investigated. Possibility (ii) might well occur with some classes of controllers - however, the ATLE with sampled, quantized inputs and a global learning strategy based on incremental weight changes and generalization were chosen to have those features of human learning most likely to be affected by the type of problem posed by the tracking task (system identification and predictive control with time delays in feedback). It is interesting to note that range of α with the ATLE controllers, 0.2 - 0.65, compares well with the range of values for human operators after 20 minutes tracking (α_1 of Table A5-1), 0.20 - 0.74; neither humans nor machines did markedly better or worse than one another in terms of absolute levels, indicating that the sampling and quantization constraints were reasonable.

If one were comparing two groups of human subjects it would be appropriate to make statistical comparisons between the results for the two populations. However, the composition of the 'population' of ATLE controllers under consideration is completely arbitrary and can be chosen at will. The four types of behaviour shown by machines A, B, C and D, described in Section 6.3.3 (together with that of complete mal-adaption) exhaust the range elicited from the ATLE controllers. However, A through D, are not just four 'subjects' but rather representatives of whole populations (obtained by variation of k and β) with closely similar behaviour. Hence, the question under consideration is whether each type of human learning behaviour is shown by one of A through D, and each type of behaviour shown by A through D is also shown by some human operators.

Machine A learnt well under the three training conditions (F,L,H) even though it had a stable state of poor performance (Section 6.3.3). This contrasts with the human operators in that none learnt under the H condition, so that at least one learning system showed a better learning capability than any of the human operators. Machine D better typified the human operators in that it learnt under the better conditions, F and L but not under H. Machine B learnt to a comparatively low level under all conditions which corresponds to a few of the human operators.

Machine C could attain an unstable state of learning under L and F, but always eventually sank back to a lower state. The only comparable results with human operators are Graphs 18 and 22 of Figure A2-1, although a rise to a high level and then a smooth progressive decline was found with one operator in the preliminary, informal experiments. It was ascribed at the time to fatigue, boredom, or some other such convenient psychological variable. In retrospect, because it is not so easy to dismiss a machine's behaviour in this way, such negative learning, or mal-adaption, appears of great importance. The learning machine did not suffer from muscular fatigue, neither did it become bored or lose concentration. One may only suppose that the changes in the sub-environment brought about by adaption of the control policy were such as to induce mal-adaption. In the human operator this phenomenom may be accompanied by complaints of boredom or fatigue, but these do not <u>explain</u> the mal-adaption.

Thus, machine D was the most appropriate to form the basis of an 'ensemble of identical machines' to evaluate the differential effects of training regimes (Section 6.1), and it was used in the experiments to determine the critical level of δ at which learning just failed to take place (δ=0.3, Section 6.3.3) and in the experiments on the effect of instructions (Section 6.3.4). In summary, the studies with ATLE controllers were of benefit in the design of the experimental system for human operators. They also gave rise to an adequate set of patterns of learning behaviour to account for the human operators who could learn the task under F and L conditions not being able to do so under H conditions. The overall results suggest that, since similar patterns were shown by humans and machines, the results obtained derived from the epistemological problems posed by the tracking task not from any particular human peculiarities in learning it. The 'fatigue' or 'boredom' of machine C, and the effect of 'instructions' on the learning of machine D, are of less weight, but illustrate possible extensions of the studies with learning systems to the modelling of other aspects of learning behaviour.

CHAPTER 7: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

7.1 Review of Objectives

The aims of the investigation and background to the objectives have been outlined in Sections 1.1 and 1.2. It is appropriate at this stage to review the objectives in the order in which the relevant results have been presented. The first objective has been to provide a rigorous foundation for studies of learning and training by developing a systematic account of the relations between behaviour, structure and purpose in arbitrary systems including men and machines. In Chapter 2 an axiomatic approach to the definition of adaptive behaviour has been established which enables operational and purely behavioural definitions to be provided of terms such as 'adaptive', and 'adapted'. In Appendix 3 the problem of deriving a structure which could give rise to observed behaviour has been analysed, and an algorithm established for constructing a minimal and observable structure cybernetically equivalent to an observed system. In the latter part of Chapter 2 these results have been used to define the 'adaption automaton' of a learning system, and to base a taxonomy of adaptive behaviour upon the properties of this automaton.

Thus, the first objective has been attained, and, in particular, the definitions of modes of adaption, the derivation of structure from behaviour in which all 'intervening variables' are measurable, and the analysis of problems in learning in terms of the 'sub-environment phenomenom' appear to break new ground and clarify difficult issues, both in animal psychology and in systems theory. The residual problems stem largely from the vast range of possible behavioural sequences generated by even a small set of descriptors, and the impossibility of empirical observation of all possible behaviours of any single adaptive system - a difficulty resulting not only from the amount of data and time taken to collect it, but also from the logical impossibility of causing an irreversible system to show all its possible behaviours. The complete resolution of these difficulties is impossible, but a practical resolution will result from the development of theories of approximation and incremental identification of general systems, and their application to adaptive systems.

The second objective has been to use these results to develop an integrated approach to the problems of training, in which a knowledge of the patterns of behaviour, structure and desired goals of a system may be used to formulate an optimal training strategy. In the first part of Chapter 3, the problem of training was formulated as one of controlling the adaption automaton of the trainee, and hence linked to the axiomatic system developed in Chapter 2. A theorem was then established which demonstrated that the necessary constraints on the adaption automaton to make training possible were also sufficient for the contruction of an effective feedback trainer. A second training theorem demonstrated that further constraints on the adaption automaton could lead to a more structured feedback trainer. In neither case did the training system require detailed knowledge of the structure of the adaption automaton of the trainee, and the trainer was seen to act as a stabilization system providing a suitable 'learning environment' rather than a detailed stimulus/response-based controller. This was also demonstrated in more abstract form in the latter part of Appendix 3 where the theory of adaption was related to that of the stability of general systems.

In the latter part of Chapter 3 the determination of constraints upon the adaption automaton of the trainee which will enable the training theorems to be applied has been related to an analysis of the epistemological problems of the trainee in attempting to control an environment whilst at the same time learning about it. An automata-theoretic statement of this problem was given, in which it was shown that any control policy restricts the environment to some sub-environment, and that the sub-environment generated by a naive controller may be unsuitable for learning. The basic training strategy was then formulated as maintenance of the sub-environment similar to that encountered by a controller which has learnt the problem. Hence, the second objective was attained in that various formulations for effective feedback trainers have been established based on the range of possible information about the trainee and the training problem.

The third objective has been to demonstrate the application of the theory to a realistic situation, and compare some of the theoretical predictions with experimental results. A high-order compensatory tracking task, related to the control of the longitudinal dynamics of aircraft, was chosen as an environment for the experiments on learning. In Chapter 4, a feedback training system was developed for this task using the hierarchical training structure of Chapter 3, and a theoretical and experimental analysis of its viability, in terms of overall behaviour and stability, was described. In Chapter 5, an experiment with human operators to determine the utility of this trainer described, in which various modes of training, fixed open-loop, and feedback, were compared, and the interaction of the mode of training with the form of instruction given was also evaluated. In Chapter.6, the same experiments were repeated with articial adaptive controllers in order to enable a comparative study to be made of human a nd machine learning.

The theoretically predicted advantages of feedback training were found, both with human operators and automatic controllers, and the effects of differing instructions were consistent with the hypothesized sub-environment phenomenom. Apart from replications of this type of experiment with other forms of task, the main directions for further research are the incorporation of the instructions within the feedback training loop, and the investigation of the applicability of feedback to real training situations, such as those of flight simulators. The present studies and experiments have not attempted to demonstrate that the concepts of adaption and training developed have application to the general range of human learning behaviour, cognitive as well as perceptual-motor skills, although the theoretical discussion has been carried out at a level of abstraction which suggests that this is so, and there is scope for major studies of the application of feedback training to cognitive skills.

7.2 Summary of Theoretical Results

An operational and purely behavioural approach to the study of adaption and learning may be established by considering the interaction between controller and environment to be segmented into a sequence of 'tasks', for each of which it is possible to say whether the interaction has, or has not, been satisfactory. The fundamental situation of an adaptive controller, to be coupled to a fixed environment and learn to control it satisfactorily, is then equivalent to the controller performing a sequence of tasks consisting of the same task repeated indefinitely, and, if its behaviour eventually becomes satisfactory and remains so, then it is said to be acceptable for the task. When the controller has reached this stable state of satisfactoriness, it is said to be adapted to the task.

Given these fundamental concepts, a variety of different modes of adaption may be distinguished when the controller may become involved in any of a set of tasks. If it is able to have an acceptable interaction with any one of the set, then it is said to be potentially adaptive to the set. In adapting to one, however, it may become unable to adapt to the others, and, hence, if this does not occur it is said to be compatibly adaptive to the set. If in adapting to one of the tasks the controller actually becomes adapted to all of them, then it is said to be jointly adaptive to the set.

In training a controller for a particular task, it may be given other tasks for which it is not required to be satisfactory, but which cause it to become adapted, or potentially adaptive, to the main task. According to the way in which the subsidiary tasks are selected, three modes of training may be distinguished: in fixed training, the controller is given only the main task, and reliance is placed on its being potentially adaptive to this task; in open-loop training, the trainee is given some training sequence of tasks before adapting to the main task, but this sequence is not varied for differences in trainees or states of learning; and, in feedback training, the trainee is given a sequence of tasks selected according to observations about its state, particularly those obtained from its performance.

In feedback training, the trainer has a control problem in taking the trainee from a state in which it is not potentially adaptive to the required task to one in which it is. These 'state' of adaption may be formally, and rigorously, defined by considering the observed interactions between controller and environment to be sequences of 'descriptors' each of which is defined by the task given and the satisfactoriness of the interaction. From the set of all possible sequences of descriptors, which may be said to define the adaptive system extensively through its behaviour, an automaton structure may be derived, the adaption-automaton, which shows the same behaviour. In particular, this structure may be chosen to be observable, in that a sufficient segment of past behaviour defines its present state, and to have a minimum number of states consistent with observability. The various modes of adaption correspond to differing forms of stability of the automaton, and training is a control-problem in the state-space of the automaton.

Although the adaption-automaton structure can, in theory, be derived from a complete set of descriptions of behaviour, in practice such a set cannot be observed for an irreversible system, and non-behavioural sources of information must be examined in order to identify the structure of the adaption-automaton. One source of such information comes from the examination of the epistemological problems of learning which are, to a large extent, independent of the nature of the learning system, and have to be faced by all controllers in the same environment with similar purposes. One major source of problems in learning is the interaction between the requirement to know how to control the environment in order to learn about it, and the requirement to know about an environment in order to control it. This interaction arises because each control policy of the learning systems restricts the environment to some sub-set of its states and state-transitions, or sub-environment, and the sub-environment generated by naive controller may be very different from that of a controller with a satisfactory policy, and learning in it may be irrelevant or even deletrious. Thus, one objective of a feedback training strategy may be to maintain the desired sub-environment.

7.3 Summary of Experimental Results

A feedback training system was developed for a third-order compensatory tracking task with dynamics consisting of an integration in cascade with a stable second-order transfer function. The damping-ratio of this latter, and the amplitude of the disturbing signal, determined the difficulty of the tracking task, and these were co-varied automatically to maintain the operator's mean error constant. The behaviour of this system, particularly its stability and speed of response, was analysed both theoretically and experimentally for non-adaptive relay controllers, and the results shown to be similar and acceptable, in that the loop behaviour was free of artifacts such as might occur from instability.

A modified version of this system, in which impulsive push-button controls which reversed polarity each time they were depressed were used to induce interactions between learning about the system and controlling it, was used in experiments to investigate the utility of feedback training. 72 operators, from a homogeneous population comprising RAF pilots at an advanced stage of selection and training, were trained under three conditions of difficulty, High, Low, or Feedback, and two forms of instruction, Weak or Strong. The High and Low difficulty groups were trained at fixed levels of difficulty and the Feedback group with the trainer. The Weak instructions gave no information about the operation of the controls, whilst the Strong instructions explained their nature. All operators were tested finally at three levels of difficulty , High Low, and Very High, and the High test was given twice in succession, the first time without the operators being informed, in order to test the effect of instruction-induced stress. All operators filled in questionnaires to evaluate their attitude to, and knowledge of, the task, and their degree of verbalization. Similar experiments were carried out with computer-simulated learning machines in order to determine whether the results were independent of the learning system, as theoretically predicted.

The main results of the experiments are as follows:-

(a) The operators trained at a High level of difficulty show little or no learning and do badly on all the tests. The Strong instructions have a significant effect in improving learning, but do not overcome the operator's basic difficulties. The High level of difficulty is not in itself unattainable, however, since over 65 per cent of the Feedback group reach it, or a much higher level, during training.

(b) The operators trained at the Low level of difficulty split clearly according to the instructions given - those with Weak instructions show little learning, whilst those with Strong instructions show a spread in performance from very good to very poor throughout the tests.

(c) The operators trained under Feedback conditions all learn to a high standard. Those with Weak instructions do not differ significantly from the group trained at a Low level with Strong instructions. The Feedback group with Strong instructions are significantly better than all other groups.

(d) The overall effect of Strong, or informative, instructions is to improve learning in all groups, but less markedly in the groups trained under the best or worst conditions. The clear-cut split in the group trained at Low difficulty demonstrates that good instructions may compensate, to some extent, for poor training conditions.

(e) The effect of instruction-induced stress is that operators trained at a High level of difficulty get worse, operators trained as a Low level do not vary appreciably, whereas operators trained under Feedback conditions get significantly better. This is the only difference in performance which differentiates the group trained at a Low level with Strong instructions, from those trained under Feedback with Weak instructions.

(f) The results with computer-simulated learning machines parallel those with human operators, in that the rank order of training conditions was the same for both, some machines could learn at Low difficulty or under Feedback, but not at High difficulty, and the effect of instructions could be to give an initial policy sufficient for learning to take place at High difficulty.

(g) Feedback training was significantly better than open-loop training for a given set of instructions. Its advantages were most pronounced when the instructions given were uninformative, and this is consistent with the supposition that the instructions aided the operators in establishing the desired sub-environment.

7.4 Conclusions, Practical Implications and Suggestions for Further Research

No single experimental study of human behaviour can give rise to definitive results, but arising out of the present studies it is reasonable to present the following broad conclusions and recommendations for further study:

(a) A feedback trainer of the type developed for these studies is most likely to be an effective training device for perceptual-motor skills which have several components, each of which is fairly difficult to perform in its own right, and which interact with one another such that poor performance of one creates a difficult situation in which to learn the other.

In the laboratory this situation was created by giving the operators unusual controls and high-order dynamics in a one-dimensional tracking task. In practice the situation is more likely to arise in the control of multi-dimensional systems in which the dynamics in each axis are different with strong cross-couplings between them.

It is predicted, therefore, that feedback training will be of value in situations where a number of skills have to be learnt and there are interactions between the performance of one and the learning of another. It is less likely to be of value in single-dimensional tasks, although they may be difficult, for example, high-order tracking in one axis; in multi-dimensional tasks where there is little interaction, for example, three-dimensional tracking with compatible controls/displays and the same dynamics in each axis with no cross-coupling; in multi-dimensional tasks where there is strong interaction but little opportunity for learning one of the interfering tasks, for example, tracking with the displayed signal immersed in random noise.

It is recommended that these predictions be validated by further experiments, and, in particular, that the efficacy of feedback training be investigated in a multi-axis system with compatible controls/ displays but differing dynamics in each axis and strong cross-coupling between them.

(b) The feedback trainer forms the basis for a sensitive test of an operator's ability to stabilize a control system. The level of difficulty at which the operator can attain a given error criterion is a measure of his ability. The homogeneous group of operators chosen for the present study was well-suited to the sutdy of the effects of different training regimes, but unsuitable for the validation of the 'trainer' as a test.

It is recommended that a small feedback training system be constructed specifically as a test device, and evaluated in a population having a normal range of abilities.

(c) The strong interaction between the effects of instructions and those of different training regimes, obtained in the present study, emphasize the importance of verbal instructions in teaching a perceptual-motor skill, and of controlling verbal effects in laboratory studies of the learning of such skills.

It is quite feasible, at the current state of technology, to incorporate an audio/visual teaching machine device in the simulator under control of the feedback training system. This device could be used to give the operator his initial instructions, evaluate his understanding of them, and give remedial instruction if necessary. It could also be used to give verbal instruction according to the level of performance, rate of learning, and control strategy of the trainee, for example, if the level of difficulty in the feedback training loop does not rise to a criterion level after a certain time. The same system could also be used to administer the questionnaires and evaluate the operator's response to, and knowledge of, the training situation.

It is recommended that future studies of feedback training incorporate a programmable audio/visual display in the training system, in order to control fully the verbal instructions to the operator, and evaluate his verbal behaviour.

(d) The theoretical studies do not distinguish between the learning of cognitive and perceptual-motor skills, and, whilst tracking tasks are most obviously amenable to feedback training through variation of 'difficulty' to maintain a desired sub-environment, the theory and its implications should apply equally to the training of cognitive skills, such as arithmetic and language. The 'paired-associate' learning employed by Gedye and Miller (1969) is an example of a cognitive skill where various levels of difficulty may be established, and indeed the trajectories of performance which they obtain are strikingly similar to those with the present feedback trainer.

It is recommended that examples of cognitive skills be analysed in terms of states of adaption and the occurrence of sub-environment phenomena, in order to examine the possible application of the theories of learning, and feedback training, developed to non-perceptual-motor skills.

(e) The theoretical studies, together with the experimental comparison of human operators and learning machines, have demonstrated that it is possible to establish a general theory of adaption and learning equally applicable to both artificial and natural learning systems. The theoretical foundations for the analysis of learning behaviour, and its relationship to the purpose and structure of the learning system, are couched in completely neutral and abstract terms, and form a mathematical rather than empirical system; the empirical content of the theory arises only in the decision as to the applicability of certain descriptions to observed behaviour.

The theory of adaptive systems and their training has been closely linked to the development of a general systems theory applicable to any system, and the further development of the theory should most fruitfully be at this general level - adaption and training will then become relations of stability and control on a system specifically derived for the analysis of learning behaviour. Experimental validation of theories relating learning behaviour to the epistemology of the environment may be tested with both human operators and with learning machines, and the increasing availability of computer-simulated learning systems implies that they will have a major role in future psychological studies.

It is recommended that theoretical studies of learning and training be placed within the framework of general systems theory and the mathematical theory of semigroups, and that increasing emphasis be placed upon the derivation of general system-theoretic results which specialize into statements about learning systems. It is also recommended that adaptive controllers and learning systems be used as subjects in experiments on human learning, both as an aid in experimental design, and for comparative purposes.

SECTION 8 : REFERENCES

ADAMS, J.A.[1964] Motor Skills, Ann.Rev.Psychol. 15 181-202

ALLPORT, D.A.[1968] Phenomenal Simultaneity and the Perceptual Moment Hypothesis, Brit.J.Psychol. 59 395-406.

ANDREAE, J.H., CASHIN, P.[1969] A Learning Machine with Monologue, Int.J.Man-Mach.Studies 1.

ANGEL, E.S., BEKEY, G.A.[1968] Adaptive Finite-State Models of Manual Control Systems, IEEE Trans. MMS-9 15-20.

ARBIB, M.A.(Ed.)[1968] Algebraic Theory of Machines, Languages and Semigroups, Academic Press.

ASHBY, W.R.[1957] An Introduction to Cybernetics, Wiley.

ASHBY, W.R.[1960] Design for a Brain, Chapman and Hall.

ASHBY, W.R.[1962] The Set Theory of Mechanism and Homeostasis, Tech.Rep.7, Electr.Eng.Res.Lab., U.Illinois.

ASHBY, W.R.[1965] Constraint Analysis of Many-Dimensional Relationships, in Progress in Bio-Cybernetics, 2, Elsevier.

AUGENSTINE, L.G.[1954] Evidence of Periodicities in Human Task Performance, in Information Theory in Psychology, Free Press, Glencoe, Illinois.

BARTLETT, F.C.[1947] Some Problems of 'Display' and 'Control', in Miscellanea Psychologica, Librairie Philosophique.

BEKEY, G.A.[1962] An Investigation of Sampled Data Models of the Human Operator in Control Systems, Doctoral Diss., University of California, and ASD-TDR-62-36.

BEKEY, G.A.[1965] Description of the Human Operator in a Control System, in Modern Control Systems Theory, McGraw Hill.

BELLMAN, R.[1957] Dynamic Programming, Princeton U.Press.

BELLMAN, R.E., DREYFUS, S.E.[1962] Applied Dynamic Programming, Princeton U.Press.

BENSON, A.J., HUDDLESTON, H.F., ROLFE, J.M.[1965] A Psycho-Physiological Study of Compensatory Tracking on a Digital Display, RAF Inst.Aviation Medicine, Farnborough,UK.

BERGE, C.[1962] Theory of Graphs and Their Applications, Wiley.

BERSH, P.J., NOTTERMAN, J.M., SCHOENFIELD, W.N.[1957] The Efficiency of Pursuitrotor Performance During Experimentally Induced Anxiety, School Aviation Medicine, Texas,USA.

BILODEAU, E.A., BILODEAU, I.McD.[1961] Motor Skills Learning Ann.Revs.Psychol. 12, 243-280.

BILODEAU, E.A.[1966] Acquisition of Skill, Academic Press.

BIRKHOFF, G.D.[1927] Dynamical Systems, Am.Math.Soc.Coll., IX.

BLAKELOCK, J.H.[1965] Automatic Control of Aircraft and Missiles, Wiley.

BODE, H.W.[1960] Feedback - the History of an Idea, Active Networks and Feedback Systems, Polytechnic Inst.Brooklyn.

BOOTH, T.L.[1967] Sequential Machines and Automata Theory, Wiley.

BOOTON, R.C.[1953] The Analysis of Nonlinear Control Systems with Random Inputs, in Nonlinear Circuit Analysis II, Polytechnic Inst.Brooklyn Press.

BREMERMANN, H.J., ROGSON, M., SALAFF, S.[1965] Search by Evolution, in Biophysics and Cybernetic Systems, Spartan Books.

BREMERMANN, H.J., ROGSON, M., SALAFF, S.[1966] Global Properties of Evolution Processes, in Natural Automata and Useful Simulations, Spartan Books.

BRIGGS, G.E.[1961] On the Scheduling of Training Conditions for the Acquisition and Transfer of Perceptual-Motor Skills, NAVTRADEVCEN 836-1.

BRIGGS, G.E.[1962] Research on Augmented Feedback and the Acquisition and Transfer of Skill, NAVTRADEVCEN 836-3.

BRIGGS, G.E.[1964] The Generality of Research on Transfer Functions, in Categories of Human Learning, Academic Press.

BRINDLEY, G.S.[1960] Physiology of the Retina and Visual Pathway, Waverly Press.

BROADBENT, D.E.[1958] Perception and Communication, Pergamon.

BROWN, I.D., POULTON, E.C.[1961] Measuring the Spare 'Mental Capacity' of Car Drivers by a Subsidiary Task, Ergonomics 4 35-40.

BUSH, R.R., MOSTELLER, F.[1955] Stochastic Models for Learning, Wiley.

BUSHAW, D.W.[1953] Differential Equations with Discontinuous Forcing Terms, Stevens Tech.Inst., N.J., Rep.469.

BUSHAW, D.W.[1967] A Stability Criterion for General Systems, Math.Syst.Theory 1 79-88.

CALLAWAY, E., ALEXANDER, J.D.[1962] The Temporal Coding of Sensory Data - An Investigation of Two Theories, J.Gen. Psychol. 62 293-309.

CARBONELL, J.R.[1966] A Queueing Model of Many-Instrument Visual Sampling, IEEE Trans. HFE-7 157-164.

CARNAP, R.[1950] Logical Foundations of Probability, University of Chigaco Press.

CARNAP, R.[1956] Meaning and Necessity, University of Chicago Press

CHERNIKOFF, R., DUEY, J.W., TAYLOR, F.V.[1959] Two-Dimensional Tracking with Identical and Different Control Dynamics in Each Co-ordinate, USNRL Rep. 5424.

CHERNIKOFF, R.[1962] Human Equalization, Concept, Machine and Measurement, in Complex Vehicular Controls, Inst. Aviation Medicine, Farnborough, UK.

CHERNIKOFF, R., LeMAY, M.[1963] Effect of Various Display-Control Configurations on Tracking with Identical and Different Co-ordinate Dynamics, J.Exptl.Psychol. 66 95-99.

CHILES, W.D.[1957] Psychological Stress as a Theoretical Construct, Wright Air Development Center, AD130942.

CLARK, R.M.[1963] Panel Discussion in Adaptive Control Systems, Pergamon Press.

CLIFFORD, A.H., PRESTON, G.B.[1961] The Algebraic Theory of Semigroups I, American Math.Soc.

CLIFFORD, A.H., PRESTON, G.B.[1967] The Algebraic Theory of Semigroups II, American Math.Soc.

CRAIK, K.J.W.[1947] Theory of the Human Operator in Control Systems, Brit.J.Psychol. 38 56-61.

DANDER, V.A.[1963] Predicting Pilot Ratings of Multi-Axis Control Tasks from Single-Axis Data, IEEE Trans. HFE 4 15-24.

DAY, R.H.[1956] Relative Task Difficulty and Transfer of Training in Skilled Performance, Psychol.Bull. 53 160.

DeLESSIO, N.X., PALIN, J.[1961] An Adaptive Model for the Human Operator, Polytechnic Inst.Brooklyn, PIBMRI-909-61.

DEUTSCH, J.A.[1960] The Structural Basis of Behaviour, Cambridge University Press.

DIAMANTIDES, N.D.[1958] A Pilot Analog for Airplane Pitch Control, J.Aero.Sciences 361-394.

DONALSON, D.D., KISHI, F.H.[1965] Review of Adaptive Control System Theories and Techniques, in Modern Control Systems Theory, McGraw Hill.

DUEY, J.W., CHERNIKOFF, R.[1959] The Use of Quickening in One Co-ordinate of a Two-Dimensional Tracking System, USNRL 5428.

EASON, R.G.[1963] Relation Between Effort, Tension Level, Skill and Performance Efficiency in a Perceptual-Motor Task, Percept.Mot.Skills 16 297-317.

ELKIND, J.I.[1956] Characteristics of Simple Manual Control Systems, MIT Tech.Rep. 111.

ELKIND, J.I., DARLEY, D.L.[1963] The Normality of Signals and Describing Function Measurements of Simple Manual Control Systems, IEEE Trans. HFE-7 115-124.

FEIGENBAUM, E., FELDMAN, J.[1964] Computers and the Brain, McGraw Hill.

FELDBAUM, A.A.[1960] Dual Control Theory I, Aut.i Tel. 21 1240-1249

FELDBAUM, A.A.[1960*] Dual Control Theory II, Aut.i Tel. 21 1453-1464.

FELDBAUM, A.A.[1961] Dual Control Theory III, Aut.i Tel. 22 3-16

FELDBAUM, A.A.[1961*] Dual Control Theory IV, Aut.i Tel. 22 129-142.

FELDBAUM, A.A.[1963] Dual Control Theory Problems, Proc.2nd Int. Congr.IFAC, Basle.

FLAMENT, C.[1963] Applications of Graph Theory to Group Structure Prentice Hall.

FLORENTIN, J.J.[1962] Optimal, Probing, Adaptive Control of a Simple Bayesian System, J.Electron.Contr., 13 165-177.

FOGEL, L.J., OWENS, A.J., WALSH, M.J.[1965] Artificial Intelligence Through a Simulation of Evolution, in Biophysics and Cybernetic Systems, Spartan Books.

FREEMAN, H.[1963] Panel Discussion in Adaptive Control Systems, Pergamon Press.

FUCHS, A.H.[1962] The Progression-Regression Hypotheses in Perceptual-Motor Skill Learning, J.Exptl.Psychol. 63 177-182.

FULLER, A.T.[1960] Phase Space in the Theory of Optimum Control, J.Electron.Contr. 8 341-400.

FULLER, A.T.[1960*] Relay Control Systems Optimized for Various Performance Criteria, Proc.1st Int.Congr.IFAC, Moscow.

FULLER, A.T.[1962] Bibliography of Optimum Nonlinear Control of Determinate and Stochastic-Definite Systems, J.Electron.Contr. 13 590-611.

FULLER, A.T.[1967] Linear Control of Nonlinear Systems, Int.J.Contr. 5 197-243.

GAGNE, R.M., FOSTER, H., CROWLEY, M.E.[1948] The Measurement of Transfer of Training, Psychol.Bull. 45 97-130.

GAINES, B.R.[1965] Automated Feedback Trainers for Perceptual-Motor Skills, Tech.Rep.1, M.o.D. Contract, Cambridge, UK.

GAINES, B.R.[1966] Teaching Machines for Perceptual-Motor Skills, Proc.Programmed Learning Conf., Loughborough, in Aspects of Educational Technology, Methuen 1967.

GAINES, B.R.[1966*] The Evaluation and Control of Human Skill, IEE Coll.Publn., Human Operator Studies, London.

GAINES, B.R.[1967] Automated Feedback Trainers for Perceptual-Motor Skills, Final Report M.o.D. Contract.

GAINES, B.R.[1967*] Techniques of Identification with the Stochastic Computer, Proc.IFAC Symp.Problems of Identification in Automatic Control Systems, Prague.

GAINES, B.R., QUARMBY, D.J.[1968] The Convergence of Adaptive Pattern Classifiers in Control Systems, Dept.Electrical Eng. Science, University of Essex, RM 68/2.

GAINES, B.R.[1968] Varieties of Computer - Their Applications and Inter-Relationships, Proc.IFAC Int.Pulse Symp., Budapest.

GAINES, B.R.[1968*] Adaptive Control Theory, in Encyclopaedia of Information, Linguistics and Control, Pergamon.

GAINES, B.R.[1968**] Training the Human Adaptive Controller, Proc.IEE

GAINES, B.R.[1968***] The Control of Human Learning, Digest IEEE Int.Conv., New York.

GAINES, B.R.[1968****] Adaptively Controlled Instruction of a Tracking Skill, NATO Symp.Major Trends in Programmed Learning, Nice.

GAINES, B.R.[1969] Stochastic Computing Systems, in Advances in Information Systems Science

GARVEY, W.D., TAYLOR, F.V.[1959] Interactions Among Operator Variables, System Dynamics and Task-Induced Stress, J.Appl.Psychol. 43 79-85.

GARVEY, W.D., HENSON, J.B.[1959] Interactions Between Display Gain and Task-Induced Stress in Manual Tracking, J.Appl. Psychol. 43 205-208.

GEDYE, J.L., MILLER, E.[1969] Automation of Psychological Assessment, Int.J.Man-Machine Studies, 1(2).

GIBBS, C.B.[1951] Transfer of Training and Skill Assumptions in Tracking, Quart.J.Exptl.Psychol. 3 99-110.

GIBSON, J.E.[1963] Nonlinear Automatic Control, McGraw Hill.

GILL, A.[1961] State-Identification Experiments in Finite Automata, Inf.Contr. 4 132-154.

GILL, A.[1962] An Introduction to the Theory of Finite State Machines, McGraw Hill.

GINSBURG, S.[1962] An Introduction to Mathematical Machine Theory, Addison-Wesley.

GON, J.J.D.van der, THURING, J.Ph., STRACKEE, J.[1962] A Handwriting Simulator, Phys.in Med.Biol., 6 407-414.

GON, J.J.D.van der, THURING, J.Ph.[1965] The Guidance of Human Writing Movements, Kybernetik 2 145-148.

GUTHRIE, E.R.[1952] The Psychology of Learning, Harper.

HAHN, W.[1963] Theory and Application of Liapunov's Direct Method, Prentice Hall.

HALL, I.A.M.[1957] Effects of Controlled Element on the Human Pilot, Aero.Eng.Lab., Princeton University, Rep.No.389.

HALL, I.A.M.[1963] Study of the Human Pilot as a Servo Element, J.Roy.Aero.Soc. 67 351-360.

HAMMOND, P.H., MERTON, P.A., SUTTON, G.G.[1956] Nervous Gradation of Muscular Contraction, Brit.Med.Bull. 12 214-218.

HARARY, F., NORMAN, R.Z., CARTWRIGHT, D.[1965] Structural Models, Wiley.

HARTMANIS, J., STEARNS, R.E.[1966] Algebraic Structure Theory of Sequential Machines, Prentice-Hall.

HELSON, H.[1949] Design of Equipment and Optimal Human Operation, Amer.J.Psychol. 62 473-479.

HENNIE, F.C.[1968] Finite-State Models for Logical Machines, Wiley.

HICK, W.E.[1948] The Discontinuous Functioning of the Human Operator in Pursuit Tasks, Quart.J.Exptl.psychol. 36-51.

HICK, W.E.[1952] On the Rate of Gain of Information, Quart. J. Exptl.Psychol. 4 11-26.

HOLDING, D.H.[1962] Transfer Between Easy and Difficult Tasks, Brit.J.Psychol. 53= 397-407.

HOLDING, D.H., MACRAE, A.W.[1966] Rate and Force of Guidance in Perceptual-Motor Tasks with Reversed or Random Spatial Correspondence, Ergonomics 9 289-296.

HUDSON, E.M.[1964] Adaptive Training and Non-Verbal Behaviour, NAVTRADEVCEN TR 1395-1.

HUNTER, W.S. [1934] Learning - IV Experimental Studies of Learning, in A Handbook of Experimental Psychology,

JEX, H.R., CROMWELL, C.H., SISKIND, R.K.[1960] Correlation of Experimental and Theoretical Limits for Pilot Control of Unstable Second Order Systems, Syst.Tech.Inc.,Calif.

JEX, H.R., McDONNELL, J.D., PHATAK, A.V.[1966] A 'Critical' Tracking Task for Man-Machine Research Related to the Operator's Effective Delay Time, NASA CR-616.

JEX, H.R., McDONNELL, J.D., PHATAK, A.V.[1966*] A Critical Tracking Task for Manual Control Research, IEEE Trans. HFE-7 138-145.

JURY, E.I.[1958] Sampled-Data Control Systems, Wiley.

KALMAN, R.E., BERTRAM, J.E.[1959] A Unified Approach to the Theory of Sampling Systems, j.Franklin Inst. 267 405-436.

KALMAN, R.E.[1960] On the General Theory of Control Systems, Proc.1st Int.Congr.IFAC, Moscow.

KELLEY, C.R.[1962] Self-Adjusting Vehicle Simulators, IRE Int. Congr.Human Factors in Electronics, Long Beach, Calif.

KELLEY, C.R.[1967] Further Research with Adaptive Tasks, Dunlap and Ass., Final Rep. Nonr-4986(00).

KELLEY, C.R., PROSIN, D.J.[1968] Adaptive Performance Measurement, Dunlap and Ass., Final Rep. Nonr-4986(00).

KILPATRICK, P.S.[1964] Comparison of Relay and Manual Controllers for Systems with High Order Dynamics, BS Thesis, MIT.

KINKADE, R.G.[1963] A Differential Influence of Augmented Feedback on Learning and on Performance, AMRL-TDR-63-12.

KNOWLES, W.B.[1963] Operator Loading Tasks, Human Factors 155-161.

KOLERS, P.A.[1962] Intensity and Contour Effects in Visual Masking, Vision Res. 2 277-294.

KOLERS, P.A.[1968] Some Psychological Aspects of Pattern Recognition, in Recognizing Patterns, MIT Press.

KOLK, W.R.[1961] Modern Flight Dynamics, Prentice Hall.

KREIFELDT, J.G.[1965] A Sampled-Data Pursuit Tracking Model, IEEE Trans.HFE-6 65-73.

KRENDEL, E.S.[1951] A Preliminary Study of the Power Spectrum Approach to the Analysis of Perceptual-Motor Performance, Wright Patterson Air Force Base, Ohio, Tech.Rep. 6723.

KRENDEL, E.S.[1952] The Spectral Density Study of Tracking Performance, Wright Patterson Air Force Base, Ohio, TR-52-11.

KRENDEL, E.S., McRUER, D.T.[1960] A Servomechanisms Approach to Skill Development, J.Franklin Inst., 269 24-42.

LANGE, G.W.[1965] Syntheses of a Model of the Human Operator Engaged in a Tracking Task, PhD Thesis, University of London.

LEMAY, L.P., WESCOTT, J.H.[1962] The Simulation of Human Operator Tracking Using an Intermittaen Model, IRE Int.Congr. Human Factors in Electronics.

LEONARD, T.E.[1960] Optimizing Linear Dynamics for Human-Operated Systems by Minimizing the Mean Square Tracking Error, IRE WESCON 4 57-62.

LEONARD, J.A.[1962] Discussion following paper by Chernikoff[1962]

LEWIS, B.N., COOK, J.A.[1969] Toward a Theory of Telling, Int.J. Man-Machine Studies, 1(2).

LI, Y.T., YOUNG, L.R., MEIRY, J.L.[1965] Adaptive Functions of Man in Vehicle Control Systems, Proc.IFAC Symp., Teddington,UK.

LIAPIN, E.S.[1963] Semigroups, Amer.Math. Soc.

LICKLIDER, J.C.R.[1960] Quasi-Linear Models in the Study of Manual Tracking, in Developments in Mathematical Psychology, Glencoe Free Press.

LINDSLEY, D.B.[1952] Psychological Phenomena and the Electro-encephalogram, EEG Clin.Neurophysiol. 4 443-454.

LIPPOLD, O.C.J., REDFEARN, J.W.T., VUCO, J.[1957] The Rhythmical Activity of Groups of Motor Units in the Voluntary Contraction of Muscle, J.Physiol. 137 473-487.

LURIA, A.[1961] The Role of Speech in the Regulation of Normal and Abnormal Behaviour, Pergamon Press.

MACDONALD ROSS, M.[1969] Programmed Learning - A Decade of Development, Int.J.Man-Machine Studies, 1 73.

MACKWORTH, N.H.[1950] Researches on the Measurement of Human Performance, MRC Rep.268, HMSO.

MAGIROS, D.G.[1966] Stability Concepts of Dynamical Systems, Inf.Contr. 9 531-548.

MARTENS, H.H.[1959] Two Notes on Machine 'Learning', Inf. Contr. 2 364-379.

McDONNELL, J.D., JEX, H.R.[1967] A 'Critical Tracking Task for Man-Machine Research Related to the Operator's Effective Delay Time, II, NASA CR-674.

McGEOCH, J.A., IRION, A.L.[1952] The Psychology of Human Learning

McNAUGHTON, R.[1961] The Theory of Automata - A Survey, in Advances in Computers, 2 379-421, Academic Press.

McRUER, D.T., KRENDEL, E.S.[1957] Dynamic Response of Human Operators, WADC-TR-56-524.

McRUER, D.T., KRENDEL, E.S.[1959] The Human Operator as a Servo System Element, J.Franklin Inst. 267 381-403, 511-536.

McRUER, D.T., GRAHAM, D., KRENDEL, E.S., REISENER, W.[1965] Human Pilot Dynamics in Compensatory Systems, AFFDL-TR-65-15.

MEALY, G.H.[1955] A Method of Synthesizing Sequential Circuits, Bell Syst.Tech.J. 34 1045-1079.

MEETHAM, A.R.[1963] Preliminary Studies for Machine Generated Index Vocabularies, Lang.Speech 6(1).

MEETHAM, A.R.[1966] Algorithms to Assist in Finding the Complete Sub-Graphs of a Given Graph, Nature 211 105.

MILLER, D.C.[1965] The Effect of Performance-Scoring Criteria on Compensatory Tracking Behaviour, IEEE Trans. HFE-6 62-65.

MINSKY, M., SELFRIDGE, O.G.[1961] Learning in Random Nets, in Information Theory, Butterworths.

MINSKY, M., PAPERT, S.[1968] Perceptrons and Pattern Recognition, MIT Press.

MIRSKY, L.[1955] Linear Algebra, Oxford, Clarendon Press.

MISHKIN, E., BRAUN, L.[1961] Adaptive Control Systems, McGraw Hill.

MOE, M.L., MURPHY, G.J.[1962] An Approach to Self-Adaptive Control Based on the Use of Time Moments and a Model Reference, IRE Trans. AC-7 82-92.

MOORE, E.F.(Ed.)[1964] Sequential Machines, Addison Wesley.

NEWELL, A., SHAW, J.C., SIMON, H.A.[1959] Report on the General Problem Solving Problem, Proc.Int.Conf.Inf. Processing, Paris.

NILSSON, N.J.[1965] Learning Machines, McGraw Hill.

NORTH, J.D.[1952] The Human Transfer Function in Servo Systems, in Automatic and Manual Control, Butterworths.

NOVIKOFF, A.[1963] On Convergence Proofs for Perceptrons, in Automata Theory, Polytechnic Inst.Brooklyn Press.

NYQUIST, H.[1932] Regeneration Theory, Bell Syst.Tech.J. 11 126-147.

ORE, O.[1962] Theory of Graphs, Am.Math.Soc.Coll.Publn. 38.

ORNSTEIN, G.N.[1963] The Automatic Analog Determination of Human Transfer Function Coefficients, Med.Electron.Biol.Eng. 1 377-387.

PASK, G.[1960] The Teaching Machine as a Control Mechanism, Trans.Soc.Instr.Tech. 12 72-82.

PASK, G.[1961] An Approach to Cybernetics, Hutchison.

PASK, G.[1964] Adaptive Teaching Machines, in Teaching Machines, Pergamon.

PASK, G.[1965] Teaching as a Control Process, Control.

PASK, G.[1965*] Comments on the Cybernetics of Ethical, Sociological and Psychological Systems, in Progress in Biocybernetics 3, Elsevier.

PEW, R.W.[1963] Temporal Organization in Skilled Performance, PhD Thesis, University of Michigan.

POINCARE, H.[1885] Sur les Courbes Definies par les Equations Differentielles, J.Math.Pures et Appl. 167-244.

POULTON, E.C.[1952] Perceptual Anticipation in Tracking with Two-Pointer and One-Pointer Displays, Brit.J.Psychol. 43 222-229.

POULTON, E.C.[1952*] The Basis of Perceptual Anticipation in Tracking, Brit.J.Psychol. 43 295-302.

POULTON, E.C.[1957] On the Stimulus and Response in Pursuit Tracking, J.Exptl.Psychol. 53 189-194.

POULTON, E.C.[1957*] Learning the Statistical Properties of the Input in Pursuit Tracking, J.Exptl.Psychol. 54 28-32.

POULTON, E.C.[1962] A Source of Nonlinearity in Pursuit Tracking, in Complex Vehicular Controls, a Symposium at Inst.Aviation Medicine, Farnborough, UK.

POULTON, E.C.[1965] On Increasing the Sensitivity of Measures of Performance, Ergonomics 8 69-76.

POULTON, E.C.[1966] Tracking Behaviour, in The Acquisition of Skill, Academic Press.

PREYSS, A.E., MEIRY, J.L.[1968] Stochastic Modelling of Human Learning Behaviour, IEEE Trans. MMS-9 36-46.

QUARMBY, D.J.[1968] The Influence of the Environment on the Convergence of a Learning Machine, Electronics Letters

RADANOVIC, L.(Ed.)[1966] Sensitivity Methods in Control Theory, Pergamon.

RAIBLE, R.H.[1963] Comments on 'An Approach to Self-Adaptive Control Based on the Use of Time Moments and a Model Reference', IEEE Trans. AC 8 270.

RIESZ, F., SzNAGY, B.[1955] Functional Analyis, Frederick Ungar.

ROSENBLATT, F.[1962] Principles of Neurodynamics, Spartan Books.

RUSSELL, L.[1951] Characteristics of the Human as a Linear Servo Element, MS Thesis, MIT.

SALTON, G., SUSSENGUTH, E.H.[1964] Some Flexible Information Retrieval Systems Using Structure Matching Procedures, AFIPS 25 SJCC 587-597.

SENDERS, J.W.[1961] Adaptive Teaching Machines, in Programmed Learning and Computer-Based Instruction, Wiley.

SENDERS, J.W.[1964] The Human Operator as a Monitor and Controller of Multi Degree of Freedom Systems, IEEE Trans. HFE-5 2-5.

SENDERS, J.W., WARD, J.L., CARBONELL, J.R.[1967] Human Visual Sampling Processes - A Simulation Validation Study, Final Rep.Contract NAS-1-5059, Bolt Beranek and Newman.

SHERIDAN, T.B.[1960] Experimental Analysis of Time Variation of the Human Operator's Transfer Function, Proc.1st Int.Congr. IFAC, Moscow.

SHERIDAN, T.B., ROLAND, R.D.[1966] A Normative Model for Control of Vehicle Trajectory in an Emergency Maneuver, Ann.Meeting Highway Research Board, Washington DC.

SIEGEL, S.[1956] Non-Parametric Statistics, McGraw Hill.

SKOLNICK, A.[1966] Stability and Performance of Manned Control Systems, IEEE Trans. HFE-7 115-124.

SMALLWOOD, R.D.[1962] A Decision Structure for Teaching Machines, MIT Press.

SMITH, O.J.M.[1962] Nonlinear Computations in the Human Controller, IRE Trans. BME-9 125-128.

SMITH, R.H.[1963] On the Limits of Manual Control, IEEE Trans. HFE-4 56-59.

SOMMERHOF, G.[1950] Analytical Biology, Oxford U.Press.

STANIER, R.Y.[1953] Adaptation, Evolutionary and Physiological, in Adaptation in Micro-Organisms, Third Symp.Soc.Gen. Microbiol., Cambridge U.Press.

STARK, L., VOSSIUS, G., YOUNG, L.R.[1962] Predictive Control of Eye Tracking Movements, IRE Trans. HFE-3 52-57.

STOCKBRIDGE, H.C.W., SIDDALL, G.J.[1956] Memorandum to the Commonwealth Advisory Committee on Defence Science.

STROUD, J.M.[1954] The Fine Structure of Psychological Time in Information Theory in Psychology, Free Press, Glencoe, Ill.

SUMMERS, L.G., ZIEDMAN, K.[1964] A Study of Manual Control Methodology with Annotated Bibliography, NASA-CR 125.

SURWILLO, W.W.[1963] The Relation of Simple Response Time to Brain-Wave Frequency and the Effects of Age, EEG Clin. Neurophysiol., 15 105-114.

SWORDER, D.D.[1966] A Study of the Relationship Between Identification and Optimization in Adaptive Control Problems, J.Franklin Inst., 281 198-213.

TAYLOR, D.H.[1964] Driver's Galvanic Skin Response and the Risk of Accident, Ergonomics 7 439-451.

TELFORD, C.W.[1931] The Refractory Phase of Voluntary and Associated Responses, J.Exptl.Psychol. 14.

THOMAS, R.E.[1962] Development of New Techniques for Analysis of Human Controller Dynamics, MRL-TDR-62-65.

THORPE, W.H.[1956] Learning and Instinct in Animals, Methuen.

TINBERGEN, N.[1951] The Study of Instinct, Oxford U.Press.

TOU, J.T., WILCOX, R.H.[1964] Computer and Information Sciences, Spartan Books.

TRUXAL, J.G.[1961] Identification of Process Dynamics, in Adaptive Control Systems, McGraw Hill,

TRUXAL, J.G.[1963] Adaptive Control, Survey Paper 2nd Int.Congr. IFAC, Basle.

TUSTIN, A.[1947] The Nature of the Human Operator's Response in Manual Control and its Implication for Controller Design, J.I.E.E.(London) 94 190-203.

VASWANI, P.K.T.[1965] Mechanized Storage and Retrieval of Information, Rev.Int.Doc. 32 19-22.

VENABLES, P.H.[1960] Periodicity in Reaction Time, Brit.J. Psychol. 51 37-43.

WALKER, N.K.[1963] The Effect of a Particular Stress on One Man's Performance of a Tracking Task, AD428112.

WARD, J.R.[1958] The Dynamics of a Human Operator in a Control System, PhD Thesis, University of Sydney, Australia.

WARD, J.L., SENDERS, J.W.[1966] Methodological Studies of Tracking Behaviour - The Effect of Various Supplemental Information Feedbacks, Nonr 4605, Bolt,Beranek,Newman.

WEIR, D.H., PHATAK, A.V.[1967] Model of Human Operator Dynamic Response to Step Transitions in Controlled Element Dynamics, NASA CR-671.

WEIZENBAUM, J.[1966] ELIZA - A Computer Program for the Study of Natural Language Communication Between Man and Machine, Comm.ACM 9 36-45.

WELFORD, A.T.[1952] The 'Psychological Refractory Period' and the Timing of High-Speed Performance, A Review and Theory, Brit.J.Psychol. 43 2-19.

WEXLER, J.D.[1970] A Teaching Program that Generates Simple Arithmetic Problems, Int.J.Man-Machine Studies, 2 1.

WHITE, C.T.[1963] Temporal Numerosity and the Psychological Unit of Duration, Psychol.Monogr. 77(12).

WIDROW, B., SMITH, F.W.[1964] Pattern-Recognizing Control Systems, in Computer and Information Sciences, Spartan.

WIDROW, B.[1966] Bootstrap Learning in Threshold Logic Systems, Proc. 3rd Int.Congr.IFAC, London.

WIENER, N.[1914] A Simplification of the Logic of Relations, Proc.Cambridge Phil.Soc. 17 387-390.

WIENER, N.[1948] Cybernetics, Wiley.

WILDE, R.W., WESCOTT, J.H.[1962] The Characteristics of the Human Operator Engaged in a Tracking Task, Automatica 1 5-19.

WYMORE, A.W.[1967] A Mathematical Theory of Systems Engineering, Wiley.

YOUNG, L.R., STARK, L.[1963] Variable Feedback Experiments Testing a Sampled Data Model for Eye Tracking Movements, IEEE Trans. HFE 38-51.

YOUNG, L.R., GREEN, D,M., ELKIND, J.I., KELLY, J.A.[1964] Adaptive Dynamic Response Characteristics of the Human Operator in Simple Manual Control, IEEE Trans. HFE-5 6-13.

YOUNG, L.R., STARK, L.[1965] Biological Control Systems - A Critical Review and Evaluation, NASA CR-190.

YOUNG, L.R., MEIRY, J.L.[1965] Bang-Bang Aspects of Manual Control in High-Order Systems, IEEE Trans. AC-10 336-341.

ZADEH, L.A.[1963] On the Definition of Adaptivity, Proc.IEEE 469-470.

ZADEH, L.A., DESOER, C.A.[1963] Linear System Theory - The State-Space Approach, McGraw Hill.

ZADEH, L.A.[1964] The Concept of State in System Theory, in Views on General Systems Theory, Wiley.

ZIEGLER, P.N., BIRMINGHAM, H.P., CHERNIKOFF, R.[1962] An Equalization Teaching Machine, USNRL Rep.5855.

APPENDIX 1 : ADAPTIVE AND LEARNING CONTROLLERS

A1.1 Introduction

A unified approach to human and machine learning has been taken in the studies reported in this thesis involving both theoretical and experimental comparisons between human behaviour and that of automatic adaptive controllers. In this appendix is gathered the background and reference material of adaptive controllers and learning systems relevant to the studies of Chapters 3 and 6.

A.1.2 Modes of Learning in Intelligent Artifacts

In Chapter 3, the fundamental structure of an adaptive controller is analysed as a two-level hierarchy, in which the lower level implements a control policy which is selected by the upper level. This splitting of what, even in the case of automatic controllers, will normally be an integrated structure, has an arbitrary element, similar to that inherent in the definition of a 'task' for the behavioural analysis of adaption. This arbitrariness resides in the definition of the family of control policies from which the upper level 'selects', and can lead to some very simple control systems being termed 'adaptive'. However, this is in itself not necessarily disadvantageous, since the analysis of very simple 'adaptive' structures may elucidate problems of learning in more complex controllers.

Most adaptive control systems perform some form of identification of their environment, although the simple linear model common in automatic control (Truxal 1961) is inadequate for more general situations. Having evaluated some characteristics of its environment, the controller must used a decision procedure to generate a control policy which is, in some sense, optimal for an environment with these characteristics. Because this strategy does not involve feedback from the actual performance to the control policy, it has been termed open-loop adaption (Freeman 1963). In Chapter 3, the epistemological problems of open-loop adaption are analysed, and it is shown that identification of an initial sub-environment, which is not the desired one, can lead to the selection of a control policy which maintains the same sub-environment and does not converge to an optimal performance.

If the performance of the controller when implementing a particular control policy is measured, and this measurement is used in the selection of another policy in an attempt to improve the performance, the adaptive strategy is termed closed-loop adaption. The earliest and simplest example of a closed-loop adaptive system is Ashby's homeostat (Ashby 1957, 1960), which changes its control strategy at random until the desired state of equilibrium is reached. Closely related to this are the 'evolutionary' adaptive systems of Bremermann (1965, 1966) and Fogel (1965), which change some characteristics of their control policy, but revert to the previous policy if this does not lead to improved performance. The epistemological problems of closed-loop adaptive systems are analysed in Chapter 3, and it is shown that maintaining an initial sub-environment which is the total environment leads to excessively slow convergence, whilst not doing so may cause convergence to local minima of the performance criterion.

There is a further mode of learning which does not readily fit into either category, and which seems to be vital to animal development and the establishment of language (Tinbergen 1951, Thorpe 1956), and this is learning by mimicing another controller. This mode of learning was one of the first copied by engineers in synthesizing 'learning machines', and is of interest, not only because of the great amount of experimental data on the use of adaptive-threshold-logic based machines, but also because of its importance in the analysis of some aspects of the effects of instructions on human, and machine, learning of control skills.

A.1.3 Adaptive-Threshold-Logic Pattern-Classifiers

The use of adaptive-threshold-logic elements (ATLE) for pattern-classification was first proposed and studied by Rosenblatt (1962) as a model of nervous processes, linking the neural nets of the brain with powerful pattern-recognition capabilities of human perception. Since then ATLE have been established as fundamental components of many artificial learning systems, and they have become a popular research topic with a large literature (Nilsson 1965).

The problem given to an ATLE is to learn to classify a set of stimuli into two (or more) disjoint classes. During the learning phase, stimuli are presented to it one by one, together with a statement of the class to which they belong. Thus, there is necessarily an independent system for evaluating the class of stimuli, and the problem of the ATLE may be viewed as one of coming to mimic this evaluator exactly.

The mechanism of learning in an ATLE depends on numeric operations, and the stimuli must be coded into a vector (or ordered array of numbers) before they may be classified by the ATLE. Figure A1-1 illustrates the learning situation of an ATLE, coupled to a stimulus generator (or

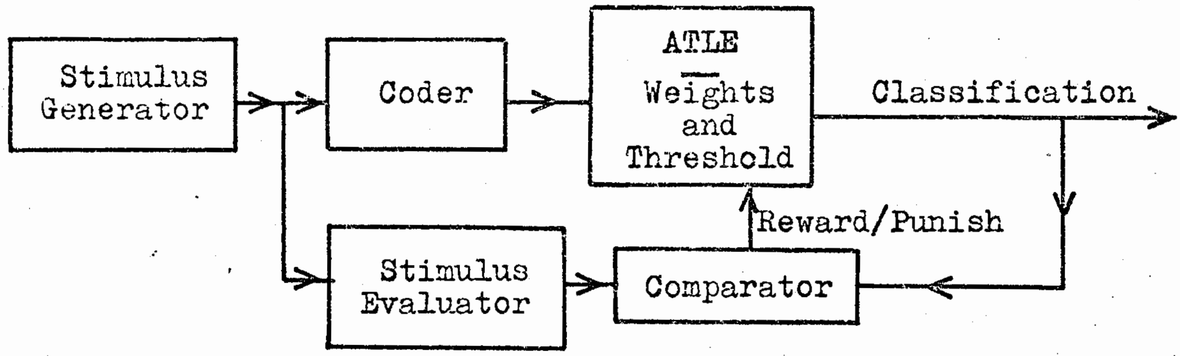


Figure A1-1 Learning Situation of Adaptive-Threshold-Logic Element

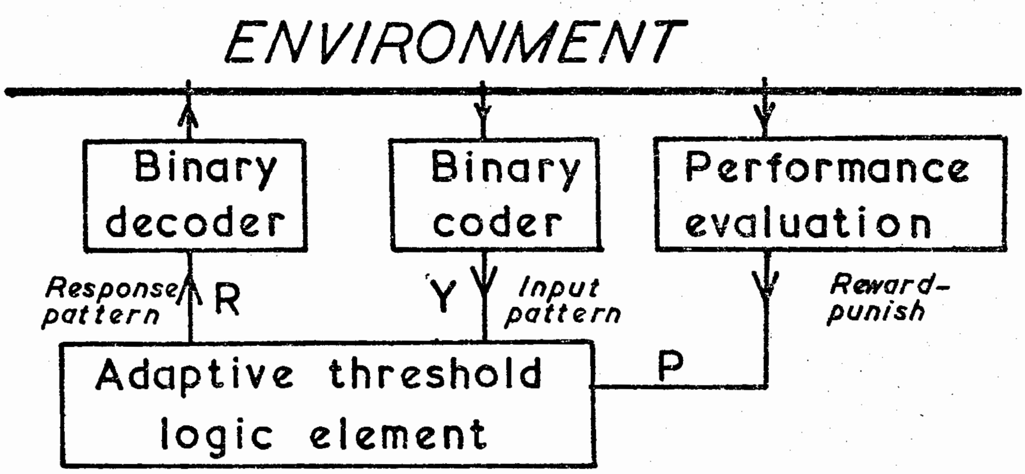


Figure A1-2 Pattern-Classifying Adaptive Controller environment) through a coder with numeric output, and receiving immediate performance feedback from an independent stimulus evaluator.

The structure and adaptive algorithm of the ATLE itself varies somewhat between different workers, but Novikoff (1963) has shown that the majority of cases may be typified by the following procedures:

(i) Coding of stimulus - let the coder represent a stimulus, S^J, from the set of possible input stimuli, by a k-vector, Y^J, whose components are $Y_i^J = \pm 1$, $1 \leq i \leq k$.

(ii) Internal weights - within the ATLE there is stored a k-vector of 'weights', W, with components, W_i, $1 \leq i \leq k$, which determine the classification adopted by the ATLE, and are adjusted with its experience.

(iii) Decision procedure - one set of weights is used to make a binary, or dichotomous, classification. However, multi-way classifications may be considered as a set of binary decisions, and hence the binary classification may be considered without loss of generality. The binary classification is a function of the scalar product between the stimulus-vector and the weight vector -

$$\text{if} \quad W.Y = \sum_{i=1}^{k} W_i Y_i \qquad \boxed{A1.1}$$

$$\text{then} \quad W.Y > \Theta \Rightarrow \text{stimulus assigned to class}_1 \qquad \boxed{A1.2}$$

$$W.Y < -\Theta \Rightarrow \text{stimulus assigned to class}_2 \qquad \boxed{A1.3}$$

$$-\Theta \leq W.Y \leq \Theta \Rightarrow \text{stimulus not assigned} \qquad \boxed{A1.4}$$

where $\Theta > 0$ is a constant, or 'threshold'.

(iv) Adaptive procedure - after each decision the ATLE is informed whether the decision was correct, or incorrect, and modifies its weight values accordingly. If W' is the new weight vector then -

$$W' = W, \text{ if the decision was correct} \qquad \boxed{A1.5}$$

$$W' = W - Y, \text{ if the decision was incorrect and } W.Y > 0 \qquad \boxed{A1.6}$$

$$W' = W + Y, \text{ if the decision was incorrect and } W.Y < 0 \qquad \boxed{A1.7}$$

Novikoff (1963) has proved that, if there exists any weight vector which will given rise to a correct decision for every stimulus-vector to be classified, then, given a sequence of stimulus sequence frequently containing all patterns to be classified, the above adaptive procedure will cause the weight vector to converge to one giving the correct classification, although not necessarily the same one, provided the weights are not bounded in their values. This convergence property of the ATLE implies that learning will be successful under a wide range of conditiions. It is of interest to consider the more general implications of the conditions necessary for convergence, and their effect on the adaption-automaton of the ATLE.

A.1.3.1 Coding Constraints on ATLE Convergence

There are three postulates necessary to Novikoff's proof of convergence, which are relevant to the adaption and training of human operators. The first is that there does exist a weight vector which gives a solution to the classification problem. This is not true in general, and two classes of vectors which may be distinguished by taking their scalar products with a weight vector and determining its sign are said to be linearly separable. The separability applies to the vectors generated by the coder, not directly to the original stimuli, which need not necessarily have any numeric connotations.

Hence, the selection of coding between the original stimuli and the k-vector inputs to the ATLE may be viewed as problem of ergonomics, similar to the problem of selecting an appropriate form of display for the human operator, for example, in detecting sonar targets. The best coding will be such that all possible dichotomizations which may be required of the ATLE are linearly separable, but that a stimulus-vector of minimum dimensions consistent with this requirement is utilized. The ATLE will then be potentially adaptive to all required tasks, and Novikoff's result demonstrates that this is true no matter what its initial state, so that it is also compatibly adaptive to the set of tasks.

A1.3.2 Stimulus-Experience Constraints on ATLE Convergence

Novikoff's second postulate is that every stimulus in the set which is to be dichotomized is frequently present in the sequence of stimuli presented during the learning phase that is, for any Y^J in the set, and for any postitive integer, N, there exists $M > N$, such that the M'th stimulus presented $Y(M) = Y^J$. This postulate is stronger than necessary, because the constraint of linear separability implies that, if some subset of the stimuli has been dichotomized into two classes, A and B which are necessarily linear separable, then a new stimulus, x, belonging to neither class, may not necessarily be assignable at will to either A or

B. It may happen that A+x is linearly separable from B, but that B+x is not linearly separable from A, or vice versa. In this event, presentation of stimuli from the set (A+B), and from the set (A+B+x), must lead to the same dichotomization of the larger set. Hence, it is unnecessary to present x, and the set (A+B) may be said to be a support (Minsky and Papert 1968) for the classification of the set, (A+B+x).

For any required dichotomization, assumed linearly separable, there will be a number of dichotomizations involving few stimuli which are supports for it, and amongst these there will be one, or more, with a least number of elements, a minimum support. Any sub-sets of the stimuli containing less elements than this cannot be a support, and there will also be sub-sets containing more elements which are not supports. Thus there is a sub-environment phenomenom, in that stimulus generators which do not generate a sufficient variety of stimuli to support the required classification may not enable the ATLE to learn that classification. In the pattern-classification situations considered so far, there is no feedback from the behaviour of the ATLE to the stimulus generator, and hence the policy, or classification, of the ATLE does not affect its environment. However, when ATLEs are used as part of an adaptive control loop, the stimulus generator is influenced by the ATLE policy, and may be forced into a state where it is not emitting a support set for the required control policy; this problem is discussed further in Section A1.4.

A1.3.3 Weight-Magnitude Constraints on ATLE Convergence

The third postulate necessary to ATLE convergence is that the values of the weights in the ATLE should be unbounded in magnitude. Again, this is stronger than necessary, and it may be shown (Gaines 1967* 1969) that the weights need take only a finite range of values, but that the range necessary for convergence is greater than the range necessary to ensure that a solution exists. When the range of weights is adequate for separation, but not adequate for convergence, then it is possible for tha ATLE, with a given training sequence, to never reach a final solution but stick in sub-optimal states with the weight values going through a repetetive cycle.

For example, the set of stimulus-vectors:

$Y^1 = (1, 1, 1, -1)$ $\qquad$ $Y^2 = (1, -1, -1, 1)$

$Y^3 = (-1, 1, -1, 1)$ $\qquad$ $Y^4 = (-1, -1, 1, 1)$

May be separated from the complementary set:

$Z^1 = (-1, -1, -1, 1)$ $\qquad$ $Z^2 = (-1, 1, 1, -1)$

$Z^3 = (1, -1, 1, -1)$ $\qquad$ $Z^4 = (1, 1, -1, -1)$

by the weight vector, W = (1,1,1,2); because -

$$\sum_{i=1}^{4} W_i Y_i^J = 1 > 1/2 > -1/2 > -1 = \sum_{i=1}^{4} W_i Y_i^J$$

so that, for $\Theta = 1/2$, Equations $\boxed{A1.2}$ and $\boxed{A1.3}$ for correct classification are satisfied.

However, if the components of the weight vector are limited to the range, $-2 \leq W_i \leq 2$, then Novikoff's convergence proof no longer applies, and equations $\boxed{A1.5}$ through $\boxed{A1.7}$, modified to take account of the 'limiting', do not necessarily lead to a solution. Thus, given the training sequence consisting of $(Y^1,Y^2,Y^3,Y^4,Z^1,Z^2,Z^3,Z^4)$ repeated indefinitely, and starting with initial weight values of zero, the weights take the following values:-

		W_1	W_2	W_3	W_4
W(0)		0	0	0	0
W(1)	Y^1	1	1	1	-1
W(2)	Y^2	2	0	0	0
W(3)	Y^3	1	1	-1	1
W(4)	Y^4	0	0	0	2
W(5)	Z^1	1	1	1	1
W(6)	Z^2	2	0	0	2
W(7)	Z^3	1	1	-1	2
W(8)	Z^4	0	0	0	2
W(9)	Y^1	1	1	1	1
W(10)	Y^2	2	0	0	2
W(11)	Y^3	1	1	-1	2
W(12)	Y^4	0	0	0	2
W(13)	Z^1	1	1	1	1
W(14)	Z^2	2	0	0	2
W(15)	Z^3	1	1	-1	2
W(16)	Z^4	0	0	0	2

Hence the weight vector goes through a repetetive cycle, always terminating with the value (0,0,0,2), which is not a linear separator for the dichotomy. It is shown in Section 2.2.5, that a sequence of stimuli constitutes a 'task' for an adaptive pattern classifier, so that the behaviour shown above has been elicited by giving the ATLE the same task repeated. Clearly the behaviour if not 'acceptable', in the sense of Section 2.2.8, because the ATLE never gives a correct response and cannot be 'satisfactory'. The state-transitions elicited by the task, in the example given, are from (0,0,0,0) to (0,0,0,2), and from (0,0,0,2) to itself - both these states are outside the region of potential adaption to the task.

ATLE with bounded weights have a richer range of behaviour than do those with unbounded weights, since they do not necessarily converge even when a solution vector exists within the range of the weights, and their convergence becomes a function of their initial state. Hence, whereas by Novikoff's result the region of potential adaption for an unbounded ATLE is either the whole state-space, or it is empty, for a bounded ATLE the region of potential adaption may be a proper sub-set of the state-space - in the example given, (1,1,1,2) is within the region, whereas (0,0,0,0) is outside it. Al the various modes of adaptive behaviour defined in Chapter 2 and 3, together with the various training techniques possible, may be illustrated with ATLE, and a comprehensive experimental study of one particular ATLE is described in Chapter 6.

A1.4 ATLEs in Control Systems

So far the ATLEs have been considered only as pattern-classifiers, not as controllers, since there is no feedback from the output of the ATLE into the environment. However, there is no reason why the stimulus generator of Figure A1-1 should not be a plant of some form which is to be controlled by the output of the ATLE - in this event, the stimuli might be the error and error-rate, and the output of the ATLE might operate an incremental actuator. Figure A1-2 shows this type of configuration, and also emphasizes that performance evaluation is not necessarily obtained by mimicing another controller, although this remains possible. However, the immediate and definite performance evaluation of each output of the ATLE is not normally available in a control situation. Not only is the performance measure averaged over many decisions, as, for example, the root-mean-square error, but its optimum value is also unknown. Thus the transition from the configuration of Figure A1-1 to that of A1-2 involves a change in the mode of learning of the ATLE from mimicing to closed-loop adaption.

If the ATLE should be required to mimic another controller, then the problem becomes very similar to that of pattern classification, except that there is feedback from the decisions made to the stimuli encountered. If it is the actions of the exemplifying controller which are fed back to the environment, then the only sources of difficulty are linear inseparability, and a restricted set of patterns; Widrow and Smith (1964) have successfully applied this technique to a second-order control system, equivalent to balancing an inverted pendulum by moving the base horizontally, using the human operator as the exemplifying controller However, if it is the actions of the ATLE which are fed back into the environment, then the sub-environment generated by it may cause the exemplifying controller to operate outside its normal range of inputs and to show incorrect behaviour.

A1.4.1 Performance Feedback to ATLE Controllers

When no exemplifying controller is available, then the criterion for adaption must be based on typical performance measures, such as the minimization of some error functional over an interval. This creates two sources of difficulty, in that the performance measure is global over a sequence of decisions, rather than local to each individual one.

Widrow (1966) has proposed a technique, called 'bootstrapping', for coverting a global performance criterion, in which a complete sequence of behaviour is evaluated as good or bad, into a local one. During an interaction, no change is made in the weights of the ATLE until a decision is available that the interaction, over some past interval, has been good or bad. This evaluation is then applied to all input/output pairs which have occurred during that interval, and the inputs are effectively presented again to the ATLE and its outputs are rewarded or punished accordingly; all input/output pairs must be stored until the performance evaluation becomes available.

Under bootstrapping, the evaluation of individual outputs of the ATLE may be incorrect and their probability of occurrence may be increaed when it should be decreased, and vice versa, but, in the long run, under certain conditions, this procedure may be shown to lead to convergence to an optimal policy (Widrow 1966). Widrow has applied this technique to the game-playing situation of 'blackjack' with results close to those predicted theoretically. Quarmby (1968) has applied the same technique to the learning of the game of 'Nim', and also finds that it leads to an optimal solution. His work is of particular interest because he has varied the form of the environment, in fact a second player, and studied the effects of this on the speed of learning. These results, which correspond to the identification of the adaption automaton of an ATLE controller, are reviewed in the following section.

Bootstrapping alone does not overcome the problem of lack of knowledge about the optimum value of a performance measure. In a game-playing situation a win/lose decision is always ultimately available, but in a control situation a value of an error-functional is not necessarily good or bad. Even if the minimal value is known, this does not enable a good/bad criterion to be established unless some tolerable deviation from it is given. The problem can be appreciated by considering a pattern-classifier in a similar situation - if the input stimuli occur at random with equal probabilities of either type, then the pattern-classifier has equal possibilities of reward or punishment - if, however, it is rewarded only for optimal decisions over the last n patterns, then the probability of reward drops to 2^{-n}, and the classifier has to learn mainly through the negative information that it has made mistakes.

One way of overcoming these problems is to consider performance over an interval good if it is an improvement over performance over some preceeding, comparable interval. However, no success has been reported in the application of this technique to control problems, and negative results are reported in Chapter 6. An alternative technique, which is not generally applicable, is described in Chapter 6 in a successful application to a particular control problem. This involves the evaluation of individual outputs of the ATLE according to whether the error at some fixed interval in the future has decreased or not.

A1.4.2 Problems of the ATLE as a Learning Controller

Examined in more general terms, the evaluation of an ATLE controller's performance by comparison with its past performance is equivalent to setting it the sub-goal of attaining a certain performance level, and varying this sub-goal as a function of its past performance. Another technique for setting a sub-goal is reported by Widrow and Smith (1964) in which an ATLE controller drives a second-order plant, with a human operator deciding when negative or positive bootstrapping should be applied to the ATLE. He states that, 'it was found that an observer familiar with control system theory, but ignorant of the plant configuration, would consistently produce training sequences leading to stable system configurations'.

This is a particularly interesting technique for overcoming the learning problems of an ATLE, since it involves feedback training by a human observer varying the sub-goals set to the machine. The observer does not require detailed knowledge of either the controller nor the controlled element, but is able to evaluate the overall performance and the effect of his own task difficulty variations upon it. In general, such a direct variation of the sub-goals of a learning system is not possible, and some form of linguistic communication must be established. An experiment on such communication with ATLE controllers is described in Chapter 10, and again the importance of feedback to the trainer is illustrated.

Quarmby (1968) has investigated the learning behaviour of a bootstrapped ATLE playing Nim against an opponent whose strategy was to play the optimum move with probability, p, and to play a random move with probability, 1-p. When playing against a fixed opponent, the mean number of games for the ATLE to reach an optimal solution, as a function of p, was:-

p =	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Number of games	723	281	104	124	68	51	49	36	35	66

Thus, from the point of view of the time taken to learn the task under conditions of fixed training (Section 3.2.1), the task increases in difficulty on either side of the value $p \simeq 0.85$. However, from the point of view of what is the most difficult opponent, in the sense of the 'best' opponent who will win most often, it is clear that the 'difficulty' of the task increases as p ranges from zero to unity. This illustrates the problems in the use of the concept of 'difficulty', discussed, in the context of human learning, in Section A4.5.1. Quarmby also evaluated the effects on learning of using the open-loop training strategy of increasing p from zero to unity uniformly over a set of M games (that is, increasing the difficulty, in the sense of a better player), and obtained the following results:-

M =	13	17	21	25	29	33	37
Number of games	30	29	28	27	27	34	35

One of the most interesting of Quarmby's results is that when two ATLE controllers play one against the other, the mean number of games for one to reach an optimum strategy is 18. This compares favourably with the 27 games for the best open-loop training policy used, and 35 games for the best fixed training policy used. The use of an opponent as part of the training environment who is also a learning system is clearly a form of feedback training, but of a very complex kind. Although Quarmby's experiments are not conclusive in themselves, they are highly suggestive of further experiments relevant to human learning, and of possible strategies for feedback training.

The main problem of ATLEs in learning control tasks may be related both to the 'generalization' properties of adaptive threshold logic itself, and to the sub-environment phenomenom. The sub-environment of a non-optimal control policy, particularly one which causes the overall system to become unstable, will generate some sub-set of the total set of input patterns to the ATLE. Because of the linear separability constraint upon the dichotomies realizable by the ATLE, any dichotomization of these patterns partially determines the response to other patterns, and a policy which is established in the initial sub-environment may imply a completely incorrect set of responses in the desired sub-environment.

APPENDIX 2: THE ALGEBRAIC THEORY OF SEMIGROUPS

A2.1 Semigroups

The basic literature on the algebraic theory of semigroups is slight, comprising two volumes by Clifford and Preston (1961,1967), one by Liapin (1963) translated from his original text in Russian published in 1960, and a recent volume on the application of semigroups in automata theory edited by Arbib (1968).

A semigroup (Clifford and Preston 1961, p.1) is a non-empty set, S, together with an associative binary operation defined on S. A binary operation is a mapping from S x S into S, and it is convenient to write the image of the ordered pair, (a,b), as the product ab; where a, b, ab ε S. If the operation is associative then:

$$\forall\, a,b,c \in S, \qquad a(bc) = (ab)c \qquad [A2\text{-}1]$$

Within a semigroup, S, there may be an element, 1, such that:

$$\forall\, a \in S, \qquad 1a = a1 = a \qquad [A2\text{-}2]$$

and such an element, if it exists, is called the identity element of S. It is possible for an element to exist which is an identity only for left (or right) operation - e is a left identity element of S if:

$$\forall\, a \in S, \qquad ea = a \qquad [A2\text{-}3]$$

and f is a right identity element of S if:

$$\forall\, a \in S, \qquad af = a \qquad [A2\text{-}4]$$

However, if S has both a left and a right identity element, then they are identical since:

$$e = ef = f \qquad [A2\text{-}5]$$

If a semigroup, S, has no identity element then it is always possible to add to it an element defined to have this function, and the extended semigroup is denoted, S^1.

Similarly, within a semigroup, S, there may be an element, 0, such that:

$$\forall\, a \in S, \qquad 0a = a0 = 0 \qquad [A2\text{-}6]$$

and such an element, if it exists, is called the zero element of S. Right and left zero elements may be defined as for identity elements, the two zero elements coinciding if they both exist, and, similarly, a zero element may be appended to a semigroup and the result denoted by S^o.

A2.2 Free Semigroups

An arbitrary set, D, is said to generate a free semigroup, F_D, by concatenation, which consists of all sequences of elements of D. The product of two sequences of elements of D, u and v, is defined to be the concatenation of these sequences, uv - hence, if $u=d_1d_2$ and $v=d_3d_4d_5$, then $uv=d_1d_2d_3d_4d_5$; concatenation is clearly associative. The identity element of F_D is the empty sequence, but no zero element is defined.

In a free semigroup, F_D, two elements, u and v, are equal if, and only if, they are identical sequences of elements of D. It is possible to obtain other semigroups from a free semigroup by superimposing on it generating relations amongst its members - that is by defining that:

$$\forall u_\lambda,\ v_\lambda \in S\ ,\ \lambda \in \Gamma\ ,\quad u_\lambda = v_\lambda \tag{A2-7}$$

where Γ is an index set.

An important free semigroup is generated by the set of partial transformations of a set, S, into itself. If u and v are mapping from two sub-sets of S to two other sub-sets, then the mapping, (uv), is defined if the range of v and the domain of u are not disjoint; if they are disjoint, then we may define a mapping, 0, whose domain and range are both empty, and define uv=0. Hence, the set of partial transformations of a set into itself, together with a mapping, 0, is a semigroup (with zero).

A useful notation for a member of a free semigroup is to let $(d_1+d_2+d_3)^*$ mean 'any sequence of elements consisting of d_1, d_2 and d_3. Then, $d_1(d_2)^*d_3$ means, 'd_1 followed by d_2 repeated any number of times (including zero), followed by d_3.

A2.3 Homomorphisms and Relations

The natural mappings from one semigroup to another, those which preserve the semigroup structure, are called homomorphisms. A mapping between semigroups, S and S', ϕ : $S \rightarrow S'$, is a homomorphism if, and only if:

$$\forall a,b, \in S \qquad (ab)\phi \quad = \quad (a\phi)(b\phi) \qquad [A2-8]$$

(where the operator is written on the right. A one-to-one homomorphism has an inverse which is also a homomorphism and such a mapping is called an isomorphism. Isomorphic semigroups are complete identical in their semigroup structures, and, for theoretical purposes, need not be distinguished.

A relation on a set, S, is some sub-set, Θ, of the product of S with itself, SxS. If $(a,b) \in \Theta$, where a and b are elements of S, then we may write - $a\Theta b$, meaning 'a bears the relation Θ to b'. The composition of two relations on S, Θ and η, is defined as - $(a,b)\varepsilon\Theta\eta$ if, and only if, there exists c such that $(a,c)\varepsilon\Theta$ and $(c,b)\varepsilon\eta$. The set of all such relations on S is a semigroup. Its identity element is the equality relation, I, such that $(a,b)\ \varepsilon I \longleftrightarrow a=b$. Its zero element is the empty relation on S.

The converse, Θ^{-1}, of a relation, Θ, is defined by - $(a,b)\ \varepsilon\Theta^{-1} \longleftrightarrow (b,a)\varepsilon\Theta$. A relation, Θ, is said to be reflexive if $I \subseteq \Theta$; symmetric if $\Theta \subseteq \Theta^{-1}$; and transitive if $\Theta\Theta \subseteq \Theta$; an equivalence if it is relfexive, symmetric and transitive. An equivalence relation partitions S into a mutually disjoint family of sets, and the natural mapping from S into its equivalence sets gives rise to the quotient set, S/Θ, of S under Θ.

If S is a semigroup, the natural mapping to a quotient set under an arbitrary relation is not necessarily a homomorphism, and major theorems in semigroup theory are concerned with determining when this is so. A relation, Θ, on a semigroup, S, is a congruence if, for $a,b\ \varepsilon S$, $a\Theta b \longleftrightarrow uav\Theta ubv$, for all $u,v\ \varepsilon S$, and Θ is an equivalence (if Θ is not an equivalence, the relation is aid to be regular). If Θ is a congruence then the quotient mapping from S to S/Θ is a homomorphism.

There are certain important theorems relating to homomorphisms and relations which will be quoted here with the numbering of Clifford and Preston (1961), where the proofs are given:

Theorem 1.5 (Main Homomorphism Theorem) Let Θ be a homomorphism of a semigroup, S, upon a semigroup, S', and let $\pi = \Theta.\Theta^{-1}$. Then π is a congruence on S, and there exists and isomorphism, μ, of S/π upon S' such that $\pi^*\mu = \Theta$, where π^* is the natural homomorphism of S upon S/π.

Theorem 1.6 (Induced Homomorphism Theorem) Let ϕ_1 and ϕ_2 be homomorphisms of a semigroup, S, upon semigroups, S_1 and S_2 respectively, such that $\phi_1 \cdot \phi_1^{-1} \subseteq_2 \cdot \phi_2^{-1}$. Then there exists a unique homomorphism, Θ, of S_1 upon S_2 such that $\phi_1\Theta = \phi_2$.

Lemma 1.28 Let F_D be the free semigroup generated by a set, D. Let S be any semigroup and let ϕ_o be any mapping of D into S. Then ϕ_o can be extended in one and only one way to a homomorphism from F_D to S.

Theorem 1.29 Let F_D be the free semigroup generated by a set, D. Let π_o be any relation on F_D and let π be the congruence relation on F_D generated by π_o. Let π^* be the natural homomorphism of F_D upon F_D/π. Let S be any semigroup and let ϕ be a homomorphism of F_D into S such that $u\phi = v\phi$ for every $(u,v) \in \pi_o$. Then there exists a homomorphism Θ of F_D/π into S such that $\pi^*\Theta = \phi$.

A2.4 Ideals

A left (right) ideal of a semigroup, S, is a non-empty sub-set, A, for S such that $SA \subseteq A$ $(AS \subseteq A)$. A two-sided ideal (or simply ideal) is a sub-set of S which is both a left and a right ideal. A semigroup, S, is called left (right) simple if S itself is the only left (right) ideal of S. Similarly a semigroup is called simple if it contains no proper (two-sided) ideal.

If X is a non-empty sub-set of a semigroup, S, then the intersection of all left ideals of S containing X is a left ideal of S containing X, and contained in every other such left ideal of S - it is called the left ideal of S generated by X; similarly for the right ideal of S generated by X and the (two-sided) ideal of S generated by X. If, in particular, X is the single element, x, then the ideals generated by it are called principal ideals generated by x.

An ideal is called 0-minimal if it contains elements other than zero, and the only ideal properly contained in it is the zero element.

APPENDIX 3 FROM BEHAVIOUR TO STRUCTURE

A3.1 The Relationship Between Structure and Behaviour

The relationship between the physical structure of a system and its observed behaviour, and to what extent one can be deduced from the other, have long been controversial topics in both philosophy and psychology. The fundamental problem of knowledge of the 'real world' inferred from sensations is the prime example of the difficulty in establishing this relationship in general, but problems of describing, modelling and predicting arise continually, not only in science, but also in everyday life. Whilst knowledge of 'reality' is a source of absolute problems because the observed behaviour is, for the individual, all that exists, other epistemological problems arise through comparison of different sources of information about the same physical structure, particularly when one source of 'information' is a set of pre-conceived assumptions.

In this appendix the problem of relating behaviour and structure is formalized, and a technique for deriving one from the other is established. The problem is treated in a general way because the results are of importance in the present context, not only in the analysis of adaptive behaviour and techniques of training, but also, at a different level of discussion, in the study of the epistemological problems of learning systems which give rise to their adaptive behaviour.

A3.1.1 Extensive and Intensive Definitions of a System

The concepts of 'behaviour' and 'structure' may be formalized by means of the logical constructs of definition by extension and definition by intension (Carnap 1956). A property is said to be defined extensively by the class of all those objects which possess the property. A property is said to be defined intensively by a rule, or decision procedure, which determines whether an object has the property. These constructs may be used to give formal definitions of a system, which correspond to its structural and behavioural connotations, respectively -

Extensive definition of a system A system is defined extensively by the class of all possible behaviours which may be shown by the system.

<u>Intensive definition of a system</u> A system is defined intensively by a rule which determines whether a particular behaviour is possible for the system.

One possible decision procedure, or rule, which defines a system by intension is that a behaviour is possible for that system if, and only if, it may be shown by another. This second system may be defined in any way whatsoever, provided the definition enables its behaviour to be generated. The 'structure' of a system, divorced from notions of physical 'reality', is nothing more than a set of rules for determining the behaviour of the system, and hence may be regarded as an intensive definition of the system. In these terms, the problem of the relationship between the behaviour and structure of a system amy be regarded as one of deriving an intensive definition from an extensive one.

A3.1.2 Cybernetic Equivalence Between Systems

To each of the two types of system definition there corresponds an equivalence relation between systems - two systems may be said to be co-extensive if they show the same class of behaviours - two systems may be said to have the same intension if the rules which define them are logically equivalent. Wiener (1914) has proposed, in the context of the analysis of logical relations, that extensive definitions be used in order to simplify the analysis of equivalence between relations: if a relation is identified with the class of n-tuples which satisfy the relation, then equivalence between relations can be evaluated by the theory of sets and requires no study of the logical structure of relations. In a later publication, Wiener (1948) extends this concept to arbitrary systems, and defines two systems to be <u>cybernetically equivalent</u> if they show the same behaviour, that is, if they are co-extensive.

The most important feature of Wiener's approach is that he proposes that it does not matter what structure we suppose for a system, if we are only interested in its behaviour, provided it is one of the set of cybernetically equivalent structures which would give rise to its observed behaviour. Deutsch (1960) has argued for the application of Wiener's approach in the study of animal behaviour. He states that -

'the central nervous system could be constructed of completely different types of components without affecting the behavioural capabilities of the machine' - and that -

'given the system or abstract structure alone of the machine, we can deduce its properties and predict its behaviour. On the other hand, the knowledge that the machine operates mechanically, electro-mechanically, or electronically does not help us very much at all'.

The logical conclusion of Wiener's arguments about the primacy of the extensive definition of a system by its behaviour might seem to be that 'structure' in itself may be neglected. However, Ashby (1965) has demonstrated the importance of the concept of 'structure' in enabling a vast, possibly infinite, number of instances of behaviour to be subsumed under the single statement of a rule, corresponding to a system structure. For example, the behaviour of a device, whose input is drawn from the field of real numbers and whose output is its input plus unity, consists of an infinite set of input/output pairs with the form - (x, x+1). Rather than tabulate all these pairs in order to define the system, it is far simpler to state the rule -

$$S \equiv \left[\langle\text{output}\rangle = \langle\text{input}\rangle + 1\right] \qquad [A3.1]$$

which constitutes a definition of the system, S, as a mathematical operator, or structure.

Consider a second system, S*, whose structure is -

$$S^{*} \equiv \left[\begin{array}{lcl} \langle\text{output}\rangle & = & x^2 - y^2 \\ x & = & \langle\text{input}\rangle + 1.25 \\ y & = & \langle\text{input}\rangle + 0.75 \end{array}\right] \qquad [A3.2]$$

Although S and S* are structurally dissimilar, they are cybernetically equivalent in that, given the same input, they will both produce the same output. This equivalence may be demonstrated by manipulation of their structural definitions, given a knowledge of the algebra of real numbers. Wiener's argument is that this knowledge is unnecessary, and that the two systems may be proved equivalent by placing the input/output pairs which constitute their behaviour in a 1-1 correspondence.

This, conceptually simple, set-theoretic procedure may be used when the system structures, or the rules governing them, are completely unknown. It is of greatest interest, however, when the structure of one system is unknown, but, by analysis of its behaviour, it may be shown to be cybernetically equivalent to another system on know structure. The second system may then be taken to subsume the behaviour of the first, or act as a 'model' for it, with the advantages in simplicity of description demonstrated by Ashby. Since the structure of the model is known, statements may be made about the properties of its behaviour without detailed examination of that behaviour, and these statements may be applied immediately to the behaviour of the cybernetically equivalent system whose structure is unknown.

A3.1.3 Mathematical Machines as Models

No one model has inherent precedence over another that is cybernetically equivalent, and additional criteria have to be applied to select a particular form of model from all those available. For biological systems, it is natural to give physiological structures primacy in the modelling of animal behaviour, but, for practical purposes,these have the disadvantage that the behaviour of large structures of neurons, for example, is difficult to determine, both because they are nonlinear elements and because their individual structures and connectivity are poorly known. A similar situation arises in the study of the behaviour of gases, where the statistical mechanics of the behaviour of individual molecules may be used to derive the thermodynamics of marco laws of behaviour, such as Boyle's law relating the pressure and volume of an enclosed gas. Boyle's law, however, was discovered and used long before the properties of individual molecules were known, and, similarly, 'laws of behaviour' may be derived long before that behaviour can be ascribed to a physiological structure.

One form of model which it is reasonable to propose is that which has no properties other than that of showing the required behaviour. An abstract mathematical model has this feature, together with the advantage that it is readily communicated in written form. Ashby (1957) has proposed and developed the theory of <u>state-determined machines</u> to provide a mathematical object which has forms cybernetically equivalent to any other system: such a 'machine' is, conceptually, a device with states, inputs and outputs, such that the present state and input determine the next state and output. It is of interest to note that the general-purpose digital computer has been developed as a machine which, by 'programming', may be made cybernetically equivalent to virtually any other system (Gaines 1968). The methematical theory of computers, called Automata Theory (McNaughton 1961), is essentially an extended form of Ashby's theory of machines. This correspondence between abstract 'machines' and computers has the advantage that models put forward in terms of the former may be physically realized in terms of the latter, and hence the behaviour of the models may be demonstrated.

A3.1.4 The Observability of Structural Variables

The structure which is derived from behaviour to serve as a model for the system producing the behaviour acts, in some sense, as an 'explanation' of that behaviour. Deutsch (1960) distinguishes this type of explanation, which he calls 'generalizatory', from an alternative form which he calls 'causal' - 'causal' explanations are those based on previously established physical laws. Brindley (1960) makes a similar distinction between explanations of human behaviour based on psychological observations, and explanation based on knowledge of physiological structure.

The distinction between 'causal' and 'generalizatory' explanations proposed by Deutsch is not fundamental, in that the laws of physics are themselves inductive generalizations from observations of matter and are not inherently more profound than the 'laws' of psychology. There is a sense in which physiological structures have a preferred status in psychology, but the preference is based on their being derived from alternative observations of the same system, not on any fundamental difference in logical status. The linking of observations on different parts of the same system is a source of difficulty, however, and it is reasonable to criticize the premature identification of 'intervening variables' in psychology with physiological constructs. There is no intrinsic reason, however, why a structure proposed from a cybernetic viewpoint, in that it will generate the observed behaviour, should not be partially, or wholly, identified with a physiological structure. This has occurred with Deutsch's own model of 'drive', where certain parts of the cybernetic structure may be identified with nuclei in the hypothalamus.

Deutsch's comments make it clear that the cybernetic structure derived from behaviour has no more predictive power than the behaviour itself, in the sense of being able to determine future behaviour. Given a knowledge of all possible behavioural sequences which may be exhibited by a system, the basis of a 'generalizatory' type of explanation, and a particular instance of a segment of observed behaviour, it is possible to match the segment with all possible sequences, delimiting the behaviour which may follow it, and hence predicting to some extent what that behaviour will be. If more of the past behaviour is known then the future behaviour may be further delimited, and hence the prediction will become more precise, or not change. Since the future behaviour of a physical system is governed by its present condition, knowledge of its past behaviour which affects predictions about future behaviour must be equivalent to more detailed information about its present condition. In the limit, it is possible that knowledge of the past behaviour may be used to specify the present condition of the system precisely.

These considerations suggest further criteria to be applied in selecting cybernetic models of behaviour. The only fundamental constraint on a structure which is to serve as a cybernetic model for observed behaviour is the very weak one, that it should show all the observed behaviour, and only that behaviour. For any given set of behaviour, there will be very many possible structures, and these will vary greatly in complexity. For example, any model can be increased in complexity without limit by the addition of internal processes which have no affect on the output, or by the addition of intervening variables within the model. The system, S', defined as -

$$S' \equiv \left[\begin{array}{lcl} \langle \text{output} \rangle & = & x - y \\ x & = & \langle \text{input} \rangle + z \\ y & = & z - 1 \\ z & = & (\langle \text{input} \rangle)^2 \end{array}\right] \qquad [A3.3]$$

is cybernetically equivalent to the system, S, of Section A3.1.2, but contains the redundant variables, x,y,z, and the redundant process defining z. This unnecessary complexity may be eliminated by requiring that any intervening variables be measurable, that is observable from past behaviour, and that the number of different internal conditions of the model is the minimum necessary to account for all the modes of behaviour. It will be demonstrated that these two conditions interact, in the sense that observability of all structural variables may necessitate more internal conditions than are necessary solely to account for the behaviour - a seemingly paradoxical result.

A3.1.5 Summary of Constraints Upon Structural Models

If one system is to act as a model for another then there is one necessary constraint upon it, and three desirable ones -

(i) Cybernetic equivalence - the model should be cybernetically equivalent to the system it is modelling: that is, there should be a 1-1 correspondence between the behaviours of the two systems.

(ii) Only sufficient properties - the model should have no properties, other than that of showing the observed behaviour of the system it is modelling. This is not essential, and mechanical or electrical models may be useful psychological models. An abstract mathematical model has the advantage, however, that there are no unnecessary features to confuse its evaluation.

(iii) Observability of structural variables - given a sufficiently long sequence of past behaviour, it should be possible to determine all internal parameters of the model precisely.

(iv) Minimum states - subject to contraints (i) and (iii), the number of possible values of the internal parameters of the model should be minimal.

In the following sections a procedure for constructing an automata-theoretic model of a system, given its behaviour alone, is derived which satisfies these constraints.

A3.2 The Behaviour of Automata

In the analysis of behaviour some formal explicatum of the concept of 'behaviour' itself is necessary. Since behaviour is essentially a sequence of observations, it is possible to provide an explicatum by setting up a calculus for the results of observations. This is done in the following postulates relating to the observation of a system, Σ. The first postulate is that the behaviour can always be described -

(i) There exists a set, D (the set of descriptors), such that a unique member of D may always be assigned to the system, Σ.

This ensures that the set of descriptors is sufficient and that a decision procedure exists for describing the behaviour with a unique descriptor: uniqueness ensures that problems of semantic relationships between descriptors in themselves do not arise.

The second postulate identifies the system, Σ, with the set of all behaviour that it may show -

(ii) A system, defined by extension, is a sub-set, Σ, of the free semigroup generated by concatenation of member of the set, D, such that every sub-sequence of a member of the sub-set is also a member of the sub-set.

This defines a system as a set of sequences of descriptions, or observations of behaviour (a sequence of descriptions being a 'behaviour'), and ensures that any part of an observed behaviour is also noted as an observed behaviour.

A3.2.1 The Semigroup of Descriptors

The term 'semigroup' used in postulate (ii) of the previous section requires further elucidation. A semigroup is a mathematical object, similar in structure to a group but with weaker postulates. Although semigroups logically precede groups, it is only during recent years that they have been studied in detail by mathematicians, whereas the study of groups has long been intense and the literature is vast. The theory of semigroups is fundamental to the study of automata and general systems, and will be used extensively in this thesis. A summary of the relevant theory and terminology is given in Appendix 2. In this section the relationship between systems and semigroups is established.

Let Σ be a system defined as in postulate (ii), such that $\Sigma \subset F_D$, the free semigroup generated by the set of descriptors, D. Then Σ, as defined, is not itself a semigroup because, given behaviours, a, b, $\varepsilon\ \Sigma$ it is possible that the behaviour, ab, does not belong to Σ - that is, it may not be shown by the system. However, it is possible to extend Σ by addition of the zero element such that it is a semigroup homomorphic to F_D.

<u>Theorem A3-1</u> If $\Sigma \subset F_D$ is a system, and Θ is the relation defined by -

$$a\Theta b \longleftrightarrow a, b \in F_D - \Sigma, \text{ or } a = b$$

then Θ is a congruence relation and the natural homomorphism from F_D onto the quotient semigroup, F_D/Θ, maps the set, $[F_D-\Sigma]$, into 0, the zero element.

<u>Proof</u> For any a, b, such that $a\Theta b$, and any $x \in F_D$, consider xa and xb. If a=b, then xa=xb, and hence $xa\Theta xb$. Otherwise $a, b \in F_D - \Sigma$. Suppose $xa \in \Sigma$, then a is a sub-sequence of a member of Σ, and hence by postulate (ii) $a \in \Sigma$. Thus, by contradiction, $xa \in F_D - \Sigma$; similarly, $xb \in F_D - \Sigma$, and hence $xa\Theta xb$. Similarly we may prove that $ax\Theta bx$.

Thus, Θ is a congruence relation and the quotient semigroup, F_D/Θ, and the natural homomorphism, Φ, from F_D to it, are defined. Let $a \in F_D - \Sigma$, and $b \in F_D$. Then $ab \in F_D - \Sigma$, and hence $a\Phi = (ab)\Phi$, under the quotient mapping, but Φ is a homomorphism so that - $(ab)\Phi = a\Phi.b\Phi$. Thus, $a\Phi = a\Phi.b\Phi$, and similarly - $a\Phi = b\Phi.a\Phi$, implying that a is the zero element of F_D/Θ.

This theorem shows very clearly the relationship between the extensive definition of a system and the algebraic properties of semigroups. A system, defined by extension, may be represented as the free semigroup generated by its descriptors, with zero adjoined, and a generating relation such that every sequence of descriptors which is not a possible behaviour is mapped into zero. Because every sub-sequence of a possible behaviour is also required to be a behaviour, this mapping is a homomorphism, the natural mapping between semigroups, and the resultant structure is itself a semigroup.

A3.2.2 Implications of Group Postulates

The relevance of semigroup structures to the study of system behaviour may be further clarified by consideration of the effects of strengthening the postulates, and representing the system by a group, rather than a semigroup, structure. The additional postulate for a group is that, for every element $a \in F_D$, there exists an element $a^{-1} \in F_D$ such that $a^{-1}a = 1$; a^{-1} is termed the 'inverse' of a. The part of the unity element, 1, in F_D is played by the empty sequence, and equality of sequences is defined only by their being identical. Hence, no inverse elements can exist in the free semigroup of descriptors, F_D, itself, because the above equation would imply that the succeeding behaviour, a^{-1}, makes an already observed behaviour, a, become unobserved. Equally, no inverse elements can exist in the quotient semigroup, F_D/Θ, defining the system, Σ , because the only additional relationship of equality is generated by the mapping into zero, and zero has no inverse.

If a further relationship of equality were defined in F_D, which implied, for example, that behaviours were equivalent which achieved the same goal, the additional group postulate would demand that any behaviour could be nullified in its effect on the attainment of the goal. This is obviously a very strong postulate implying a degree of reversibility which, whilst present in many systems studied in physics and chemistry, is not always found in animal behaviour, and is certainly very rare, almost by definition, in the learning process. This explains the necessity for the mathematics of semigroups, rather than that of groups, in the analysis of adaptive behaviour.

A3.2.3 Structure of Automata

One of the most general structures to have been investigated in recent years is the sequential machine, or automaton (Gill 1962, Ginsburg 1962, Moore 1964, Hartmanis and Stearns 1966, Booth 1967, Hennie 1968). Any physical system, examined at discrete intervals, can be represented as an automaton, and automata theorists are developing techniques for analysing general, nonlinear systems (Wymore 1967), which may eventually become as powerful as comparable techniques for linear systems at present (Zadeh and Desoer 1963). Since the discrete, decision-making mechanisms of animal behaviour are not amenable to analysis with linear systems theory, the development of automata theory is particularly important in biology. In particular, an automaton structure is an adequate representation of a biological system, and the problem of relating behaviour and structure may be reduced to that of determining an automaton whose behaviour matches that of the observed system.

An automaton is a device with inputs, states and outputs, whose present state and input determine its next state and output. This may be formalized (Mealy 1955) in the following terms - an automaton, or sequential machine, is characterized by -

(i) A set S of states.

(ii) A set I of inputs.

(iii) A set G of outputs.

(iv) A mapping, $\sigma: I \times S \quad S$, called the next-state function.

(v) A mapping, $\pi: I \times S \quad G$, called the output function.

Hence a particular machine may be characterized by the 5-tuple, $(I.S.G.\sigma.\pi)$. The dynamic behaviour of the machine is determined by the following transition equations:-

$$s' = \sigma(i,s) \tag{A3.4}$$

$$g = \pi(i,s) \tag{A3.5}$$

- where s is the present state, s' is the next state, and g is the present output.

A3.2.4 Descriptor Semigroup of an Automaton

Since the state of an automaton is an internal variable which may not be directly observed, its overt behaviour may be completely described in terms of its inputs and outputs. In the terminology of the previous section, a complete description of the behaviour of the automata in its present condition is -

$$d \equiv (i,g) \tag{A3.6}$$

It is apparent that a set of sequences of such descriptors, corresponding to a set of observed behaviours of the automaton, obey the postulates for an extensive definition of a system defined in Section A3.2. It is also apparent that any restrictions on the initial state of the automaton, or upon the input which may be applied when a certain output is present, will generally restrict the observed behaviour to some sub-set of all possible behaviour, and hence will give rise to an extensive definition of a different system. This is important in the context of the problems faced by any learning system in attempting to learn about its environment (which may be regarded as an automaton), and is analysed in Section 3.5.

For any description of the behaviour of the automaton, $d \equiv (i,g)$, there will be a set of states, R, in which the input, i, may be applied. Consider the sub-set of R, $Q \subset R$, such that the output, g, occurs -

$$Q \equiv s : s \in R, g = \pi(i,s) \quad [A3.7]$$

Let Q be the domain of a mapping, M_d, whose range, T, is the set of states which may follow the state, $s \in Q$, and the input, i, so that -

$$T \equiv [s' : s \in Q, s' = \pi(i,s)] \quad [A3.8]$$

The mapping, $M_d : Q \longrightarrow T$, is then defined -

$$s' = M_d(s) \quad [A3.9]$$

- where $s \in Q$, and $s' = \sigma(i,s)$, $g = \pi(i,s)$.

Thus, with every description, d, of the behaviour of the automaton, $M \equiv (I,S,G,\sigma,\pi)$, it is possible to associate a unique mapping, M_d, from the class of partial transformations of the set, S, into itself. This association may be written as the mapping, λ, such that -

$$M_d = d\lambda \quad A3.10$$

- where the range of λ is M, some sub-set of S^T, the set of partial transformations over S.

This mapping does not, in itself, demonstrate a relationship between the semigroup of partial transformations over S, and the semigroup of descriptors defining the behaviour of the automaton; an arbitrary mapping between semigroups does not necessarily imply any similarity between their structures. However, the following theorem shows that the mapping, λ, is an isomorphism, so that the semigroups of partial transformations is identical in structure with the semigroup of descriptors.

Theorem [A3-2] The free semigroup of partial transformations over S, F_M, generated by M_d for all $d \in D$, is isomorphic to the quotient semigroup, F_D/θ of Theorem [A3-1], under the transformation, λ.

Proof λ maps the generating set, D, of the free semigroup, F_D, into the free semigroup generated by the set, $M \subset S^T$, the set of partial transformations over S. Hence, by Lemma 1.28 of Clifford and Preston (1961), λ may be extended to be a unique homomorphism from F_D to F_M by setting -

$$(abc\ldots n)\lambda = M_a M_b M_c \ldots M_n = a\lambda . b\lambda . c\lambda \ldots n\lambda \quad [A3.11]$$

Consider a behaviour, $u = ab \in F_D - \Sigma$. The absence of ab from the set of possible behaviours implies that the input-output pair corresponding to b cannot follow that corresponding to a. Thus, for every possible state in which the machine may be after having shown the behaviour, a, (all states in the range of M_a), either the input corresponding to b cannot be applied or the output corresponding to b does not occur. In either event, these states cannot lie in the domain of M_b and hence the range of M_a and the domain of M_b do not overlap, so that the mapping,

M_aM_b , takes its domain into the empty set, Φ .

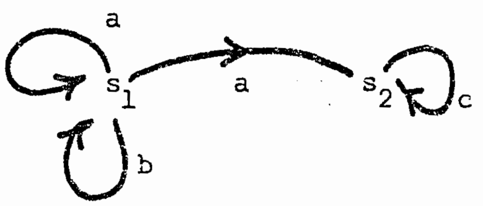
Any mapping in F_M which takes its domain into the empty set plays a part of the zero element in the semigroup of partial transformations, and hence we have that any behaviour, $u \varepsilon F_D - \Sigma$, is mapped into 0 by λ. Thus, considering the relation, Θ , of Theorem A3-1 , if $u\Theta v$, then either $u=v$ so that $u\lambda = v\lambda$, or $u,v \varepsilon F_D - \Sigma$ so that $u\lambda = v\lambda = 0$. Hence, by Theorem 1.29 of Clifford and Preston (1961), (see Appendix A2.3), there exists a homomorphism, λ' , of F_D/Θ into F_ST, such that $\Theta\lambda' = \lambda$. By the nature of Θ and λ, λ' is clearly λ itself extended over F_D/Θ . The mapping between F_D/Θ and F_{ST} is one-to-one and hence λ' is an isomorphism.

Thus the two semigroups, F_D/ Θ, corresponding to the extensive definition of the automaton by its behaviour, and $F_M \subset F_{ST}$, corresponding to the intensive definition of the automaton by its structure, are completely equivalent. Thus the problem of determining an automaton, cybernetically equivalent to a system defined by extension, may be regarded as one of finding a suitable mapping from the descriptors of the system to the set of partial transformations over some set, such that there is an isomorphism between the semigroup structures on them both.

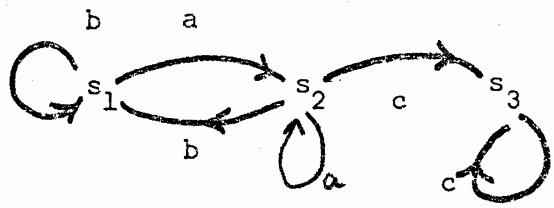
A3.2.5 Choice of Structure for a System Defined by Extension

There are many possible automata, with differing numbers of states and differing transition equations, which are cybernetically equivalent to any given automaton. Without further constraints there is no basis for selecting between these automata. One possible constraint is to demand that the automaton have the minimum number of states necessary to show the required behaviour, but this may lead to a structure which is peculiar in that no amount of information about its past behaviour will allow one to deduce its present state; in control-theoretical terms, the automaton has unobservable states (Kalman 1960, Wymore 1967 p.277).

For example, consider the two automata whose state transitions and corresponding behaviour are shown in Figure [A3-1] as directed graphs; each node of the graph corresponds to a state of the autotaton, and each line of the graph coreesponds to a possible transition between states, the emitted behaviour being indicated by a letter (descriptor).



(i). Minimum State



(ii) Observable

Figure A3-1 Different Structures for Equivalent Automata

Any behaviour of either of the two automata is of the form - (a+b)*a c* (in the notation of Appendix 1), and hence both systems show the same range of possible behaviour and are cybernetically equivalent. The first system, (i), has only two states, which is clearly the minimum necessary to restrict behaviour to this form. Given that the past behaviour is of the form - (a+b)*a, however, it is impossible to determine whether the automaton is in state s_1, or s_2. With the second system, (ii), any behaviour terminating in a leaves it in s_1; behaviour terminating in b leaves it in s_1 ; and in c leaves it in s_3. Thus any sequence of past behaviour is sufficient to determine its present state.

Thus although the second system, (ii), has one more state than system (i), all its states are observable, and it is possible to determine its current state from its past behaviour. This condition not only gives an operational definition of the 'states' of the hypothetical structure giving rise to the observed behaviour, in that all 'intervening variables' are measurable, but leads to what is, in many ways, a more 'realistic' structure. For example, it appears from the transition diagram of system (i) that behaviour, c, occurs only when the automaton is in state, s_2, and hence that s_2 is particularly informative about the occurrence of c. However, since no sequence of past behaviour terminating in a, no matter how long, is sufficient to determine whether the system is in s_1, or s_2, this apparent 'information' is rather misleading. On the other hand, the alternative structure, (ii), makes it immediately apparent that behaviour, c, may be emitted when the automaton is in either of states, s_2 or s_3, but that behaviour, b, may also be emitted when it is in s_2.

This further exemplifies the desirable constraint, (iii), proposed in Section A3.1.5, that all internal parameters of the model, its state variables, should be measurable from a sufficiently long sequence of observations of past behaviour. Given that this contraint upon the automaton structure is satisfied, it is reasonable to further impose contraint (iv) of that section, and minimize the number of internal states of the automaton. In the following sections, a construction is established for deriving an automaton structure for a system defined by extension, which satisfies these constraints.

A3.3 Construction of a Minimal, Observable Automaton

A set of observable states for a system defined by extension may be determined by noting that the present state of an automaton is essentially that which contains all the known information about its future behaviour, and considering the manner in which knowledge of past behaviour restricts future possible behaviour.

Consider a system defined by extension as a sub-set, Σ, of the free semigroup, F_D, generated by the set of descriptors, D. Any behaviour, $u \in F_D$, generates a sub-set of Σ, $N_u \subset \Sigma$, defined to be -

$$N_u \equiv \{v : v \in F_D,\ uv \in \Sigma\} \qquad [A3.12]$$

- so that N_u is the set of potential future behaviours which may follow the behaviour, u. It will be noted that if u does not belong to Σ, the N_u is the empty set, Φ.

<u>Theorem</u> [A3-3] For all $u,v \in F_D$, $N_{uv} \subset N_v$.

<u>Proof</u> For any $w \in N_{uv}$, we have by definition that $uvw \in \Sigma$. By postulate (ii) for a system (Section A3.2), the sub-sequence of uvw, vw, also belongs to Σ. Hence $w \in N_v$, and thus $N_{uv} \subset N_v$.

This result implies that the addition of successive descriptions of the past behaviour of the system creates an ordered sequence in the lattice of sub-sets of F_D, ordered by inclusion. We will assume that this sequence is bounded, and hence that sets of the form N_{uv} eventually stabilize so that -

$$\forall t \in F_D, \text{ either } N_{tuv} = N_{uv}, \text{ or } N_{tuv} = \Phi .$$

This is stated formally in the third system postulate -

(iii) To any descriptor, $d \in D$, there corresponds a family of sub-sets of F_D, $d\eta$, possibly including the empty sub-set, Φ, such that -

- if $N \varepsilon d\eta$, $u \varepsilon F_D$ -

(1) there exists $v \varepsilon F_D$, such that, for all $t \varepsilon F_D$

such that $tvd \varepsilon \Sigma$, $N_{tvd} = N$.

(2) there exists $N' \varepsilon d\eta$, such that $N_{ud} \supset N'$.

The first part of the postulate only requires that all members of $d\eta$ are actually attained bounds of the sequences created by behaviour terminating in d - that is, the sub-sets of possible future behaviour are as small as possible, and none are included which do not occur. The second part of the postulate ensures that the members of $d\eta$ are sufficient always to bound the effects of prior knowledge about the behaviour of the system.

This postulate is not inherent in the general concept of a system, but is a reasonable one, being implied by many other contraints upon the system. For example, if the system has a finite memory span so that information about the sufficiently remote past is irrelevant to its future behaviour, or if it shows only a finite number of possible behaviours, then postulate (iii) is satisfied. This last condition is always implied in practice, because only a finite number of observations may be made before a model is formed. It is interesting to note that the postulate is not implied by the descriptor set, D, being finite in number. The justification for the postulate, in the present context, is that it is satisfied in all cases of interest, for one reason or another, and that systems which do not satisfy it generate an infinite variety of behaviour and are, thus, not subject to complete experimental observation.

A3.3.1 Assignment of States

The concept of the 'state' of a system plays a very important role in modern control and general systems theory. It was first formalized by Poincaré (1885), in the context of dynamical systems and thermodynamics, in the late nineteenth century. His definition was naturally based on structural considerations, being a set of parameters defining the positions and momenta of a system of particles, and it is only in recent years that the purely behavioural connotations of the concept have been studied. In particular, Zadeh (1964) has analysed the conditions for the assignment of values to the state-variables of a system to be consistent with given behaviour of that system.

Zadeh takes the content of the concept of state to be - 'a number or set of numbers, which collectively contains all the information about the past of the system that is relevant to the determination of its future behaviour'

In a structural context, the definition of the 'state' of an automaton in Section A3.2.4 clearly satisfies this statement - it is a variable which, together with the sequence of future inputs to the automaton, is sufficient to determine all future outputs, and all future states, of the automaton. In a behavioural context, Theorem A3-3 demonstrates how further information about the past behaviour of a system restricts its potential future behaviour, and hence must reduce the set of states in which the system may be after the observed behaviour.

The problem of state-assignment thus consists of giving a set of possible values, or designators, to a system which contains all the known information about its future behaviour. Thus, if S is an abstract set and Σ is a system defined extensively as a sub-set of the free semigroup, F_D, generated by the descriptor set, D, then a <u>state assignment</u> to the system is a mapping from F_D to the family of sub-sets of S, 2^S, which satisfies certain consistency requirements. These are set out in the following definition.

A3.3.2 Definition of a Consistent State Assignment

The mapping, $\gamma : F_D \longrightarrow 2^S$, is a consistent state assignment if, and only if -

(i) $u \in \Sigma \longleftrightarrow u\gamma \neq \Phi$.

(ii) $\forall u,v \in F_D, \quad (uv)\gamma \subset v\gamma$.

(iii) $\forall u,v \in F_D, \quad u\gamma \cap v\gamma \neq \Phi \longrightarrow \exists t \in F_D : ut, vt \in \Sigma$.

The first requirement implies that there is no state of the automaton which can follow impossible behaviour. The second is less trivial, and implies that further information about past behaviour can only reduce the possible present states of the automaton. The third requirement implies that if the same state is possible after two different behaviours then there is at least one possible future behaviour common to both. This last requirement is stronger than is strictly necessary - it is sufficient to require that the state sets assigned to behaviours with differing future possible behaviours are not identical. However, the only effect of weakening the requirement is to allow states to be assigned which the automaton can never enter, a trivial possibility in the present context.

Theorem A3-4 In the notation of Section A3.2.4, the state assignment to the behaviour of an automaton, $M \equiv (I,S,G,\sigma,\pi)$, given by -

$$u\gamma = \text{Range } (u\lambda) \qquad [A3.13]$$

- is a consistent state assignment.

Proof This is trivial, following from Equation [A3.11] where λ is defined as a sequence of mappings. Part (i) of the definition follows because there is no mapping corresponding to $u\lambda$ if u does not belong Σ. Part (ii) follows because the range of the product to two maps is included in the range of the last. Part (iii) follows because any common state in the range of two maps allows a map with this state in its domain to follow both.

If the state variables are not only to be consistent but also observable, then it must be possible to assign a single member of S to a behaviour by examination of sufficient behaviour preceeding it. By postulate (iii), part (1), of Section A3.3, the sub-sets belonging to $d\eta$, which each correspond to different sets of potential future behaviour, are each irreducible by further information about the past behaviour of the system. Hence it is reasonable to assign a single member of S to each distinct sub-set belonging to the union of all the $d\eta$. This observable state assignment is stated formally in the following definition.

A3.3.3 Definition of an Observable State Assignment

Let $\eta' \subset 2^{F_D}$ be a family of sub-sets of F_D which is the union of $d\eta$ for all d -

$$\eta' \equiv [X : \exists\, d,\ X \varepsilon\ d\eta] \qquad [A3.14]$$

- and consider the equivalence relation on η', that two sub-sets belonging to η' are equivalent if and only if they are constituted of identical elements. Let η^* be the quotient set of η' by this relation, i.e. the distinct sub-sets in the union. Let δ be a one-to-one mapping from η^* to an arbitrary, abstract set, S, so that for $X \varepsilon \eta$, $\exists\, s \varepsilon S$ -

$$s = X\delta \qquad [A3.15]$$

Then the mapping, $\gamma : F_D \longrightarrow 2^S$, defined by

$$u\gamma \equiv \left[(N_{wu})\delta \;:\; \forall\, t \in F_D \text{ such that } twu \in \Sigma\,,\; N_{twu} = N_{wu}\right] \qquad [A3.16]$$

- is defined to be an <u>observable state assignment</u>.

The defining points, $(N_{wu})\delta$, belong to S, even though the domain of δ is η^*, because the second part of the defining equation is equivalent to part (1) of postulate (iii) in Section [A3.3]. Hence, since any u may be expressed as - $u = xd$, where x is a sequence and d is a descriptor, we have that $N_{wu} \in \eta(d)$ for some d; and thus $N_{wu} \in \eta^*$.

This assignment is observable in that, for a behaviour, u, such that the potential future behaviour cannot be further reduced by knowledge of behaviour preceeding u, $N_u \in \eta^*$, and hence - $u\gamma = (N_u)\delta$, a single state belonging to S.

<u>Theorem [A3-5]</u> An observable state assignment is consistent.

<u>Proof</u> Using the notation of Definitions, A3.3.2 and A3.3.3, part (i) of the consistency requirements follows because there is no sub-set in the domain of δ which can follow impossible behaviour, and hence $u\gamma$ is empty if $u \in F_D - \Sigma$. Part (ii) follows because, if N_{wuy} is irreducible, it may be regarded as $N_{(wu)v}$ and hence it is also one of the irreducible sub-sets for $v\gamma$. Thus, $(N_{wuy})\delta = (N_{(wu)v})\delta$, is a member of both $u\gamma$ and $v\gamma$, and hence $(uv)\gamma \subset v\gamma$. Part (iii) follows from δ being a 1-1 correspondence, for if there is a common state to $u\gamma$ and $v\gamma$, $s = (N_{wu})\delta = (N_{w'v})\delta$, then $N_{wu} = N_{w'v}$, and hence there exists t belonging to both N_{wu} and $N_{w'v}$ such that ut and vt both belong to Σ.

It is interesting to note that an observable state assignment is not necessarily minimal, because the assignment does not take into account the inclusion relations between the sets in η^*. For example, given $a,b,c \in D$, it is possible that $a\eta \cap b\eta = \Phi$, $a\eta \cup b\eta = c\eta$, and hence it would be reasonable to assign s_1 to $a\eta$, s_2 to $b\eta$, and both s_1 and s_2 to $c\eta$. This is a consistent assignment which requires less states than the observable one, since the latter would require a further state, s_3, to be assigned to $c\eta$. However, with the minimal assignment, it would be impossible, given uc such that $N_{uc} \supset c\eta$, to determine whether the system is in state s_1, or s_2.

Having taken, η^*, effectively as the set of (observable) states, it remains to be shown that there is a one-to-one correspondence between partial transformations of these states and the set of descriptors, D, which extends to an isomorphism between the semigroup, F_D/θ , and the semigroup of partial transformations.

A3.3.4 Partial Transformations Corresponding to Descriptors

Consider a system defined extensively by its behaviour as a sub-set, Σ , of the free semigroup, F_D, generated by the set of descriptors, D. Let η^* be the family of distinct, irreducible sub-sets of future possible behaviour; let δ be a one-to-one mapping from η^* to an arbitrary, abstract set, S; and let the mapping, $\gamma : F_D \longrightarrow 2^S$, be an observable state-assignment - all as defined in Section A3.3.3.

Suppose that a mapping, M_d , from one sub-set of S to another such sub-set, is to be associated with each descriptor, $d \in D$. Take the range and domain of M_d to be -

$$\text{Range } (M_d) \equiv d\gamma \qquad [A3.17]$$
$$\text{Domain } (M_d) \equiv \{s : s=X\delta \ , \ X \in \eta^*, d \in X\} \qquad [A3.18]$$

- so that, the range of M_d is the set of states which may arise after behaviour terminating in d, and the domain of M_d is the set of states which may precede behaviour commencing with d.

The mapping M_d may now be defined. Suppose that the state, $s \in S$, belongs to the domain of M_d, then, from Equation [18] , there exists $X \in \eta^*$, such that $d \in X$ and $s=X\delta$. By the definition of η^*, it is a quotient set of η', and hence, from Equation [14] , $X \in a\eta$ for some $a \in D$. Hence, by the definition of η in postulate (iii) part (1) of Section A3.3, there exists $v \in F_D$ such that, for all $t \in F_D$ such that $tva \in \Sigma$, $N_{tva} = X$. Consider now the sequence of behaviour, vad. Since the possible future behaviour following va is irreducible, so is that following vad, and since $d \in X = N_{va}$, $vad \in \Sigma$. Hence, $N_{vad} \in \eta^*$, and thus $s' \equiv (N_{vad})\delta$ is defined, and $s' \in S$. The state, s', is to be taken as the image of s under the mapping, M_d, and hence it is necessary to prove that s' is unique. Suppose that v'a' is a second sequence such that, for all $t \in F_D$ such that $tv'a' \in \Sigma$, $N_{tv'a'}=N_{v'a'}=X$. As before, we have $N_{v'a'd} \in_N \eta^*$, and, for the sake of contradiction, will suppose that $N_{v'a'd} \neq N_{vad}$. Then there must exist a sequence, w, which belongs to one, but not to the other - suppose, without loss of generality, that it belongs to N_{vad}. Hence, we have - $vadw \neq 0$,

$v'a'dw = 0$, but the first equation implies that dw belongs to N_{va}. This, however, is identical to $N_{v'a'}$, and hence $dw \in N_{v'a'}$, which implies that $v'a'dw \neq 0$. Thus, by contradiction, we have $N_{v'a'd} = N_{vad}$, so that $(N_{v'a'd})\delta = (N_{vad})\delta = s'$ is unique.

Thus, the mapping, M_d, is defined such that -

$$s' = sM_d \qquad [A3.19]$$

- or, equivalently, such that -

$$(N_{vad})\delta = (N_{va})\delta \; M_d \qquad [A3.20]$$

- whenever the left hand side of the equation is defined.

The association of M_d with d may be written as a mapping, λ , such that -

$$M_d = d\lambda \qquad [A3.21]$$

- where the range of λ is $M \subset S^T$. This is identical to Equation [A3.10] of Section A3.2.4, and, as noted there, λ may be extended to be a unique homomorphism from F_D to F_M by setting -

$$(abc\ldots n)\lambda = M_a M_b M_c \ldots M_n = a\lambda . b\lambda . c\lambda \ldots n\lambda \qquad [A3.22]$$

- so that, for any $u \in F_D$, we may write -

$$M_u = u\lambda \qquad [A3.23]$$

The domain and range of M_u may be determined from the following considerations. Suppose $s \in$ Domain (M_u) and $s' \in$ Range (M_u), where $u = abc\ldots n$. Then $s \in$ Domain (M_a), so that, from Equation [20], there exists va (=w, say), such that $s = (N_w)\delta$. The image of s under M_a is $(N_{wa})\delta$, and the image of this under M_b is $(N_{wab})\delta$, and so on, so that, finally we have -

$$s' = (N_{wu})\delta \qquad [A3.24]$$

Hence, from Equation 16 -

$$\text{Range } (M_u) = u\gamma \qquad [A3.25]$$

and, since $wu \in \Sigma$ so that $u \in N_w \in \eta^*$, we have -

$$\text{Domain } (M_u) = [s : s = X\delta \,,\; X \in \eta^* \,,\; u \in X] \qquad [A3.26]$$

These two equations are extended forms of Equations [17] and [18] , and the first of them shows that, as required, the assignment of states to the sequence u by the range of M_u is an observable, and consistent, state assignment.

Theorem $\boxed{A3-6}$ The mapping $\gamma : F_d \longrightarrow S^T$, defined in Equation $\boxed{23}$, is an isomorphism.

Proof γ is an isomorphism if, and only if, for $u,v \in \Sigma$, $uv \in F_D - \Sigma$ $\Longleftrightarrow$ Range $(M_u) \cap$ Domain $(M_v) = \Phi$. Suppose $uv \in F_D - \Sigma$, but the range of M_u and the domain of M_v are not disjoint, then there exists s belonging to both. Since $s \in$ Range (M_u), we have $s=(N_{wu})\delta$, for some w. Since $s \in$ Domain (M_v), we have $s=(X)\delta$, where $v \in X$. Since δ is one-to-one, $X=N_{wu}$, and hence $v \in N_{wu}$, so that $wuv\ \Sigma$, and thus $uv \in \Sigma$. Hence, by contradiction, the result is proved from left to right. Conversely, suppose that the range and domain of the mappings are disjoint, but $uv \in \Sigma$. Then $v \in N_u$, and hence by postulate (iii), there exists w such that $v \in N_{wu} \in \eta^*$. Then $s=(N_{wu})\delta$ satisfies the conditons for belonging both to the range of M_u and to the domain of M_v. Hence, by contradiction, the result is proved from right to left.

A3.3.5 Transition and Output Equations of Equivalent Automaton

A semigroup of partial transformations having been constructed which is isomorphic to the semigroup of descriptors defining the system, transition and output equations may be established for a cybernetically equivalent automaton. Whilst the equations are similar in form to those of Equations [3] and [4], for a system defined by its structure, they are not idential because the equivalent automaton has a certain type of indeterminacy in its governing equations.

If the behaviour of the system consists of input/output pairs generated by some automaton, then we have the set of descriptors, D, which is a sub-set of the product set between the set of inputs, I, and the set of outputs, G - $D \subset I \times G$. Thus each $d \in D$ may be written, as in Equation [6], $d \equiv (i,g)$, where $i \in I$, and $g \in G$.

With each $d \in D$ there is also associated a mapping, M_d, according to Equation [21]. Let s belong to the domain of M_d, as defined in Equation [18], and let s' be the image of s under M_d, as defined in Equation [19]. s' may clearly be considered as the 'next state' of a machine, cybernetically equivalent to the system, Σ, whose current state is s and whose current input is i. However, s' is not necessarily a unique 'next state', because the pair (i,s) need not itself be unique to d and M_d. It is possible that there exists a second output, $g^* \neq g$, such that $(i,g^*) \equiv d^* \in D$, and $s \in$ Domain (M_{d^*}). Let $s^* = sM_{d^*}$, then, when the machine is in the state, s, and receives the input, i, its next state and output may either be s' and g, respectively, or s* and g*, respectively. Hence a transition equation, of the form of Equation [4] , is not necessarily determinate, and $\sigma(i,s)$ is a set of possible states rather than a single unique state; equally, the output, $\pi(i,s)$, also becomes a set of possible outputs.

The next state and output are not independently indeterminate, and there is a correspondence between the state and output sets, in that s' and g, or s* and g*, occur together. Hence, once the output is known, the 'next state' is well defined, but the converse does not hold, because it may happen that s' = s*, in which event the output alone is indeterminate. Because of this asymmetry, the indeterminacy is best assigned solely to the output by writing the transition equations:-

$$g \quad \varepsilon \quad \pi(i,s) \qquad [A3.27]$$

$$s' \quad = \quad \sigma(i,s,g) \qquad [A3.28]$$

The generation of an automaton structure which is indeterminate in that the present state and input do not determine the next state and output uniquely, is reasonable in the present frame of reference. Inputs and outputs have been inextricably mixed in descriptors, and there is no reason why the automaton's behaviour, in some circumstances, should not be characterized solely by its outputs. The automaton, even though its behaviour is indeterminate, remains observable, in that its output is sufficient to determine its state. It is, however, essentially uncontrollable, in that, even with unrestricted control over its inputs, it is not necessarily possible to force the automaton into a particular one of its potential future states. Accepting this indeterminacy, it is now possible to define formally an observable automaton structure which is cybernetically equivalent to a system defined by its behaviour, Σ .

In the notation of the preceeding sections, let $s \ \varepsilon\ S$, and $i \ \varepsilon\ I$, and define the set -

$$\Delta(i,s) \;=\; \{d : \exists\, g \;\varepsilon\; G,\; s \;\varepsilon\; \text{Domain}\,(M_d),\; d \equiv (i,g)\} \qquad [A3.29]$$

$\Delta(i,s)$ may be empty, in which event, the input i never occurs when the automaton is in state, s. Let

$$\pi(i,s) \;=\; \{g : d \;\varepsilon\; \Delta\,(i,s),\; d \equiv (i,g)\} \qquad [A3.30]$$

and $\sigma(i,s,g) = sM_d$, where $d \equiv (i.g)$ $\qquad$ [A3.31]

Then Equation $\boxed{30}$ and $\boxed{31}$ define the output and next state functions of Equations $\boxed{27}$ and $\boxed{28}$, for an indeterminate, but observable, automaton, cybernetically equivalent to the system defined by its behaviour as a sub-set, Σ, of the free semigroup, F_D, generated by the descriptor set, D.

A3.3.6 Problems in the Construction of the Automaton

The preceeding section gives the main results of this appendix - that from the system postulates (i), (ii) (Section A3.2.0) and (iii) (Section A3.30), it is possible to derive a structure for a system defined by its behaviour, which is minimum-state, observable, and cybernetically equivalent to the original system. In psychological terms, it is possible to eradicate all intervening variables which cannot be operationally defined and measured. Hence, it is possible, with full rigour, to re-state the behavioural definitions of modes of adaption, given in Chapter 2, in terms of the cybernetically equivalent structure. Thus is particularly important when it is desired to control adaption by varying the learning environment as discussed in Chapter 3. The system-theoretic results obtained in this chapter are also relevant, not only to the problems of the trainer dealing with an adaptive system, but also to the problems of the system itself in learning to cope with its environment, for example in the "dual-control" problem (Section 3.4.1).

However, there are difficulties in applying the results of this appendix to practical situations. At the most mundane level, the problem of manipulating the semigroup of sequences of descriptors to determine the mapping η (Section A3.3) of descriptors into irreducible sub-sets is itself computationally demanding, although it may be solved for simple cases with present computers. More fundamental is the problem of collecting the information as to which behaviours belong to Σ (A3.2(ii)). We have effectively solved a problem of complete induction in this appendix, and assumed that all possible behaviour is known. There are both practical and theoretical objections to basing a theory on this assumption. At a fundamental level there can be no operational procedure for collecting all possible behavioural sequences of the automaton, firstly if the automaton is irreversible so that it cannot be taken back to earlier states, and secondly if the maximum number of possible states is unknown so that it cannot be determined whether all possible behaviour has been elicited.

These difficulties are increased in practical situations where the time available for observation will be limited as are the possibilities for manipulating the behaviour. Both theoretical and practical obstacles generate the need for an incremental approach to determination of approximate structures.

In reality data about a system is gathered incrementally, item by item, and it is generally necessary to specify a structure for the system before its behaviour is completely known. If a minimum-state, observable automaton is constructed, which is cybernetically equivalent to some sub-set of a systems behaviour, then the question arises as to how this structure will change when new behaviour is observed. Alternatively, since new behaviour is expected to be observed, it is reasonable to add some examples of what may be observed to those which have already been observed, and base the automaton construction on this, particularly if this greatly simplifies the structure. In this event, it is necessary to know the effect on the structure of deciding that some behaviour is definitely now shown by the system.

The problems of incremental identification and approximation of an automaton structure may be stated formally -

given two systems defined as extensively as sub-sets, Σ, Σ', respectively, of the free semigroup, F_D, generated by the set of descriptors, D, and which satisfy system postulates, (i), (ii) and (iii), what is the relationship between the minimum state, observable automata, cybernetically equivalent to systems generating Σ and Σ', induced by the inclusion relation, $\Sigma \subset \Sigma'$.

If Σ and Σ' are identical apart from a few sequences then we would expect the structure determined for one to be a good approximation for the other. This may be formalized by considering any structure whose behaviour lies between Σ and Σ' as a tolerable approximation, and posing the question -

what is the minimal-state, observable automaton whose behaviour is a subset, $\Sigma_a \subset F_D$, such that -

$$\Sigma \subset \Sigma_a \subset \Sigma' \qquad \boxed{A3.32}$$

the problem of approximation is clearly closely related to that of incremental identification, and solutions to neither problem are available at present. The most relevant work is that on the simplication, and representation, of non-parametric data structures, such as those obtained in information retrieval systems, using graph-theoretical techniques (Salton and Sussenguth 1964, Meetham 1963, Vaswani 1965, Meetham 1966). The state transitions of an automaton, and the sequences of behavioural descriptors, form directed graphs (Ore 1962, Berge 1962, Flament 1963, Harary 1965), and techniques developed for the matching and simplification of graphs may be applied to the approximation of automata. Unfortunately, the theories of automata, semigroups, and graphs, are in comparable stages of early development, and no one is able to make a major contribution to another - it is in their synthesis that future advances in system theory may be expected.

3.4 Semigroups of States

In this appendix the input/output pairs of an automaton have been analysed as a semigroup. It is possible also to consider the sequences of states through which an automaton passes as a semigroup, and this enables a more powerful approach to be made to the problems of adaption and training than is possible with the purely set-theoretic definitions given in Sections 2.4.2 and 3.4.3.

Let $(I,S,G,\sigma\,\pi)$ be an automaton, defined as in Section A3.2.3, except that it may be indeterminate, so that σ and π are mappings into 2^S, rather than S itself, and the behaviour of the machine is determined by the equations:-

$$s' \in \sigma(i,s) \qquad [A3.33]$$

$$g' \in \pi(i,s) \qquad [A3.34]$$

Starting in a given state, for any particular input control policy - that is, for any procedure for selecting the inputs which may involve feedback and be dependent on past inputs, outputs and states, the automaton will pass through a sequence of states. This sequence will not necessarily be the same if the automaton is started in the same state again, and it is possible to conceive of an ensemble of identical automata being started in the same state to generate all possible state-sequences commencing with that state. If this is done for all possible states, then a set of state-sequences is obtained which, if the states are taken as descriptors, satisfies system postulates (i) and (ii) of Section A3.2.0. Hence, by Theorem A3-1, of all states sequences from the free semigroup of state sequences, F_S, which do not occur, are mapped into the zero element, 0, then the resultant structure is a semigroup.

3.4.1 Analysis of State-Semigroups in Terms of Ideals

The ideals of a semigroup (Appendix 2), sets which are invariant, or contract, under concatenation with any other element in the semigroup, are fundamental to the characterization of the semigroup structure. In the context of the state-semigroups of an automaton, the right ideals, in particular, characterize restrictions on the future behaviour of the automaton, and the stability, and some features of the controllability, of the automaton may most conveniently be expressed in terms of ideals.

Let Σ^* be a sub-semigroup of the free semigroup F_S^O generated by the set of states, S (Section 3.4). Let β be a transformation from the set of sub-sets of Σ^* into itself, such that:

$$\text{if } U \in 2^{\Sigma^*}, \quad U\beta \equiv [u : u=vw,\ v \in U,\ w \in \Sigma^*] \qquad [A3.35]$$

$U\beta$ is the set of sequences which commence with a sequence from U, and may be written as the concatenation of the sets U and Σ^* $U\beta \equiv (U)(\Sigma^*)$. It is clearly a right ideal of Σ^* because $(\Sigma^*)(\Sigma^*) \subset \Sigma^*$ so that $((U)(\Sigma^*))(\Sigma^*) \subset (U)(\Sigma^*)$, and may be called the right ideal generated by U; β may clearly be restricted as a mapping from either Σ^* or S into 2^{Σ^*}, and the same symbol will be used for the three mappings.

An ideal is called 0-minimal (Clifford and Preston 1961 p.66) if it contains elements other than zero, and the only ideal properly contained in it is the zero element. Hence, if W is a 0-minimal ideal, then either $M^2 = M$ or $M^2 = 0$ - in terms of state semigroups, the elements of M are either recurrent or terminal sequences. An ideal of Σ^* is characterized by 0-minimal ideals contained in it, and the union of the 0-minimal ideals contained in an ideal, M, will be denoted M . For consideration of stability, a state-semigroup may be taken to have no terminal states (those which cannot be followed by any other state, including themselves), and hence $(M\alpha)(M\alpha) = M\alpha$. Thus, the set of sequences present in the 0-minimal ideals is invariant, and any sequence of states contained in them may recur.

A3.4.2 Stability and Adaption

The concept of 'stability', in its conventional application to the behaviour of dynamical systems, involves an existent topology on the inputs and outputs of the system - a system is stable if a 'small' disturbance at the input produces a 'small' disturbance at the input; even in recent attempts to extend the notion of stability to more general systems, some form of inbuilt topology remains (Magiros 1966, Bushaw 1967). However, the inference from topology to stability, found in formal definitions, is the opposite to that actually used in discovering the properties of real systems and modelling their behaviour. If, under one set of conditions, the system behaviour has certain properties, and, after a transient change in these conditions, the property is retained then this change is a 'small' perturbation - that is, we discover the topology of the input by observing its effects on the output.

Hence, any topology upon the input of a system is useful only in so far as it reflects the effect of the input upon relevant properties of the system behaviour. For example, the Euclidean distance between two acoustic waveforms, regards as points in a space, is almost irrelevant to their properties as speech-points close together sound alike, but points sounding alike may lie far apart; the situation is worse for the sensation of pitch, since points close together may sound very different. Thus, the important problem in system stability is not the stability itself, and indeed every system is stable in some topology, but the relationship between extrinsic topologies applied to the system, and intrinsic ones derived from the system behaviour itself. In extending the notion of stability to general systems, it is this concept of intrinsic stability and the calculus associated with it that requires abstraction and extension, and the state-semigroups with their associated 0-minimal ideal structure provide the means to do this.

When considering the stability of linear systems, there is generally some preferred input, the zero input, which is normally present and about which perturbations occur. For the more general automata, as defined in Section A3.2.3, the 'steady-state', or preferred, input is arbitrary and may be taken to be any possible input, $i \in I$ say. Let the structural state-semigroup of the automaton, generated by the input sequences consisting of i repeated be Σ_i. For any initial state, s_o, the right ideal generated by s_o is defined as $s_o\beta$, and consists of the sequences of states starting with s_o which may occur when the input is i^n. The set of sequences $s_o\beta\alpha$ is also defined, and states occurring in these sequences form the 'confluence sets' used by Ashby (1960 Ch.14) in his discussion of habituation, a stability phenomenom. In a stable linear control system with no noise injection, $s_o\beta\alpha$ may be a sequence consisting of the 'zero' point in the phase space repeated indefinitely, and will be independent of s_o; whereas in nonlinear control of a linear system it will be the sequence of points on a limit cycle, and may be a function of s_o.

Suppose now that some transient disturbing sequence, u, is injected into the system in place of some segment of the input, i^m. The state after this input will belong to the set $\sigma(u,s_o\beta)$, where the next-state function has been extended to input sequences and state sets in a natural way. It is possible that this set does not coincide with $s_o\beta$, since the disturbance may give rise to passage through new transient states, and the question of interest is whether the sets, $s_o\beta\alpha$ and $(\sigma(u,s_o\beta))\beta\alpha$, coincide. If they do coincide, then the disturbance has had no effect on the ultimate state sequence and may be said to be 'small' - if the sets do not coincide, then their symmetric difference, $D_u = s_o\beta\alpha \cup (\sigma(u,s_o\beta))\beta\alpha - s_o\beta\alpha \cap (\sigma(u,s_o\beta))\beta\alpha$, is an indication of the 'size' of the disturbance, u. Certainly if $D_u \subset D_v$ then the effect of u is less than that of v - however, if no inclusion relationship holds, then a measure over the set of states enables a completely quantitative assignment of 'size' to a disturbance to be made.

This calculus of stability is completely intrinsic to the automaton and does not require any imposed topologies if the partial order over disturbances is accepted as an adequate description of their relative effects. One form of extrinsic criterion may be to require that the state of the automaton is ultimately within some set, $R \subset S$. This will be so if $s_o\beta\alpha \subset F_R$, and the disturbance, u, will be 'small' in its effect if $(\sigma(u,s_o\beta))\beta\alpha \subset F_R$. In the definitions of 'adaption sets' given in Section 2.4.2, if R is taken to be W(t), then A(t) is the set of states of the form, s_o, such that the system is extrinsically stable, for an input i=t, and no disturbances. C(T) is the set of states in which the system is intrinsically stable for each $t \in T$, with respect to a transient disturbance of the form $u \in F_T$.

Thus the calculus of adaption developed in Section 2.4 is no more than an application of a generalized theory of stability to the adaption automata derived from the learning behaviour. Unfortunately, nonlinear stability theory itself has not yet progressed to the point where it may serve as the foundation of learning systems theory. However, the results on training obtained in Chapter 3, particularly Section 3.3, demonstrate the value of a "stabilization" approach to training, and the potential for future extensions of nonlinear stability theory and its applications.

APPENDIX 4: THE HUMAN CONTROLLER

A4.1 Introduction

In this appendix studies of the human controller are reviewed which are relevant to the experiments on the learning of a tracking skill, described in Chapters 4 and 5; related material on human behaviour, studies of training, and particularly so-called 'adaptive' training is also incorporated in this appendix.

The long-lag type of tracking task used in the experiments of Chapters 4 and 5 is similar in its dynamics to the longitudinal motion of an aircraft, and linear models of the human control strategy with plant of these dynamics have been extensively investigated in the aircraft industry. Although the results do not give much insight into the actual human behaviour, they provide some useful approximations on which to base the designs of Chapters 4 and 5, and are critically reviewed in the following section.

A4.2 Foundations of Linear Modelling

In mathematical terms, a linear transformation is a mapping between vector spaces which obeys the superposition principle, in that the transform of the sum of two vectors is the sum of the transforms of each of the two individual vectors (Mirsky 1955). Functions of time over an interval form a vector space which is infinite-dimensional, and the operations of addition, scaling, integration, differentiation and time-delay may be shown to be linear operations (Riesz and Nagy 1955). A linear dynamical system is one whose action may be represented in terms of these operations alone (Birkoff 1927), and these systems have been extensively studied in linear systems theory. In particular, linear functionals from the space of linear functions to a complex algebraic variable have been developed, such as the Laplace transform, which enable linear operators on time-functions to be manipulated in an algebraic manner with full mathematical rigour.

Because linear system theory is so well-developed and contains such a powerful body of techniques for studying system behaviour, when a nonlinear system is to be analysed it is convenient to attempt to approximate its behaviour by that of some linear system. If the linear approximation is good, in some sense, then much of the behaviour of the nonlinear system may be predicted from a linear model whose behaviour is readily determined. In control engineering, techniques have been developed for the analysis of the stability of nonlinear systems using a linear approximation, or 'describing function' (Gibson 1963). The first techniques developed were based on an analysis of the behaviour of the nonlinear system when excited by simple harmonic waveforms at various frequencies. Booton (1953) extended these results to systems excited by noise-like signals, and it is his technique which has been used to derive linear approximations to human control policies.

A detailed mathematical analysis of the describing function technique is not relevant to the present studies, but certain assumptions made, and their applicability to the human controller, are important in evaluating the utility and implications of linear models of the human operator, and these assumptions will be outlined here. The configuration envisaged for linear modelling is shown in Figure A4-1: a nonlinear system, N, drives a linear system, G, the output of which, c(t), is subtracted from the input signal, r(t), and fed to the nonlinear system as an error signal, e(t); the output of the nonlinear element, m(t), is assumed to be made up of two components, one of which is correlated with the up of two components, one of which is correlated with the error, and the other of which, n(t), is independent of it.

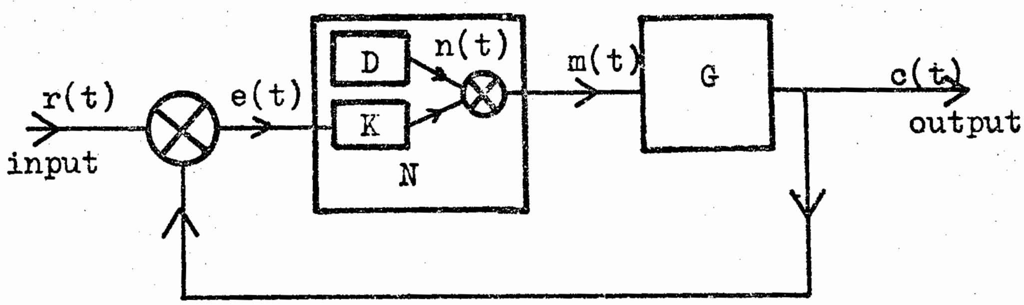


Figure A4-1 Linear Analysis of Nonlinear Feedback System

Booton's analysis depends on the assumption that the signals in the system, particularly e(t), are Gaussian processes. Even if e(t) is Gaussian, m(t), the output of the nonlinear element, will not be, and hence neither will c(t). However, if the linear element G is narrow band with respect to the input spectrum, by the central limit theorem its output will more closely approximate a Gaussian process (Gibson 1963 p.387), and hence, if r(t) is Gaussian, then so will be e(t). Thus, the use of describing function techniques to model a nonlinear system is dependent on the assumptions that the input is Gaussian, and that the nonlinear element is followed by a filter which is narrow-band with respect to the input spectrum.

There are alternative views of the describing function technique which throw some light on the meaning of the assumptions made. The overall linear model is an approximation to the transfer function between the input, r(t), and the output c(t). This will be good to the extent that G is a narrow-band filter which eliminates frequencies, n(t), which are not present in the input. Since G is known, K, the linear model of the nonlinear element is effectively available from the closed-loop response. However, the relationship between e(t) and m(t), predicted from a knowledge of K, will only account for that part of m(t) which is <u>not</u> filtered out by G. Hence, to the extent that G is narrow-band and enables the describing function technique to be used, it also restricts the model of N to account for only a small part of the behaviour of the nonlinear element.

A further effect on the type of linear model obtained for N is dependent on the amplitude of e(t) compared with that of r(t) (more strictly on the ratio of rms amplitudes). Since n(t) and r(t) are uncorrelated random processes, any part of n(t) which passes through G and is fed back to form e(t) increases the error, on average. Hence, for the controller to perform well and maintain a small error between overall input and output, it is necessary for the nonlinearly generated part of its output which passes through G to be small. In the context of the human controller, this implies that a good linear model may be obtained for the overall loop behaviour of an operator controlling a linear system; the model will not account for any components of the operator's output which have little effect on the system.

The linear approximation to a nonlinear system varies with system variables, such as the mean amplitude of the input - for example, a relay switching function whose output is the sign of its input has a constant rms output, and hence its 'equivalent gain' is inversely proportional to the rms input amplitude. Similar dependencies on e(t), and hence on both r(t) and G, occur for any form of nonlinear element, N, and the measured describing function will be found to be a function of the input and controlled element. Thus N will appear to be 'adaptive' to the input and controlled element, but this 'adaption' is completely open-loop and unrelated to any effort by N to improve its performance - the 'adaption' is an artifact resulting from the linear modelling, rather than adaptive behaviour on the part of N.

To conclude this critical examination of the describing function technique, it is worth quoting Gibson's remark (Gibson 1963 p.388) that, 'under a wide set of circumstances the use of the Gaussian describing function to compute closed-loop response is invalid'. Although the technique is based on a mathematical analysis which looks both impressive and plausible, in practice its derivation is based on highly restrictive assumptions, and, even when these apply, the meaning of the results obtained is not clear. Although this critique has been largely destructive, it is essential to consider linear modelling in some depth because it is the most obvious, and most readily applied, technique for analysing human control policies. Equally, any other approach to modelling must be able to withstand similar criticism, and the defects of linear modelling can be most readily overcome if they are throughly analysed.

A4.2.1 Results of Linear Model Studies of the Human Controller

The earliest study of the human operator as a linear servomechanism is that of Tustin (1947) who proposed that, despite amplitude nonlinearities, temporal discontinuities and haphazard fluctuations, there might be an 'approximate linear law' that would describe the main part of the operator's behaviour. Since that time, there have been many studies, including those of Russell (1951), Krendel (1951,1952), Elkind (1956), and McRuer and Krendel (1959). The early studies have been reviewed by Licklider (1960) and more recent reviews have been given by Summers and Ziedman (1964), Young and Stark (1965), and McRuer et al (1965). Hall (1963) has published a concise study covering the main aspects of linear models of the human operator in a flying situation, and this is summarized here for reference in Chapters 4 and 5.

All the studies of linear models referenced have taken control situations in which it is reasonable to expect the human operator to act linearly: the error, e(t), has been displayed on an analogue, positional display, such as an oscilloscope or pointer-meter; the operator's output m(t) has been applied to an analogue positional control, such as a joy-stick. In most cases only the error, e(t), has been displayed to the operator (compensatory tracking), but in a few studies, Elkind's in particular, r(t) and c(t) have been displayed on the same scale (pursuit tracking). Hall studied compensatory tracking with G, the controlled element, having a form corresponding to the short-period motion in the longitudinal dynamics of aircraft - its transfer-function was of the form:

$$G(s) = \frac{L(1 + 0.6s)}{s(1 + \frac{2K}{w_n} s + \frac{1}{w_n^2} s^2)} \qquad A4.1$$

where L is the gain of the controlled element, w_n is its undamped natural period in radians/second, and k is the damping ratio; in Hall's experiments, $0 \leq k \leq 1$ and $0 \leq w_n \leq 7$. The input, r(t), was a Gaussian random signal passed through a low-pass filter of the form $1/(1+s)^3$. Hall's operators were highly skilled pilots, used to controlling elements with dynamics of this form.

Hall's main results are illustrated in Figure A4-2 by plots of various variables in a plane with damping ratio, k, as abscissa, and undamped natural frequency, $F_n = w_n/2\pi$, as ordinate. He considered the operator to be acting linearly if the rms level of the 'remnant' terms, n(t), was less than five per cent of the rms operator output, m(t). From A4-2(i), it may be seen that the operator acted linearly, by this criterion, for the higher values of natural frequency and damping ratio. Figure A4-2(ii) shows contours delimiting regions of similar tracking performance in terms of mean error - this plot gives some indication of the variation in 'difficulty' of the tracking task with variation of F_n and k, and is important to the design of the adaptive trainer described in Chapter 4.

Hall found radical changes in the form of linear model associated with different controlled elements, and the regions associated with different models are delimited in Figure A4-2(iii); the forms of model are given in Table A4-1. The term $e^{-.2s}$, occurring in all four models, is a pure time delay of 200 msec similar in magnitude to a simple visual/motor reaction-time. The terms in s in the numerators of the transfer-functions

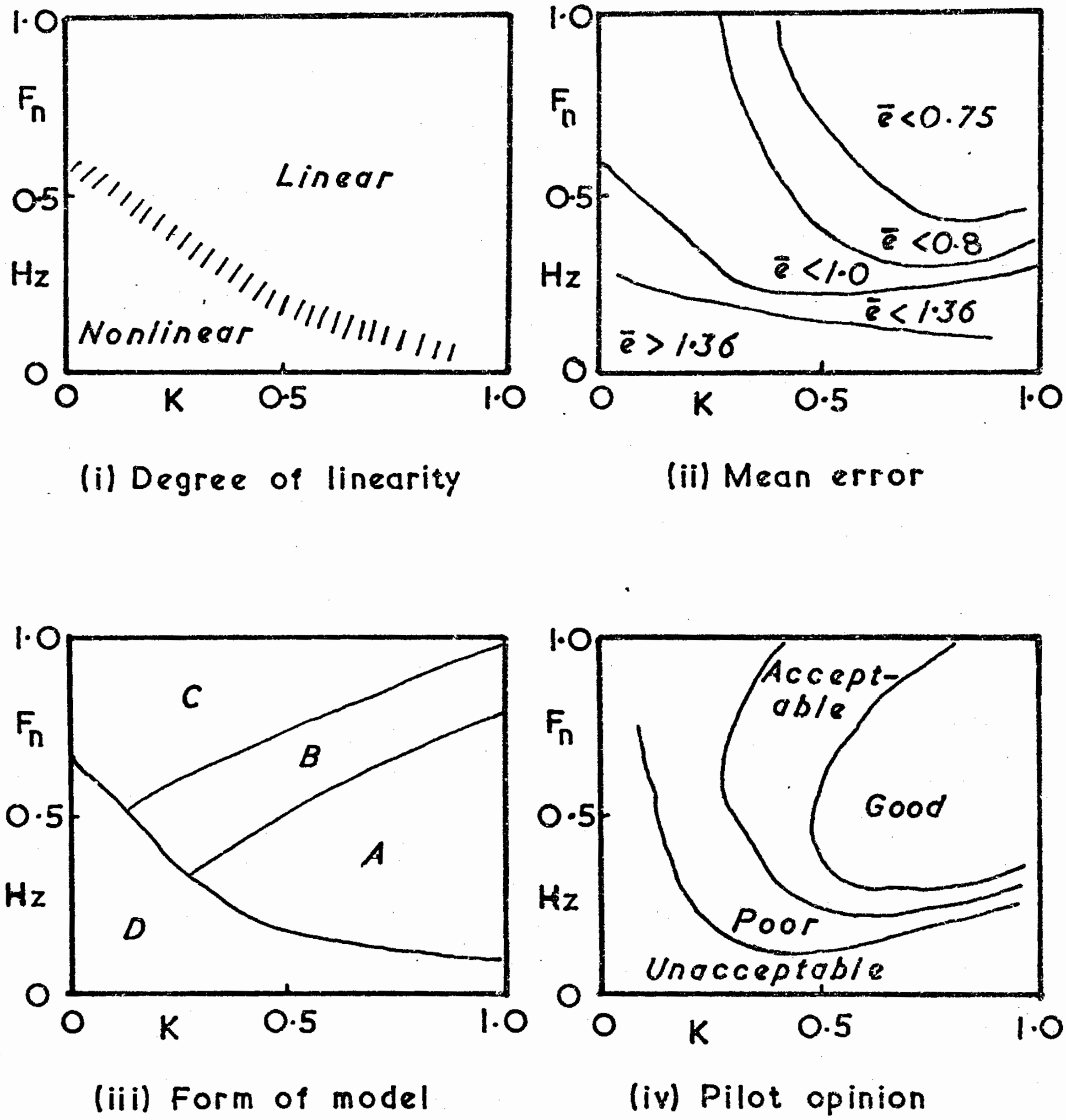


Figure A4-2 Results of Experiments on Linear Modelling

(Hall 1963)

Region -	Form of Model - $K(s)$	(Hall 1963)
A -	$0{\cdot}16(1+0{\cdot}67s)e^{-0\cdot2s}/(1+0{\cdot}2s)$	
B -	$0{\cdot}25\ e^{-0\cdot2s}$	
C -	$0{\cdot}5(1+0{\cdot}77s)e^{-0\cdot2s}/(1+2{\cdot}5s)$	
D -	$0{\cdot}0625(1+2s)^2e^{-0\cdot2s}/(1+0{\cdot}5s)^2$	

$$\text{where } K(s) \sim \frac{\text{stick output (inches)}}{\text{error signal (degrees)}}$$

Table A4-1 Various forms of Model for Human Controller

correspond to a phase-lead, or dependence upon the input velocity, and similar terms in the denominator correspond to a phase-lag, or smoothing of the error signal. Region A is one of high damping and medium-speed response, and the model has a predominant lead term showing that the operator is using the velocity of the error to predict ahead. Region C is one of low damping and fast response, where the higher frequency components of the input are very apparent, and the model has a dominant lag term showing that the operator is filtering out, or responding less, to these components. In Region B, between these two, the model suggests that the operator, apart from his reaction-tim delay, is acting as a pure gain element. In region D, where the system is show and underdamped, second-order terms in s appear in the numerator, showing that the operator is now making use of information about the acceleration of the error.

Hall asked the experienced pilots who acted as experimental subjects to rate the controlled element for its 'handling qualities' as if it were an aircraft, and the consensus of these ratings is shown in Figure A4-2(iv). It is from comparison of plots (iv) and (i) that the main justification for the utility of linear models of the human operator in the aircraft industry is derived - Hall states it thus

'If the pilot's opinion is 'good' the pilot is acting linearly. .... if a system is studied which has to be altered so that the pilot opinion will be high, the better the configuration becomes in terms of handling qualities the more accurate a linear analysis will be'.

This is a very fair, and a virtually complete, assessment of the utility of linear models of the human controller, emphasizing the restricted, but useful, conditions under which they apply.

One final question which Hall considered was to suppose that the operator was acting linearly and determine the 'source' of the remnant term, n(t). Records of the pilot response, m(t), for configurations in region D of Figure A4-2(iii) showed that there was superimposed on the output predicted by the linear model a 'rather frantic switching' mode, with the pilot alternating 'somewhat arbitrarily between the two hard over stock' positions. Hall reports that the switching did not occur regularly and was not correlated with error zero crossings - he matches the remnant component in spectral density with a random telegraph waveform having a mean time between switching of 3.5 seconds. Diamantides (1958) has reported a similar effect under the same conditions, and ascribes it to pilots attempting to obtain 'informative feedback' about the controlled element dynamics by injecting a signal into the loop. He also reports that the injected signal is more apparent with less skilled pilots, and, in one operator at least, the signal disappeared with learning as an 'exponential function of time'.

A4.2.2 Utility of Linear Models in Human Operator Studies

Because of the strong theoretical constraints upon the circumstances under which the describing function is meaningful, several workers have studied the validity of the necessary assumptions in experiments with the human operator, and the extent to which overall behaviour, such as stability, may be predicted from the measured linear models. Elkind & Darley (1963) measured the deviations from a Gaussian distribution of the operator's output, m(t), the remnant, n(t), and the error signal, e(t), for a controlled element, G, which was a pure gain, with an input r(t) which was band-limited Gaussian noise. He reports that the output 'obtained with all inputs and the error and remnant signals obtained with medium bandwidth inputs appear to be approximately normally distributed'. Hall (1957), in a similar experiment but with a controlled element of the form given in Equation A4.1, found that the amplitude distribution of the error signal was approximately normal, but the distribution of the operator's output appeared rectangular and even bimodal when the bandwidth and damping of the controlled element were low.

Jex, Cromwell and Siskind (1960) and Smith (1963) have compared the stability boundary for the human operator, computed from describing function measurements, with the actual boundary found by experiment.

They used the results obtained by Krendel and McRuer (1960) to predict the stability of the controlled system for second-order unstable dynamics (negative damping ratio), and find it necessary to introduce an input-predictive mode of operation to account for the experimental results. Skolnick (1966) has used measured data on the human operator describing function to determine 'capability bounds' on the human controller, and has proposed techniques for optimizing the performance of a control system containing a human operator using these bounds.

Various workers have studied the effect of verbal instructions to an operator on the parameters of a corresponding linear model, and these effects are summarized by McRuer and Krendel (1957) in what is still the most comprehensive and detailed discussion of linear modelling and its relevance to human operator studies. McRuer and Krendel used two sets of instructions, one of which emphasized 'speed' in reducing the effect of disturbance, and the other of which emphasized 'accuracy' in nulling the disturbance. For one operator they found no change in the measured describing function under the two conditions, whereas for another they found a distinct change corresponding to: higher dc gain and shorter smoothing time-constant when the emphasis was on 'speed'; lower dc gain and smoothing time-constant triple what it was for 'speed' when emphasis was on accuracy. Russell (1951) measured the change in linear models parameters after the operator had drunk a substantial quantity of alcohol, and reported lower dc gain and greatly reduced capability to introduce lead, that is, to estimate error-velocity.

Sheridan (1960) has used a technique for measuring the describing-function on-line to follow changes in the model parameters when those of the controlled element undergo a step variation. He reports that, 'the experienced operator adapts almost instantaneously if the parameters of controlled process or the type of display suddenly change'. This is in accord with the more recent studies of Young et al (1964) who investigated the time taken for the operator to adapt to changes in the gain of the controlled element, and sense of the error, in a simple compensatory tracking task. They report that, 'adaption generally occurs in 0·4 - 0·8 sec following a controlled element change, and the resulting error is usually reduced to its asymptotic level in 1-3 sec following transition. Krendel and McRuer (1960) have outlined a developmental approach to the learning of a tracking skill in terms of the parameters of the describing function at various stages. Fuchs

(1962) has put forward a 'progression-regression' hypothesis suggesting that in learning the parameters of the higher time derivatives of the error will be gradually given more weight, whilst under stress their relative weights will be reduced. DeLessio and Palin (1961) put forward a program to identify the time-variation of the parameters of an operator's describing function, and hence form 'an adaptive model for the human operator', but this program has not been carried out.

Briggs (1964) has defended the linear describing function as a powerful methodology for human operator studies, but claims that there has been too much effort expended on the development of the technique, and too little on 'the descriptive quantification of behaviour and analytic tests of hypotheses about behaviour'; he fears that the same mistake will be made in future developments of nonlinear models. This is an important criticism which bears equally on the development of models of human adaptive behaviour, and in the following section some details of the relationship between actual behaviour and its equivalent linear model are examined.

A4.2.3 Nature of the Linear Approximation and Constraints Upon It

In the studies of the feedback trainer, described in Chapters 4 and 5, it has been natural to use controlled elements whose parameters vary progressively, but rapidly, over a range of values, and hence to obtain records of the behaviour of the same operator under different conditions within a short span of time. These records serve to illustrate some of the characteristics of the describing function discussed in Section A4.2. Figure A4-3 shows the input, $r(t) = \sin(\pi t/5)$, a sine-wave of 10 seconds period, and the operator's output, for a tracking task with continuous manual input and continuous visual display of error, and controlled element dynamics of the form:

$$G(s) = L/s(s+1/T)^2 \qquad [A4.2]$$

that is a second-order lag of time-constant, T, followed by a pure integration; this is similar to Hall's dynamics, Equation A4.1, with $k=1$ and $w_n=1/T$ ($F_n=1/2\pi T$).

In the upper part of the figure, which shows the response for a short time lag ($T=0{\cdot}25$ sec), the operator's response has an overall shape which is similar to that of the sine-wave input but lagging it in phase. This response is clearly made up of a number of discrete

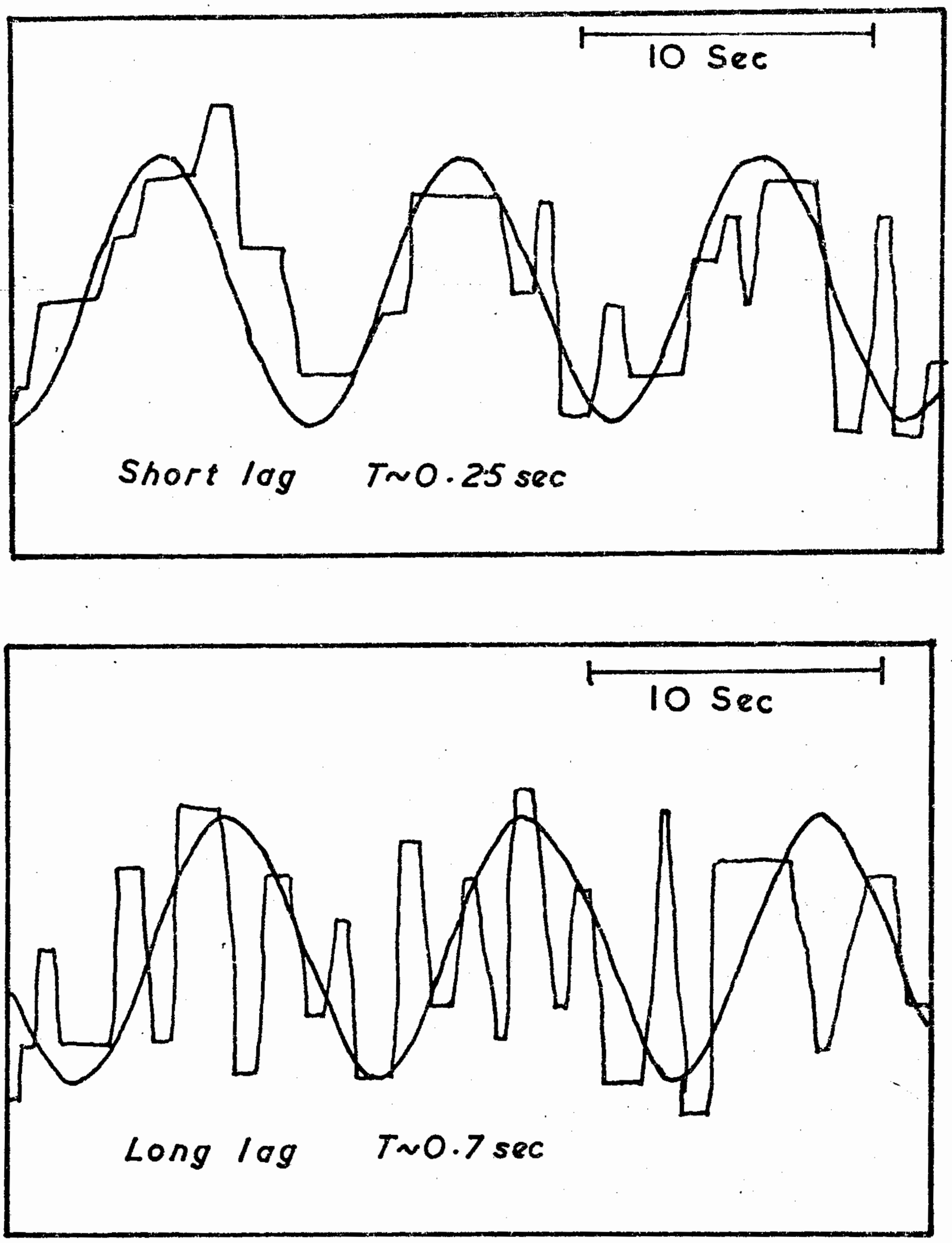


Figure A4-3 Sine-Wave Tracking Through Cascaded Lags movements, however, and is not the continuous sinusoid response which would be obtained from a linear servo in the same situation. If the Fourier transform of this response, however, it will clearly be found to have a very high percentage of its energy in the expected sinusoid, and very little in the 'perturbation' due to the discrete movements. The converse is true of the response at high lags (T=0·7 sec), shown in the lower part of Figure A4-3. The operator is now responding so rapidly with such large amplitude movements that his output appears closer to a pulse-width modulated signal than a sinusoid. A Fourier transform would still show a phase-lagging signal at the input frequency, but this is now lower in amplitude and accounts for a minority of the energy in the response.

A correlational analysis of the control behaviour partially shown in Figure A4-3 would result in a good linear fit to the controller at short lags and a bad fit (high remnant) at long lags. More importantly, the model would differ greatly for the two situations, and yet it is plausible, both from an examination of the records and from the fact that they were taken within a few seconds of one another from the same operator, that the operator has not changed his control strategy in the least. This is the gravest defect of the describing function - that the linear model of an operator may vary widely as function of his environment without hsi control strategy changing at all.

The intermittency and discreteness of the human operator's response is not a newly discovered phenomenom - Telford (1931) reported a 'refractory phase' in the motor responses to two stimuli presented within an interval of about 0·5 seconds of one another, and Craik(1947) described the type of response in a tracking task, shown in Figure A4-3, as 'intermittent corrections' consisting of 'ballistic movements'. However, whilst many workers have followed up Telford's discoveries of a central refractory period in simple discrete stimulus/response situations, lack of development of both the theoretical and technological tools has made it impossible to go further with Craik's analysis until recently. Even now only a few steps forward have been taken, and no comprehensive and complete structure, equivalent to the describing function, is available for nonlinear studies of the human operator. In the following section, work on nonlinear models of the human operator is reviewed for its relevance to improved models of human adaptive behaviour.

A4.3 Nonlinear Models of the Human Controller

The evidence for a fundamental discrete-time, discrete-action basis for human perceptual-motor skilled behaviour has been presented in general reviews by Summers and Ziedman (1964), Young and Stark (1965), and Poulton (1966), and also in theses proposing sampled-data models of the human controller (Bekey 1962, Lange 1965). Definitive evidence has been gathered of discrete-action and discrete-time phenomena in peripheral behaviour such as hand and eye movements, and theoretical models of these have been explored in depth. Less firm evidence has been adduced for discrete-time phenomena in perception and decision-making, but no models have yet been proposed which can account for all the experimental data. In the following section work on discrete behaviour in sub-structures of the human controller is reviewed, whilst further sections outline sampled-data and 'bang-bang' models of overall tracking behaviour.

A4.3.1 Discrete Phenomena in Human Peripheral Dynamics

In moving his hand from one position to another, or in rotating his eye from one fixation to another, the human operator has to vary the location of a mass using the force exerted by his muscles which is limited in its maximum value. Dynamically, the hand or eye is virtually a pure mass, with low dissipation of energy through friction, and low storage of potential energy through spring-like behaviour. A simple servomechanism, in controlling the location of an object, applied a force to it proportional to the deviation of the location from the desired one, in such a direction as to reduce the deviation. Bushaw (1953) showed that the control policy of the linear servomechanism was not time-optimal, in that it did not reduce the error in location to zero as rapidly as possible, and he showed that a 'bang-bang' controller, applying maximum available force in one direction for half the time and then applying it in the other, gave improved performance. From 1953 onwards, a number of workers proved, with increasing generality, that the minimum-time control of a linear system was achieved by a controller which applied either maximum or zero force (Fuller 1960). This result has been extended to the optimization of performance criteria other than settling time, such as error-functionals (Fuller 1960*).

In 1962, Smith (1962) and Wilde and Wescott (1962) published papers giving experimental evidence that the human operator used bang-bang control in moving his hand and arm, and at the same time van der Gon, Thuring and Strackee (1962) described a 'handwriting simulator' which accurately reproduced the movements of the hand in writing using bang-bang controllers in two dimensions. In a later paper, van der Gon and Thuring (1965) reported that the controllers worked at a fixed force within a movement pattern, rather than at constant maximum force. They state that, 'to write the same word involves the use of the same timing and that the instruction of change of size is interpreted as change of force'; since the size of the writing varies as the square of the force, small changes in force are adequate to produce large changes in size. Equally, the human operator does have an ultimate limit in the force applicable, and loading the hand or arm with more inertia reduces the speed of movement (Smith 1962). The minimum time of application of the force for an unloaded limb was found by all workers to be about 90msec, which tallies with the response-time of the muscle servo (Hammond, Merton and Sutton 1956) and the rate at which nervious pulses are sent to the muscle (Lippold, Redfearn and Vuco 1957). A similar discrete-action servomechanism has been discovered in the control of eye-movements (Stark, Vossius and Young 1962, Young and Stark 1963), with independent control of positional saccades and velocity pursuit motion, again with forces applied for about 90 msec in turn.

Apart from the clearly defined discrete phenomena in human limb and eye movements, there is considerable circumstantial evidence for discrete phenomena, 'data-sampling' or a 'psychological movement' in perception itself. Experiments on the 'psychological refractory period' (Welford 1952), on choice reaction times (Hick 1952), on temporal numerosity (White 1963), on periodicities in simple reaction times (Stroud 1954, Augenstine 1954, Venables 1960), on backward masking of one stimulus by a succeeding one (Kolers 1962), and on the reaction time to the cessation of a repetetive stimulus (Callaway and Alexander 1962), all suggest that visual perception is not a continuous process. Various authors have suggested, on the basis of such data, that the brain works in terms of a moment of time, in duration about 90 msec, within which events are confused in their temporal relationships. As Kolers (1968) and Allport (1968) have pointed out, however, no simple model of such discreteness in time can account for more than a minority of the known phenomena, although it is clear that some form of discontinuity is present. Some workers (Wiener 1948, Lindsley 1952, Surwillo 1963) have attempted to link the hypothesized periodicity in perception of 90 msec with the similar periodicity in the alpha rhythm of the brain, and indeed Surwillo has described definite experimental evidence of a strong correlation between alpha period and simple reaction time over a population. However, no incontrovertible evidence of such a link has been obtained.

A4.3.2 Sampled-Data Models for Human Tracking Behaviour

The evidence of temporal discontinuities in human perception and movement, together with the observed nonlinearities in human tracking behaviour (Craik 1947, Hick 1948, Poulton 1962) when perception and movement are coupled closely together, has lead to a number of proposals for data-sampling models of the human controller in which the display is sensed intermittently and a motor-pattern released according to what is observed. In control engineering (Kalman and Bertram 1959, Jury 1958) such sampled-data control systems became of practical importance with the use of digital computers in control loops, and for the case where the sampling frequency is constant a theory of linear sampled data systems has been developed based on the z-transform, which is similar in pwoer to the theory of continuous linear systems based on the Laplace transform. Because such a theory exists, it has been customary to base recent human operator models on sampled-data systems with constant sampling-frequency, although this does not fit the experimental data (Lange 1965), and attempts have been made to develop techniques to deal with more complex sampling criteria (Bekey 1962).

The earliest sampled-data model was that of North (1952) who took Tustin's model of the human operator and replaced the differential equations by difference equations. North matched the behaviour of his model against that of the human operator in terms of power spectra only, and the first study in which the behaviour was matched in the time domain was that of Ward (1958). More recently Bekey (1962), Lange (1965) and Kreifeldt(1965) have proposed sampled-data models for human tracking behaviour, and Bekey(1965) has reviewed some of this work. Lange's work was a continuation of the work of Wilde and Lemay (Wilde and Wescott 1963; Lemay and Wescott 1962), and has the most detailed experimental backing; the main points of his model and results are outlined in the following paragraph.

Lange considered compensatory tracking through a simple gain, with continuous manual control and visual display, of zero-mean Gaussian noise with a cut-off frequency of 3·8 radians/seconds, and used highly trained operators as subjects. In his model, the operator samples both position and velocity of error instantaneously, regularly at about 150 msec intervals, and attempts to reduce both to zero by a 'bang-bang' output to actuate his hand. The qualitative nature of the output of the model is a far better match to the operator's output than that of a linear model. Quantitatively, correlations of between 0·8 and 0·9 were obtained between model errors and operator errors, corresponding to cross correlations between their outputs of between 0·98 and 0·99. The match in the time-domain could have been improved by taking a varying sampling interval, and Lange suggests an extended model with random variation of the sampling frequency.

A4.3.3 Bang-Bang Models of Human Controller for High-Order Systems

Data-sampling models ofthe human control provide a good representation when the controlled element is a pure gain and the operator is effectively required to match a difficult waveform. In this situation, the movement of the hand to match the waveform, and the movement of the eye to track it, are clearly the main variables, and the tracking models are closely related to those of the hand and eye alone. The situation is also a very natural one, to which hand and eye co-ordination should have become well-suited during the course of evolution, and it is not surprising that the movement time of the eye, the reaction time delay between visual stimulus and motor response, and the movement time of the hand, are all similar in magnitude at about 180 msec - it would be no advantage to the system to have one very much less than the others. Hence, a 'sampling interval' of the same order is a reasonable approximation in these simple situations. When the lags in the controlled element become very much greater than those in the operator, however, the eye and hand in themselves become of less importance; and the problem-solving capability of the brain in between them comes to dominate the behaviour.

It was noted in Section A4.3.1 that bang-bang, or maximal force, control is the optimum strategy for the control of a pure second-order system which approximates to the dynamics of an eye or limb. This result has been extended to the time-optimal control of any linear system, and the maximum-effort controller is becoming as ubiquitous in the literature as the linear controller (Fuller 1960, Fuller 1962). Pew (1963) and Young and Meiry (1965) presented experimental evidence that in the control of both stable and unstable second-order systems the human operator adopts a bang-bang control strategy, and have shown that tracking improves if this strategy is forced upon the operator by giving him a two-position only control.

Because the two-level output of a bang-bang controller is far simpler to monitor than the continuous output of a linear controller, and the control strategy can be represented by those points in the state-space of the controlled element at which the controller changes from one output value to the other (the 'switching-line' in the position/velocity 'phase-plane' for a second-order system), it is comparatively simple to measure the control policy of the human operator working in a bang-bang mode. In particular, the adaption of the control policy during learning is readily followed, and since a plot of individual decision-points is obtained as a function of time it is possible to clarify the effects of indeterminacy in the policy (in the 'search' phase), indeterminacy in the measurement of the policy (since only a limited number of data points are available), and time-variation of the policy with learning. In linear modelling by correlational techniques, the smoothing of data over time causes these factors to be inextricably mixed.

Li, Young and Meiry (1965) have described qualitatively the variation of the human operator's switching line in learning to control an unstable second-order system. Weir and Phatak(1967) have measured the time-variation of the switching line in response to step changes in the controlled element dynamics. However, as yet, there does not appear to have been published any detailed study of the learning of a high-order control skill, where a bang-bang control policy is either forced by the nature of the controls, or expected to appear.

A4.3.4 Adaptive Nonlinear Models of the Human Controller

Angel and Bekey (1968) have described a simple finite-state machine for the control of a pure second-order system, based on experimental studies of discrete actuation in human limb movements (Section A4.3.1), which provides a qualitative match to many of the characteristics of human hand motion, and has self-adjusting properties giving it an adaptive capability; so far, they have not presented studies of the goodness of fit of the model to human tracking behaviour and its adaption. Preyss and Meiry (1968) have described a 'stochastic model' of human learning behaviour in controlling a pure second-order system, in which the output is bang-bang and its polarity is switched on the basis of probabilistic estimates of the efficacy of so doing. These estimates are themselves built up from prior experience using Bayes rule (Minsky and Selfridge 1961) to weight the evidence obtained from sensors giving quantized position and belocity information from the controlled element. This model learns to control the second-order system, and its behaviour, both in tracking and in learning is qualitatively similar to that of the human operator -again, no detailed analysis of goodness of git is presented.

Gaines (1967) and Gaines and Quarmby (1968) have presented comparative studies of human and machine learning behaviour, in which the learning model was an adaptive-threshold logic pattern-classifying adaptive controller; details of these studies are reported in Chapter 6 for comparison with the human operator experiments of Chapter 5 - only goodness of fit to the learning behaviour is considered. Studies of learning system models of the human operator are currently limited only by the availability of suitable learning systems in a utilizable form. As more learning machines become generally available, preferably as computer programs for small, on-line process control machines, it will be possible to evaluate their utility as human operator models.

A4.3.5 Tracking with Nonlinear Controls

Since the human operator of high-order systems adopts a strongly nonlinear control policy, it is of interest to consider whether his performance is enhanced through the use of a control which naturally induces this type of policy; for example, a two or three position joystick rather than a continuously variable control. Young and Meiry (1965) have noted that, in high-order systems, the error is dependent on the integral of the control movement, and the operator must keep track of this quantity. With a continuous control, this involves integration of a continuous function of time; with a two or three position controller, it involves summation of the time intervals when the output is positive, and subtraction of those for which it is negative; for a pulsing controller, which gives out fixed-duration, fixed height pulses, either positive or negative in sign, it involves only counting the excess of one type of pulse over the other. Thus these three types of control should be successively easier to use, provided the integrations in the system are adequate to filter out the quantization noise of the nonlinear controls.

Pew (1963) found in his studies that the performance of the human operator in controlling a pure second-order system was similar with a continuous joystick and a two-position switch. Kilpatrick (1964) found that when the controlled element dynamics were of the form, $1/s^2$ or $1/s^2(s+3)$, there was no significant difference between the two types of controller, whereas with a very difficult controlled element, requiring more lead, such as $1/s^2(s+1)$, the rms error for the continuous control was fifty per cent higher than for the bang-bang control. Young and Stark (1965) note that, in Kilpatrick's studies, that 'even though the operator uses the continuous controller in a more or less bang-bang fashion, he is able to use the bang-bang controller in a <u>pulse</u> control fashion'.

Gaines (1966,1967) has reported that the use of pulsing controls not only improves the performance of the human operator in high-order systems, but is also less fatiguing; some experimental results are described in Chapter 5. He was interested in obtaining a control for use in studies of training, which was itself difficult to use and involved interactions between the learning of the tracking task and learning to use the control. Building memory into the pulsing control system, such that the sign of a pulse obtained from one of two push-buttons depended upon that last pressed, gave a control with the required characteristics. The control consisted of a pair of push buttons, one held in each hand, such that pressing one push-button would give out a positive impulse. The polarity of the two push-buttons was not constant, however, and changed each time either was pressed. Hence, to obtain a stream of pulses of constant polarity, it is necessary to alternate between the two push-buttons.

Initially this control feels most awkward and unnatural to use, but eventually, after ten to thirty minutes of use under reasonable conditions, it becomes as simple and natural to use as the non-reversing push-buttons. The problems in using this control may be appreciated by considering the situation in which the operator has pushed a button and the error has increased - his natural tendency is to push the other button, but the correct response is to push the <u>same</u> button again. Gaines (1966) describes the various stages of learning to use the push-button controls, from an almost entirely verbal strategy, through the build-up of response structures, to a highly-skilled, non-verbalizable control strategy. This type of control was used in the experiments of feedback training, described in Chapter 5, and these provide further information about the problems of learning associated with it.

A4.3.6 Control Strategies in Multi-Variable Situations

Whilst much research effort has been devoted to the study of the dynamics of the peripheral mechanisms of the human controller, and to the linking of these by control strategies for compensatory tracking, and much progress has been made in the understanding and modelling of the human operator in simple situations, there is no comparable understanding of the control strategies adopted in more realistic situations, where the operator has multiple, diverse and interacting tasks to be performed either simultaneously, or sequentially. However, as Pask (1960,1965) has noted, these are the situations for which training is required, where a number of interacting sub-skills have to be learnt, and the training problems for the simple, laboratory tracking skills used in modelling the human controller may be very different from those encountered in more realistic situations.

At the level of overall performance, there have been a number of studies of two-dimensional tracking tasks, with the error on one axis presented as the horizontal axis on an oscilloscope, and the error on the other presented as the vertical axis - the control being a two-dimensional, continuous-output joystick (Chernikoff, Duey and Taylor 1959, Duey and Chernikoff 1959, Chernikoff and Lemay 1963). The main result of these was that tracking in both axes deteriorated as the task dynamics in the two became more different, and that a two-dimensional task with the same dynamics in both axes was similar in difficulty to the equivalent one dimensional task. At a similar conceptual level, Dander (1963) has investigated the possibility of predicting pilot ratings of multi-axis control tasks from single-axis data. The interference between widely differing tasks, which do not in themselves interact, has been extensively investigated in studies to improve the sensitivity of performance measures through the use of secondary tasks (Knowles 1963). The factors which make 'secondary loading' techniques useful operate to to make the modelling and prediction of interference difficult, since there is generally a level below which a secondary task shows no detectable effect - it is using up the operator's 'spare capacity'.

At a detailed level,the main problem in modelling human control strategies in multi-variable situations is that of measuring and simulating 'attention switching' as the operator multiplexes his control capabilities to various parts of the total system. One of the more accessible and important indications of attention is the instrument at which the operator is looking, and much effort has been devoted to measuring and modelling the human controller's visual behaviour in a many-instrument, multi-dimensional tracking task (Senders 1964, Carbonell 1966, Senders, Ward and Carbonell 1967). Senders originally proposed, and tested experimentally, a model in which the frequency of sampling an instrument was proportional to the potential information flow-rate through that instrument regarded as a communication channel. The later studies extend this to models which take account of the 'queueing' of instruments for attention, and the risk taken in not reading a particular instrument.

Even pursuit tracking, where the operator is shown not only the error but also the input, or disturbing, signal, is itself a multi-variable tracking situation, and poses far greater difficulties in the analysis of the operator's behaviour than does compensatory tracking. Poulton (1952, 1952*, 1957, 1957*) has studied the difference between behaviour in compensatory and pursuit tracking in great detail, and suggests that the advantage of the pursuit situation stem from its enabling a complete separation to be made between the demanded input to the system, the 'track', and the operator's own input through the system. This separation aids both the prediction of future system behaviour, and the modelling by the operator, as part of his learning process, of the demand signal and system dynamics.

The pursuit tracking situation becomes even more complex when the operator can see not only the immediate value of the demand signal, but also some segment of its future values, for example, in a car-driving situation. Classical control theory gives no indication of how advantage may be taken of such a pre-view, and hence it has been impossible, until recently, to approach the modelling of human control behaviour in the vehicle-steering situation from a control-theoretical point of view. Sheridan and Roland (1966) have now used the modern control technique of 'Dynamic Programming' (Bellman and Dreyfus 1962) to obtain a normative model of the optimum control strategy in this situation for comparison with that of the human operator. This interesting and powerful approach has also been expounded in some detail by Thomas (1962), and offers the opportunity for a substantial advance in the understanding of human control behaviour. Dynamic programming is essentially a computational technique which enables very general control problems to be solved, given a criterion of optimality, by numerical algorithms, its main disadvantage is the amount of computation and data-storage required, but this is far less than that for a complete search of all possible control policies.

A4.4 Overall Performance of the Human Controller

Work on the detailed modelling of human behaviour, whether linear or nonlinear, has been the exception rather than the rule, and psychologists have tended to concentrate on the effects of variables such as training conditions, stress, and auxiliary tasks on performance, rather than on the actual control behaviour. Adams (1964) has noted that American research has dealt with tracking performance, and given less emphasis to the underlying mechanisms of the skilled activity, whilst British research has taken a more molecular view.

In this section, some of the results of studies on the measurement of performance, and the effect of variables neglected in the control-theoretical models of tracking behaviour, such as instructions and stress, are outlined.

A4.4.1 The Measurement of Performance

The taxonomy of adaptive behaviour introduced in Chapter 2 involves the definition both of a task, and the satisfactory performance of a task. The measurement of satisfactory performance on a particular task is not itself difficult, because an operational definition of 'satisfactoriness' is built into that of a task. However, in practice the question of interest is more likely to be to determine the range of tasks, out of a set of possible tasks, to which the operator is adapted; in a real training situation, there is rarely a single well-defined task, and in more basic studies it is necessary to make maximum use of the experimental data.

When the range of tasks is such that there is a single parameter of difficulty, and, for a given operator, increasing the difficulty decreases the performance, then the problem of determining the range of adaption reduces to that of determining the task of greatest difficulty which the operator can perform at a specified level. In practice, it has been easier to measure the operator's performance at a number of levels of difficulty, rather than the difficulty for constant performance, and, as Poulton (1965) has noted, this may lead to lack of differentiation between operators with different capabilities if a sufficient span of difficulty is not included in the tests. When the task of interest is not readily varied in difficulty, or no natural continuum of difficulty exists, it may be possible to create an equivalent effect by giving the operator a second task, assumed not to interact physically with the first, which can itself be varied over a continuum of difficulty. The performance on the secondary task may then differentiate between operators, even if that on the primary task does not (Knowles 1963); the situation is also more realistic than that of giving the operator a task of high difficulty, since practical problems, such as flying, generally involve multiple, rather than individually difficult, tasks.

One assumption about the human operator which may be made to give some theoretical foundation to the use of secondary tasks to increase the sensitivity of performance measures, is that the operator is a single-channel system (Broadbent 1958) whose 'capacity' is constant, so that the secondary task measures the amount of channel capacity surplus to the requirements of the primary task. It is reasonable to extend this model and hypothesize that the stress on the operator increases as high channel capacity is taken up by a task, leaving less capacity for emergencies, and hence, even without a secondary task, physiological indicators of stress may be used to establish the degree of effort, or 'channel capacity' required by the main task. Thus, where Brown and Poulton (1961) used a secondary task to measure the demands of different road situations on car drivers, Taylor (1964) used the operator's galvanic skin response to the same end. Similarly, Benson, Huddleston and Rolfe (1965) have used a variety of physiological measures to determine the relative difficulty of tracking tasks with analogue and digital altimeters.

A4.4.2 Variables Affecting Performance

There are a number of factors affecting an operator's performance, such as the instructions given, fatigue, stress, and so on, which are not taken into account in control-engineering models, but may have a major effect on performance. In particular, lack of control of these factors may invalidate the results of experiments involving the measurement of human performance. For example, the instructions given to an operator required to perform a control task are generally indadequate to define the optimal automatic controller for the task. Hence, one must assume that the operators make additional assumptions in performing the task, and these may differ between operators. In these circumstances, a difference in performance between operators may be due to their using performance criteria differing from those of the experimenter, rather than, for example, due to a difference in learning capabilities in different training regimes.

Leonard (1960) hypothesizes that the human operator adopts a mean square error criterion in optimizing his performance, but the evidence he presents, whilst not negating this, does not necessarily indicate that this is so, since all error measures are closely correlated, and, in many situations, optimization of one automatically optimizes several others. Miller (1965) demonstrated that a human operator changed his control strategy radically in response to changes in the performance measure, indicated to him by continuous visual feedback of performance. Ward and Senders (1966) demonstrated that the operator, in a similar situation to that adopted by Miller, was able to obtain better performance when <u>instructed</u> about the performance criteria than when they were displayed to him continuously.

The effects of instructions and linguistic interaction in general on the learning and performance of perceptual-motor skills is not well understood, although the studies referenced above, the developmental studies of Luria(1961), and the effects on the parameters of linear models, referenced in Section A4.2.2, show that major effects occur. Lewis and Cook (1969) have suggested that an analysis of human verbal inter-action may be simplified by restricting it to the act of 'telling', and that 'telling' may occur by signs which are not necessarily verbal - for example, the supplementary performance feedback used by Miller (1965). Lewis and Cook emphasize that in telling the person emitting information has no feedback as to the use made of it by the recipient, and this phenomenom has been a major one in studies of the effects of supplementary performance feedback on learning. Kinkade (1963) reviewed previous experiments, some of which indicated that performance indication to the operator improved the level of performance which he had reached when the supplementary feedback was removed, and others of which indicated that his performance deteriorated to that of a control group, and suggested that performance feedback also could be used as error feedback in a 'hill-climbing' control strategy if normal error feedback was poor; his subsequent experiments confirmed this hypothesis.

The effect of stress on performance is also difficult to determine, mainly because 'stress' is a term covering a variety of phenomena (Chiles 1957), and there is no reason, in advance, why they should all cause similar effects. Whilst, in everyday discourse it would be accepted that the danger of imminent death, for example is a situation creating stress, in psychology the term has come to mean almost any effect causing deterioration in the operator's performance of the main task, not induced by the performance of other physical or mental skills. Garvey and Henson (1959) generalized the term even further by calling the effect of secondary tasks, 'task-induced stress', a use not inconsistent with the possibility of using physiological measures of stress as alternative to secondary tasks.

In practice, the experimental interest in stress relative to human performance is not so much in the nature of stress itself, but in its induction in order to test the robustness of an acquired skill. To this end Mackworth (1950) used the effects of tear gas; Bersh, Notterman and Schoenfield (1957) used classical conditioning of an electric shock to a tone; Walker(1963) used an auditory 'shadowing' task; and Eason (1963) and Taylor (1964) have measured the 'stress' induced by the main tracking task, using generalized muscular activity, and galvanic skin response, respectively.

A4.5 Training

The acquisition of perceptual-motor skilled behaviour, and the variables affecting it, have been subject to much study, and the main body of experimental data has been reviewed by Bilodeau and Bilodeau (1961) and Bilodeau (1966). In the following sections the concepts of transfer-of-training and task-difficulty are analysed, experiments on the use of performance feedback, guidance and pacing are reviewed, and previous experiments on feedback training are examined.

A4.5.1 Transfer of Training

Superficially, the evaluation of the relative merits of different training techniques is straightforward - in concept, it is resolved by a comparison of the performance of an operator after training under each of the possible regimes. However, in practice the evaluation is made very complex by the irreversibility of human learning, the variety of performance criteria possible, and the variations in experimental design possible within the same abstract framework. Some of these factors have been analysed theoretically in previous chapters, and, in particular, the analysis of Chapters 2 and 3 provides a suitable foundation for resolving the complexity.

The impossibility of erazing human learning entails that one operator cannot be compared against himself, but that populations of operators, assumed homogeneous, must be compared. This, in itself, creates problems - for example, an operator may perform one task to a higher standard than he does another, and yet a second operator may have a reverse range of performances. It is clear that learning and the relative difficulties of different tasks will vary from individual to individual, and, by comparing populations one creates anomalies. The scope for possible anomaly is immensely widened by the variety of performance criteria which may be applied. For example, suppose that the numerical values of a performance measure after training under two different regimes are - Group A (1,1,1,1,3,3,3,3), Group B (2,2,2,2,2, 2,2,2). If the criterion of satisfactoriness is set at 2.5, then the training that produced Group A is clearly best. If it is set at 1.5, then that producing Group B is best, and if the 'mean' performances of the groups are compared, there is no difference.

The variety of possible experimental designs is best illustrated by temporal effects in learning. If the learning of a particular skill is dependent only on the time spent in practicing relevant tasks, then an experimental design in which one group practices one task, and then is compared with a control group in learning or performing a second task, may show that the first task is relevant to the second, in that positive transfer occurs, but does not indicate whether training on the first task is useful - as Day (1956) notes, 'transfer' from one task to itself (called 'fixed training' in Chapter 3) may be better than transfer from other tasks. The situation becomes more complex if differential rates of learning are coupled with differential final levels of acquired skill. If one training techniques leads to a rapid initial rise in performance but a low level of final skill, and another technique has the converse effect, then the mean performances of operators trained by the two techniques will cross over at some point in time. Hence, comparison of the two groups will be entirely dependent on the length of the experimental training period.

The methodological problems in the evaluation of the evaluation of transfer of training have been analysed in great detail by Gagné, Foster and Crowley (1948). However, in later studies problems of a semantic nature have become apparent, particularly those concerned with the effects of relative task difficulty on transfer. Bartlett (1947) noted that tasks have a natural topology such that wide variations can be made in physical factors without any corresponding change in human performance, and states that, 'The fundamental features of a performance will remain stable over a certain range of its conditions. Outside this range they will change often in a dramatic or radical manner'. Helson (1949) has made a similar suggestion, and has demonstrated, in a limited range of tracking tasks, that a U-shaped curve is obtained on plotting performance against variation of a variety of task parameters.

Gibbs (1951) extends the concept of zones of equal performance to account for variations in ease of transfer between learning on one task and performance of another. In stating his conclusions, Gibbs uses the term, 'task difficulty', which he earlier introduces implicitly as being determined by differences in the mean performances of two groups of operators, assumed matched. He states, 'It appears that the amount of transfer between two equally difficult tasks may be equal, whereas the transfer between two unequal tasks may be unequal. There may be greater transfer from the difficult to easy task than from the easy to difficult, if the same kind of ability is required in both tasks and learning is carried on until the total possible skill is closely approached in both tasks.'

Day (1956) and Holding (1962) review experiments on the transfer of training between 'easy'and 'difficult' tasks, and vice versa, and conclude that no simple and universal prediction of asymmetrical transfer in these terms is possible. Holding states that, 'the use of the concept of difficulty must give may to far more detailed analysis of the appropriate skills, if asymmetrical transfer is to be predicted.' However, the concept of 'difficulty' is a natural and attractive one, which cannot be outlawed from the design and analysis of experiments, even if it is outlawed from the published discusssions of them, and it is worth noting some possible sources of the diverse results in transfer experiments and their implications.

Holding (1962) notes, as mentioned above, that both the ultimate level of performance and the rate of learning may vary with task difficulty. Gibbs (1951) specifically excludes effects due to the variation in rate of learning from his statement of expected transfer, by requiring that learning continue until the total possible skill is approached. He also suggests that, 'there is an optimal level of task difficulty ... for every kind of learning material', presumably one which maximizes the rate of learning. Taken together with his explanation of asymmetrical transfer in terms of a hierarchy of abilities, the higher of which encompass performance of the lower, a reasonable statement of the direction of asymmetrical transfer would be of the form -

If there is a natural continuum of tasks generated by a physical parameter of the environment, and there is either a U-shaped or monotonic variation in the difficulty in the tasks along this continuum - as measured by the performance levels of a group of operators performing each of the different tasks, then learning of a more difficult task on the continuum, providing it takes place to its ultimate level, will give better performance on the transfer task than learning on an easier task. However, the maximum rate of learning does not necessarily take place with either the task of best performance or the task of maximum difficulty.

This statement, although it sets out the theses and qualifications of previous hypotheses in some detail, is still amenable to different interpretations, largely because of the vagueness of the terms, 'difficulty' and 'rate of learning'. If an operator is performing one task as training for another, then it is not the rate of learning of the first which we wish to maximize, but rather the rate of learning of the second - conceptually, the operator's control policy should be 'frozen' at intervals whilst he is learning the first task and his performance measured on the second task to plot a true 'learning curve'; one advantage of working with automatic adaptive controllers is that this conceptual procedure can actually be carried out. Furthermore, inherent in the requirement that the rate of learning should be maximized is the assumption that an operator's state, in so far as his further learning is concerned, is determined solely by his performance. Although this assumption is not valid in general, and counter-examples may always be found, it is a useful working hypothesis in practical training situations.

The greatest problem in the use of terms such as 'difficulty' and 'rate of learning' is that performance measures give only ordinal, not interval scales, and that any monotone transformation of a performance measure is valid. Hence, the only valid comparison of rates of learning is in terms of the time interval to change from one level of performance to another, and only the relative ease or difficulty of two tasks is defined. The possible effects of a monotone change in the scale of performance are further compounded by the possibility of a similar change in the scale of the underlying physical variables - hence, the <u>shape</u> of a U-shaped curve of performance is meaningless. Furthermore, some common monotone transformations, particularly logarithmic ones, may eradicate parts of the scale of physical variables and turn a U-shaped function into a monotonic one, giving a false appearance of graded difficulty - for example, in tracking tasks involving exponential lags, positive lags only are generally considered and increasing difficulty with increasing lag is discovered; increasing negative lag, however, corresponding to an unstable system, also gives increasing difficulty.

A4.5.2 Feedback Training

Whilst the arguments of the previous section indicate that, ultimately, a detailed analysis of learning behaviour, of the type discussed in Chapters 2 and 3, is essential to predict all the phenomena of transfer, the concepts of transfer between levels of 'difficulty', and maximizing the 'rate of learning', outlined in the previous section, form the basis of an 'approximation' of the type discussed in Appendix 3. The basis for determining the optimum level of difficulty for learning, in terms of maintaining the desired sub-environment has been discussed in Chapter 3. Accepting the statement of degree of transfer and rate of learning, given in the last section, as a reasonable approximation, however, one point it brings out clearly is that the level of difficulty in training which gives the best transfer is not necessarily that which maximizes the rate of learning. Hence, if training in the shortest possible time to the highest possible level is the objective, the difficulty of the task should vary with time. Since the future learning of an operator is also assumed to be determined by his present performance of a particular task, the variation in task difficulty should be a function of the present difficulty and performance - this leads directly to the 'feedback trainer' described in Chapter 3.

The earliest mention of the possibility of feedback training for perceptual-motor skills appears to be that of Stockbridge and Siddall (1956), who suggested that a guided weapons tracking trainer be used in which, 'the difficulty of the task is proportional to the success of the subject'. A short exposition by Senders (1961) of the principles of 'adaptive teaching machines' is an example of the many studies of feedback training which have been reported only informally. In a number of papers, Pask (1960,1961,1964,1965*) has made available a deep and comprehensive analysis of automated training, and has placed it in the general context of interactions between self-organizing systems; he has applied these principles to cognitive skills rather than perceptual-motor skills.

Actual equipment for feedback training has been described in detail by Ziegler, Birmingham and Chernikoff (1962), as a 'teaching machine for the selection and training of operators of higher order vehicles'. The task was compensatory tracking in two dimensions with the same third-order dynamics in both axes, consisting of three integrators in cascade with variable feed-forward (so-called 'quickening') around them. The amount of feedforward was controlled by a servomechanism driven by the smoothed sum of the mean error modulus in both axes. Hence, as operators learnt their mean error was expected to decrease and the task-difficulty to increase. No formal experiments have been reported with this system, although Chernikoff (1962) indicates that it appears to result in improved training; the discussion following his paper is particularly informative. A similar feedback trainer was proposed at the same time by Kelley (1962).

Other studies of equipment relevant to feedback training have been carried out by Jex, McDonnell and Phatak (1966, 1966*, McDonnell and Jex 1967), using a first-order divergent system with variable divergence rate; this system has been investigated solely for testing. Kelly (1967, and Prosin 1968) has carried out extensive tests of various forms of task with performance feedback for personnel evaluation, and his equipment and the ensuing discussion are both relevant to the design of feedback trainers.

Some training aids which have previously been shown to affect learning, and whose magnitude may be varied along a continuum, are obvious candidates for feedback training since the aid must eventually be withdrawn and performance feedback may be used to schedule the withdrawal. For example, Holding and Macrae (1966) have demonstrated that a 'hinting' device, which makes the joy-stick in a complex tracking task easier to move in the correct direction, has a profound influence on the rate of learning; descreasing 'hinting' to a subliminal level as learning proceeds offers obvious possibilities for feedback training. Similarly, the degree of 'augmented feedback' is another variable susceptible to continuous variation, and Briggs (1961,1962) has investigated the scheduling both of display aiding and extra performance feedback. In the first paper he reports that there is an optimum level of display aiding, and that experience on systems with either too much, or too little, aiding leads to reduced learning. He suggests that, 'optimum schedules for display aiding be determined for each device or task'. In the later paper, Briggs (1962) investigates the effect of slowly withdrawing augmented feedback in a tracking task according to various schedules, but finds no significant improvement in learning over a control group without augmented feedback.

The only major experimental study of feedback training is that of Hudson (1964) who trained some 72 operators for ten hours each (in fifteen minute periods) on a third-order system with variable parameters including both feedforward and feedback. Hudson mechanized an automatic feedback training loop in the same way as Ziegler et al (1962), by relating the parameters of the task directly to the mean error. However, he found this was not successful since the parameters both varied widely with the error, and the mean performance required became excessively high if the ultimate levels of difficulty were to be attained. In his conclusions, he suggests that the automatic loop should have been set up to keep the mean error constant by varying the difficulty of the task.

In his experiments, Hudson maintained approximately constant error conditions for some groups of operators by putting himself in the adaptive loop, and adjusting the parameters of difficulty by hand so that the operator's mean score of 'out-of-controls' (error becoming so great that spot leaves screen and is reset) was constant over each fifteen minute training run. Hudson's main result is that a plot of the final test performance on the criterion task against the mean level of performance during training is strongly U-shaped, and there is a very clear optimum level of difficulty for maximum transfer. Another result of particular interest is that Hudson used a variety of plant parameter variations to maintain the performance constant, but only the actual level of difficulty seems to affect the main result.

Significant as they are, Hudson's results may be criticized on a number of counts: firstly, because of the variety of conditions used, the number of subjects under each condition is small (between 4 and 6) so that the results are only marginally significant; secondly, each 'out-of-control' took a minimum of five seconds of the operator's training time whilst the equipment was reset, since the mean number of times this occurred (over the ten-hour training period) was 77 in a fifteen minute practice session, there is clearly the possibility of a major effect due to the differences in actual training times - this is not unrealistic, in the sense that it is clearly a loss of the 'desired sub-environment', but represents a trivial variation in the difficulty of the task; thirdly, although highly suggestive, the experiments do not test the viability of automatic feedback training, and, since the experimenter was part of the loop, the results are not replicable.

Gaines (1965, 1966, 1966*, 1967, 1968*, 1968**, 1968***, 1968****) has published a series of reports and papers on automated feedback training, its theoretical foundations, viability and utility. The material from these is incorporated in the various chapters of the present volume, which represents a complete account of this work.

APPENDIX 5 EXPERIMENTAL RESULTS

A5.1 Introduction

The experimental situation is described in Chapter 5, together with an analysis of the results. This appendix contains the raw data and statistics of the experiment with 72 RAF pilots, conducted to investigate the utility of feedback training. It also includes plots of the trajectory of difficulty against time for the 32 operators who underwent feedback training.

A5.2 Notes on Raw Data

Table A5-1 gives the results of the experiments on training RAF pilots. The subject numbers are individually arbitrary, but operators have been grouped into the six groups, Hw, Hs, Lw, Ls, Fw, Fs, defined in Table 5-2. The test results are the mean error scores over four minutes of $Test_1$ through $Test_5$ of Table 5-3, and the level of difficulty is also noted. The units of error are the same as those of Chapter 4, but multiplied by 1,000 for convenience. α_1 and α_2 are the peak levels of difficulty attained by the feedback training group during the first and second training sessions respectively, multiplied by 100.

TIME is the estimate given on the questionnaire of Figure 5-2, in minutes. The next four readings are the distances in millimetres from the left-hand edge of the scales of Figure 5-2 at which the subjects made a mark (these scales were 100 mm long on the original questionnaires); they represent the operator's estimates of his INTerest in the task, the DIFFiculty of the task, his PRESent level of performance, and his FUTure level of performance. The OPTimal time reading is from the second questionnaire, Figure 5-4, and taken from the marked scale, assuming the 'present length' to be 25 minutes. The total number of WORDs written on all three questionnaires is given in the next column. The RAF score is the total mark obtained from normal ratings during the operator's RAF training of personal qualities, examination performance and flying ability; in the present context it is used only to check the comparability of the six experimental groups, membership of which was assigned 'at random' since the RAF data was not available before the experiment. The final column, GRAPH, for the feedback group, refers to the trajectory of difficulty in Figure A5-1.

A5.3 Notes on Statistics

Table A5-2 gives the means and variances of the data of Table A5-1, within each of the six groups, and enables groups to be compared for differences in mean and variance by the t-test and variance-ratio test; values significant at the one per cent level are bracketed, one-tailed for the six results on performance.

These statistics are intended only as an indication of the magnitude of the differences between the groups apparent in Figures 5-5 and 5-6, and in the means and variances. Non-parametric tests would give a more correct estimate of the probability of chance occurrence of the phenomena of interest.

A5.4 Notes on Graphs

Figure A5-1 is a complete set of trajectories of difficulty against time for the 32 operators under feedback training. Numbers 1 through 16 are Fs, and 17 through 32 are Fw; the graphs may be related to other data through the last column of Table A5-1.

NO.	MODE	Test Results 1 (0.5)	2 (0.5)	3 (0.5)	4 (0.25)	5 (0.7)	α_1	α_2	TIME	INT.	DIFF.	PRES.	FUT.	OPT.	WORD	RAF
01	Hw	586	534	576	193	640	--	--	30	76	34	27	63	25	030	721
02	Hw	598	585	513	555	670	--	--	10	69	42	24	80	31	047	716
03	Hw	564	577	583	712	700	--	--	12	51	11	00	53	25	131	705
04	Hw	596	585	583	357	666	--	--	25	71	26	13	21	28	099	686
05	Hw	571	559	606	645	647	--	--	20	21	50	43	43	05	044	699
06	Hw	562	567	587	730	671	--	--	10	37	42	32	29	12	113	634
07	Hw	565	565	573	777	625	--	--	20	88	36	31	32	18	091	710
08	Hw	563	625	617	250	600	--	--	20	81	20	01	39	52	024	640
Mean		576	573	580	527	652	--	--	18	62	33	21	45	25	073	689
09	Hs	597	554	545	721	692	--	--	18	58	25	25	44	21	212	696
10	Hs	529	464	499	427	654	--	--	30	37	19	23	48	12	113	646
11	Hs	573	474	411	164	563	--	--	30	73	43	12	37	25	204	725
12	Hs	571	571	605	701	684	--	--	20	28	26	24	24	17	157	769
13	Hs	572	571	607	748	648	--	--	15	77	58	39	68	15	138	724
14	Hs	534	536	571	377	660	--	--	20	22	31	06	17	09	233	718
15	Hs	581	551	580	607	654	--	--	25	83	32	28	60	18	190	697
16	Hs	535	507	487	330	630	--	--	22	25	42	25	25	16	86	666
Mean		561	529	538	510	648	--	--	23	50	35	23	40	17	167	705
17	Lw	582	538	579	550	648	--	--	10	25	16	35	66	22	259	668
18	Lw	572	488	421	240	604	--	--	15	39	22	46	61	25	084	699
19	Lw	611	554	567	512	651	--	--	20	82	50	26	45	17	161	687
20	Lw	554	605	590	515	655	--	--	23	03	08	01	00	11	010	711
21	Lw	284	222	282	152	402	--	--	10	70	38	59	73	25	178	701
22	Lw	644	595	580	601	638	--	--	20	58	47	49	48	25	130	720
23	Lw	579	448	476	183	620	--	--	26	58	44	44	61	22	064	716
24	Lw	563	464	395	170	487	--	--	20	30	29	22	33	90	166	694
25	Lw	628	573	575	748	632	--	--	20	81	33	18	41	21	168	785
26	Lw	564	571	561	775	659	--	--	26	85	41	52	51	23	151	648
27	Lw	583	563	568	740	650	--	--	25	60	45	34	39	25	033	680
28	Lw	556	505	473	295	596	--	--	25	51	02	10	06	20	021	674
Mean		560	510	506	457	603	--	--	20	54	31	33	44	27	119	699
29	Ls	590	345	371	214	592	--	--	25	68	50	25	49	25	033	729
30	Ls	318	203	220	122	388	--	--	20	83	50	16	34	19	130	690
31	Ls	580	287	241	145	390	--	--	20	74	50	04	83	25	056	705
32	Ls	356	201	193	104	269	--	--	20	96	51	27	68	24	109	652
33	Ls	478	296	252	140	394	--	--	06	98	47	76	82	25	232	663
34	Ls	317	227	228	157	326	--	--	25	79	66	69	78	19	356	675
35	Ls	431	402	347	183	580	--	--	24	45	50	59	66	19	143	689
36	Ls	534	578	507	346	591	--	--	20	59	48	50	64	25	246	717
37	Ls	360	288	274	197	449	--	--	20	99	50	72	69	25	026	673
38	Ls	615	439	455	223	562	--	--	20	48	50	23	14	27	089	740
39	Ls	479	191	219	131	302	--	--	25	86	50	69	80	25	135	725
40	Ls	622	551	612	492	641	--	--	20	99	35	40	41	25	133	656
Mean		473	334	327	204	457	--	--	20	78	50	44	61	24	141	693

Table A5-1 Performance and Responses of Experimental Subjects

NO.	MODE	Test Results 1 (0.5)	2 (0.5)	3 (0.5)	4 (0.25)	5 (0.7)	α_1	α_2	TIME	INT.	DIFF.	PRES.	FUT.	OPT.	WORD	RAF	GRAPH
41	Fw	532	311	270	114	353	39	52	15	99	43	28	28	21	081	768	20
42	Fw	451	232	250	181	377	49	68	08	60	45	38	67	16	072	740	17
43	Fw	189	232	185	083	197	74	70	15	98	40	38	65	25	157	758	19
44	Fw	403	353	215	127	322	30	65	30	64	50	64	88	25	071	709	25
45	Fw	590	520	530	251	591	33	47	30	70	39	29	35	25	299	756	26
46	Fw	498	578	451	191	564	28	40	30	57	39	39	38	25	085	711	28
47	Fw	580	499	390	212	520	31	38	15	49	43	50	52	25	149	728	18
48	Fw	610	450	450	226	568	30	43	30	48	50	56	68	18	058	714	23
49	Fw	338	308	252	092	484	53	73	15	99	39	27	44	25	096	650	31
50	Fw	541	525	496	206	622	30	30	15	69	43	40	59	21	095	717	30
51	Fw	260	398	239	145	497	62	62	14	66	45	31	42	16	071	652	21
52	Fw	541	438	388	254	507	29	40	20	33	49	54	82	16	046	754	32
53	Fw	481	358	345	175	534	43	52	15	69	35	28	66	28	135	664	22
54	Fw	388	233	168	076	282	49	67	30	99	50	18	28	19	073	692	29
55	Fw	642	572	508	292	581	20	28	25	46	32	17	17	20	099	650	24
56	Fw	474	283	218	214	447	50	53	26	57	46	52	48	23	131	644	27
Mean		470	393	334	177	465	41	52	21	68	43	38	52	22	107	706	--
57	Fs	358	357	167	129	238	52	62	20	50	40	44	56	25	078	742	11
58	Fs	519	332	155	071	203	32	68	15	69	49	54	80	18	120	668	08
59	Fs	563	475	446	192	532	33	41	15	79	45	11	49	19	091	705	02
60	Fs	417	333	255	134	353	46	58	30	10	29	41	53	32	072	723	14
61	Fs	515	403	340	134	523	35	48	20	33	43	50	66	18	102	721	12
62	Fs	572	421	301	122	501	35	48	15	00	24	27	36	12	048	680	10
63	Fs	452	287	221	090	319	35	66	20	85	50	36	74	25	134	685	06
64	Fs	282	225	167	075	241	58	76	25	71	42	51	50	18	094	690	13
65	Fs	205	227	142	105	176	58	82	22	99	44	47	69	25	089	654	15
66	Fs	326	196	201	108	333	63	76	10	68	65	31	51	25	072	678	07
67	Fs	326	261	278	128	378	55	69	24	99	67	54	55	27	094	704	09
68	Fs	346	149	147	099	256	56	78	22	71	46	20	41	35	061	744	05
69	Fs	233	171	137	085	201	58	83	15	76	50	69	90	25	093	702	03
70	Fs	486	375	356	267	508	33	43	05	84	37	21	34	20	126	677	16
71	Fs	370	226	142	063	311	66	80	15	71	40	45	66	17	179	660	01
72	Fs	510	400	363	186	474	37	53	17	77	57	70	84	16	143	626	04
Mean		405	303	238	124	347	47	64	18	65	46	42	60	22	100	691	--

Table A5-1 Performance and Responses of Experimental Subjects

Performance on First Test

			t-tests						var-ratios					
Mode	Mean	Var.	Hw	Hs	Lw	Ls	Fw	Fs	Hw	Hs	Lw	Ls	Fw	Fs
Hw	0.576	0.000	0.0	1.4	0.5	2.5	2.3	(4.1)	1	3	(37)	(59)	(72)	(60)
Hs	0.561	0.001		0.0	0.0	2.1	2.0	(2.7)		1	(14)	(22)	(27)	(23)
Lw	0.560	0.008			0.0	2.1	2.1	(3.8)			1	2	2	2
Ls	0.473	0.012				0.0	0.1	1.5				1	1	1
Fw	0.470	0.015					0.0	1.5					1	1
Fs	0.405	0.013						0.0						1

Performance on Second Test

			t-tests						var-ratios					
Mode	Mean	Var.	Hw	Hs	Lw	Ls	Fw	Fs	Hw	Hs	Lw	Ls	Fw	Fs
Hw	0.573	0.001	0.0	2.5	1.7	(5.0)	(4.1)	(7.5)	1	2	(14)	(23)	(19)	(13)
Hs	0.529	0.002		0.0	0.5	(4.0)	(3.0)	(6.2)		1	(06)	(10)	(09)	(06)
Lw	0.510	0.010			0.0	(3.6)	(2.7)	(5.4)			1	2	1	1
Ls	0.334	0.016				0.0	1.2	0.7				1	1	2
Fw	0.393	0.014					0.0	2.4					1	2
Fs	0.303	0.009						0.0						1

Performance on Third Test

			t-tests						var-ratios					
Mode	Mean	Var.	Hw	Hs	Lw	Ls	Fw	Fs	Hw	Hs	Lw	Ls	Fw	Fs
Hw	0.580	0.001	0.0	1.6	2.1	(5.2)	(5.4)	(9.5)	1	5	(10)	(20)	(17)	(11)
Hs	0.538	0.004		0.0	0.8	(4.1)	(4.3)	(7.7)		1	2	4	4	2
Lw	0.506	0.009			0.0	(3.7)	(3.9)	(7.1)			1	2	2	1
Ls	0.327	0.017				0.0	0.2	2.0				1	1	2
Fw	0.334	0.015					0.0	2.3					1	2
Fs	0.238	0.009						0.0						1

Performance on Fourth Test

			t-tests						var-ratios					
Mode	Mean	Var.	Hw	Hs	Lw	Ls	Fw	Fs	Hw	Hs	Lw	Ls	Fw	Fs
Hw	0.527	0.046	0.0	0.2	0.7	(4.2)	(5.8)	(6.8)	1	1	1	4	(11)	(18)
Hs	0.510	0.040		0.0	0.5	(4.2)	(5.8)	(6.9)		1	1	4	(10)	(15)
Lw	0.457	0.052			0.0	(3.7)	(4.5)	(5.4)			1	5	(13)	(20)
Ls	0.204	0.011				0.0	0.8	(2.5)				1	3	4
Fw	0.177	0.004					0.0	(2.6)					1	2
Fs	0.124	0.003						0.0						1

Performance on Fifth Test

			t-tests						var-ratios					
Mode	Mean	Var.	Hw	Hs	Lw	Ls	Fw	Fs	Hw	Hs	Lw	Ls	Fw	Fs
Hw	0.652	0.001	0.0	0.2	1.7	(4.2)	(4.2)	(6.7)	1	2	7	(18)	(17)	(17)
Hs	0.648	0.001		0.0	1.5	(4.0)	(4.0)	(6.6)		1	4	(11)	(10)	(11)
Lw	0.603	0.006			0.0	(3.3)	(3.4)	(6.2)			1	3	3	3
Ls	0.457	0.015				0.0	0.2	2.3				1	1	1
Fw	0.465	0.014					0.0	(2.7)					1	1
Fs	0.347	0.015						0.0						1

Table A5-2 Statistics of Experimental Results

Performance Change Under Stress

Mode	Mean	Var.	t-tests						var-ratios					
			Hw	Hs	Lw	Ls	Fw	Fs	Hw	Hs	Lw	Ls	Fw	Fs
Hw	-0.005	0.001	0.0	0.2	0.6	0.8	(3.1)	(3.0)	1	1	1	1	2	3
Hs	-0.010	0.001		0.0	0.8	1.0	(3.3)	(3.2)		1	1	1	2	3
Lw	0.005	0.001			0.0	0.2	(3.0)	(3.0)			1	1	2	2
Ls	0.007	0.001				0.0	(2.8)	(2.8)				1	2	2
Fw	0.059	0.003					0.0	0.3					1	1
Fs	0.064	0.003						0.0						1

Estimated Time of Training Session

Mode	Mean	Var.	t-tests						var-ratios					
			Hw	Hs	Lw	Ls	Fw	Fs	Hw	Hs	Lw	Ls	Fw	Fs
Hw	18.4	51	0.0	1.3	0.6	0.7	0.7	0.1	1	2	2	2	1	2
Hs	22.5	29		0.0	1.0	0.9	0.6	1.7		1	1	1	2	1
Lw	20.0	30			0.0	0.2	0.3	0.8			1	1	2	1
Ls	20.4	22				0.0	0.2	1.1				1	2	2
Fw	20.8	65					0.0	1.1					1	2
Fs	18.1	33						0.0						1

Interest in Tracking Task

Mode	Mean	Var.	t-tests						var-ratios					
			Hw	Hs	Lw	Ls	Fw	Fs	Hw	Hs	Lw	Ls	Fw	Fs
Hw	61.7	472	0.0	0.9	0.7	1.7	0.7	0.3	1	1	1	1	1	2
Hs	50.4	565		0.0	0.3	(2.8)	1.8	1.2		1	1	2	1	1
Lw	53.5	599			0.0	(2.7)	1.6	1.1			1	2	1	1
Ls	78.0	359				0.0	1.3	1.3				1	1	2
Fw	67.9	427					0.0	0.3					1	2
Fs	65.2	762						0.0						1

Estimated Difficulty of Tracking Task

Mode	Mean	Var.	t-tests						var-ratios					
			Hw	Hs	Lw	Ls	Fw	Fs	Hw	Hs	Lw	Ls	Fw	Fs
Hw	32.6	156	0.0	0.3	0.2	(3.9)	(2.8)	(2.5)	1	1	2	4	5	1
Hs	34.5	147		0.0	0.5	(3.6)	2.3	2.2		1	2	3	5	1
Lw	31.3	233			0.0	(3.7)	(2.8)	(2.8)			1	(6)	(9)	2
Ls	49.8	41				0.0	(3.0)	1.2				1	2	3
Fw	43.0	32					0.0	0.8					1	4
Fs	45.5	126						0.0						1

Estimated Present Performance

Mode	Mean	Var.							var-ratio					
			Hw	Hs	Lw	Ls	Fw	Fs	Hw	Hs	Lw	Ls	Fw	Fs
Hw	21.4	218	0.0	0.2	1.5	2.3	(2.7)	(2.9)	1	2	1	3	1	1
Hs	22.7	91		0.0	1.5	2.3	(2.8)	(3.0)		1	3	(7)	2	3
Lw	33.0	292			0.0	1.3	0.9	1.4			1	2	2	1
Ls	44.2	575				0.0	0.8	0.3				1	3	2
Fw	38.1	187					0.0	0.7					1	2
Fs	41.9	277						0.0						1

Table A5-2 Statistics of Experimental Results

Estimated Future Performance

Mode	Mean	Var.	t-tests Hw	Hs	Lw	Ls	Fw	Fs	var-ratios Hw	Hs	Lw	Ls	Fw	Fs
Hw	45.0	332	0.0	0.5	0.1	1.6	0.8	1.9	1	1	1	1	1	1
Hs	40.4	290		0.0	0.4	2.2	1.4	(2.6)		1	2	2	1	1
Lw	43.7	461			0.0	1.9	1.1	2.2			1	1	1	1
Ls	60.7	438				0.0	1.1	0.1				1	1	2
Fw	52.3	388					0.0	1.1					1	2
Fs	59.6	266						0.0						1

Estimated Optimal Training Time

Mode	Mean	Var.	t-tests Hw	Hs	Lw	Ls	Fw	Fs	var-ratios Hw	Hs	Lw	Ls	Fw	Fs
Hw	24.5	175	0.0	1.5	0.3	0.2	0.7	0.5	1	(8)	2	(23)	(12)	5
Hs	16.6	23		0.0	1.4	(4.0)	(2.8)	2.3		1	(17)	3	2	2
Lw	27.2	375			0.0	0.6	1.1	0.9			1	(50)	(26)	(10)
Ls	23.6	8				0.0	1.4	0.7				1	2	5
Fw	21.8	15					0.0	0.3					1	2
Fs	22.3	34						0.0						1

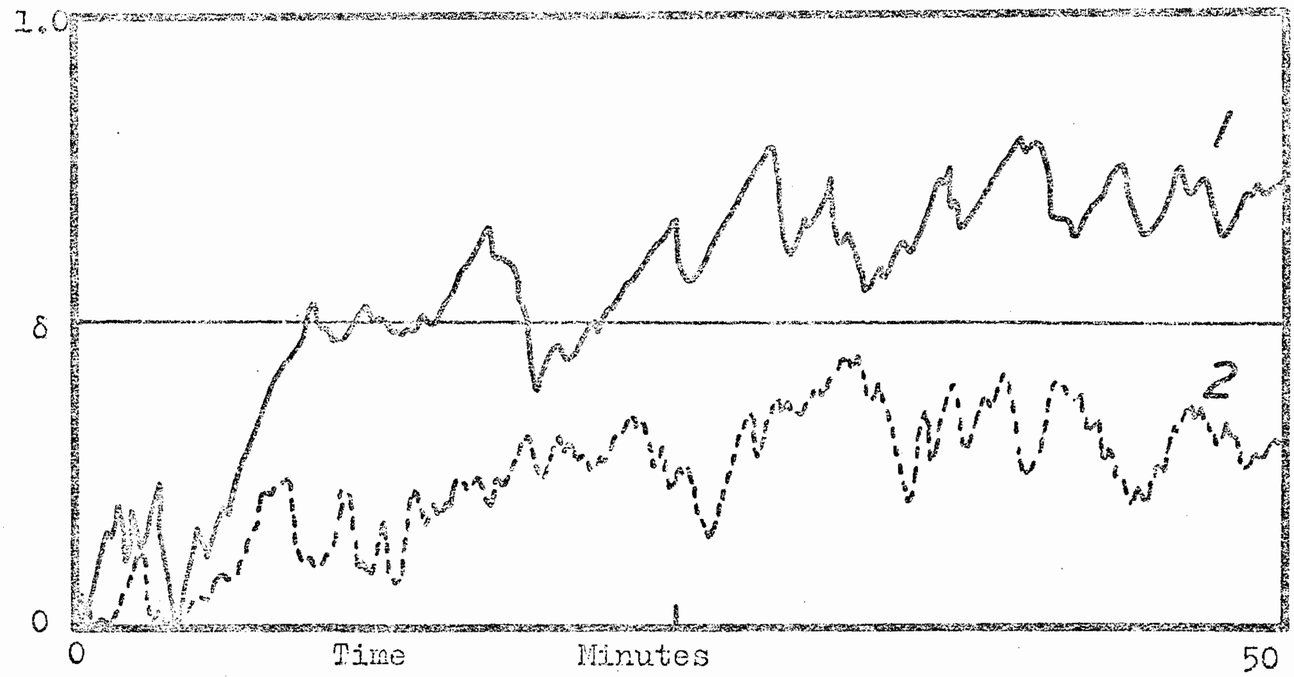
Total Words Written on Questionnaires

Mode	Mean	Var.	t-tests Hw	Hs	Lw	Ls	Fw	Fs	var-ratios Hw	Hs	Lw	Ls	Fw	Fs
Hw	73.5	1620	0.0	(3.9)	1.5	1.8	1.4	1.7	1	2	3	5	2	2
Hs	166.5	2340		0.0	1.6	0.7	2.4	(3.8)		1	2	3	2	2
Lw	118.8	5320			0.0	0.6	0.4	0.9			1	2	2	(5)
Ls	140.7	8470				0.0	1.1	1.6				1	3	(8)
Fw	107.4	3420					0.0	0.4					1	3
Fs	99.6	1080						0.0						1

RAF Evaluation

Mode	Mean	Var.	t-tests Hw	Hs	Lw	Ls	Fw	Fs	var-ratios Hw	Hs	Lw	Ls	Fw	Fs
Hw	68.9	9.9	0.0	0.9	0.6	0.3	0.0	0.1	1	1	1	1	2	1
Hs	70.5	12.8		0.0	0.4	0.8	0.6	0.9		1	1	1	1	1
Lw	69.9	10.8			0.0	0.4	0.4	0.6			1	1	2	1
Ls	69.3	8.3				0.0	0.2	0.1				1	2	1
Fw	70.6	15.2					0.0	0.9					1	2
Fs	69.1	9.6						0.0						1

Table A5-2 Statistics of Experimental Results



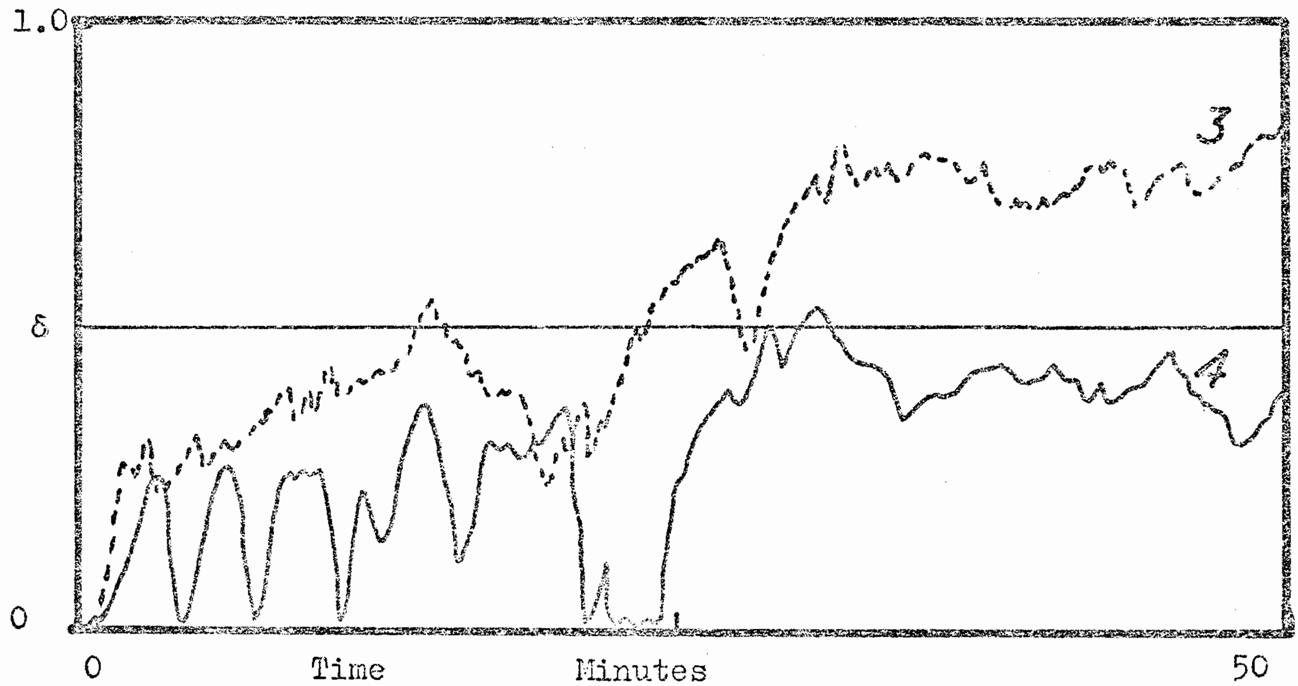


Figure A5-1 Trajectories of Difficulty

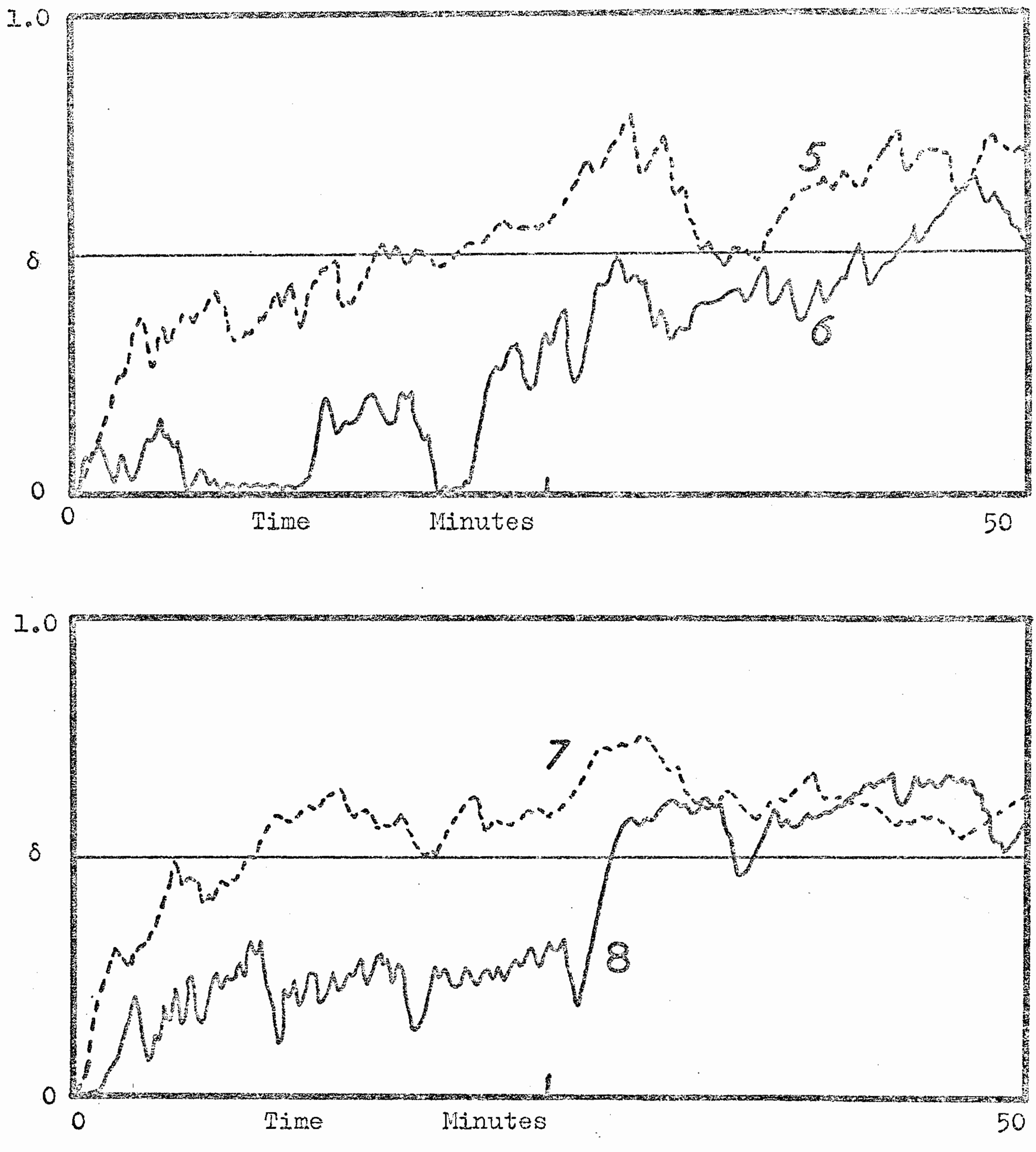


Figure A -1 Trajectories of Difficulty

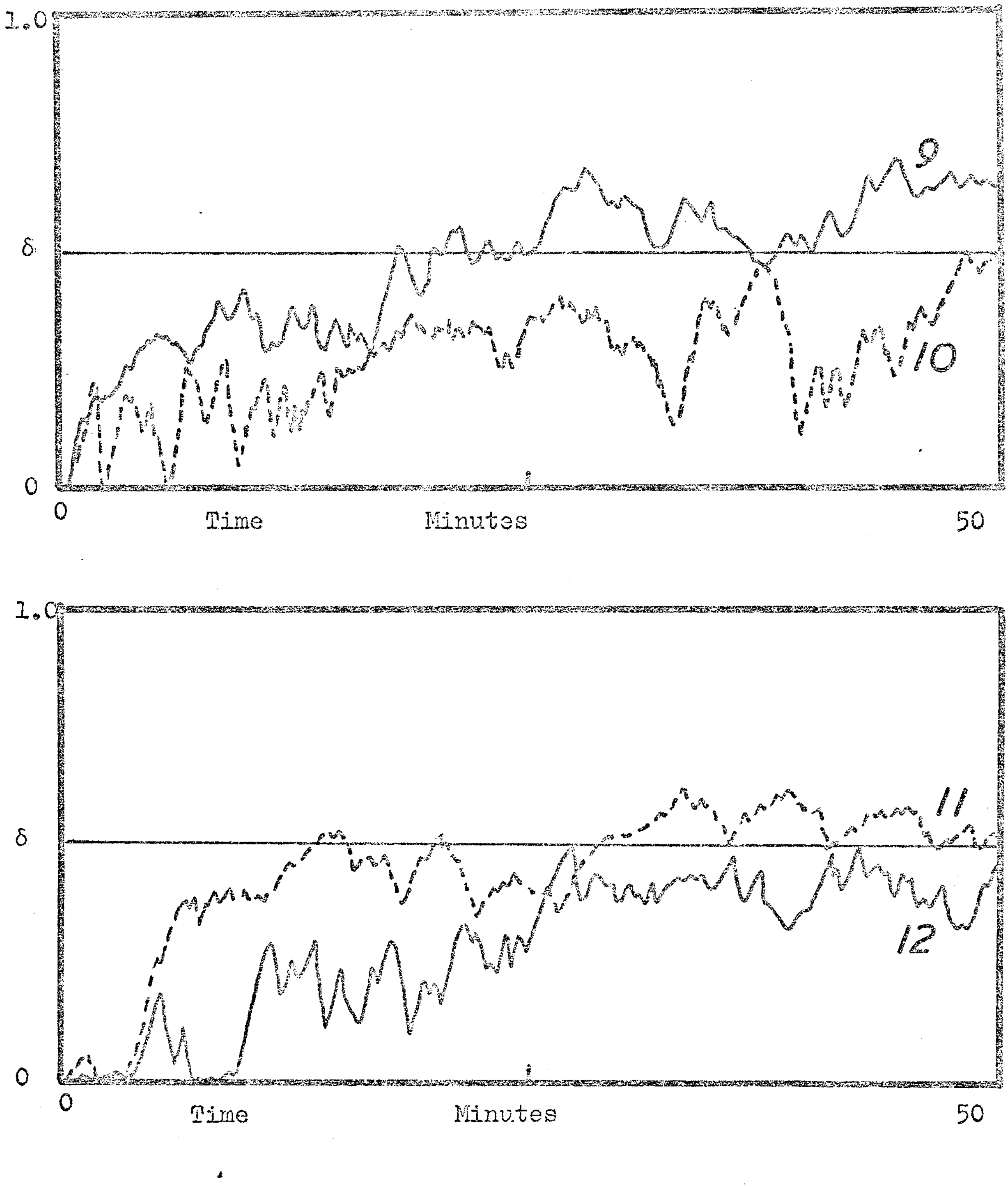
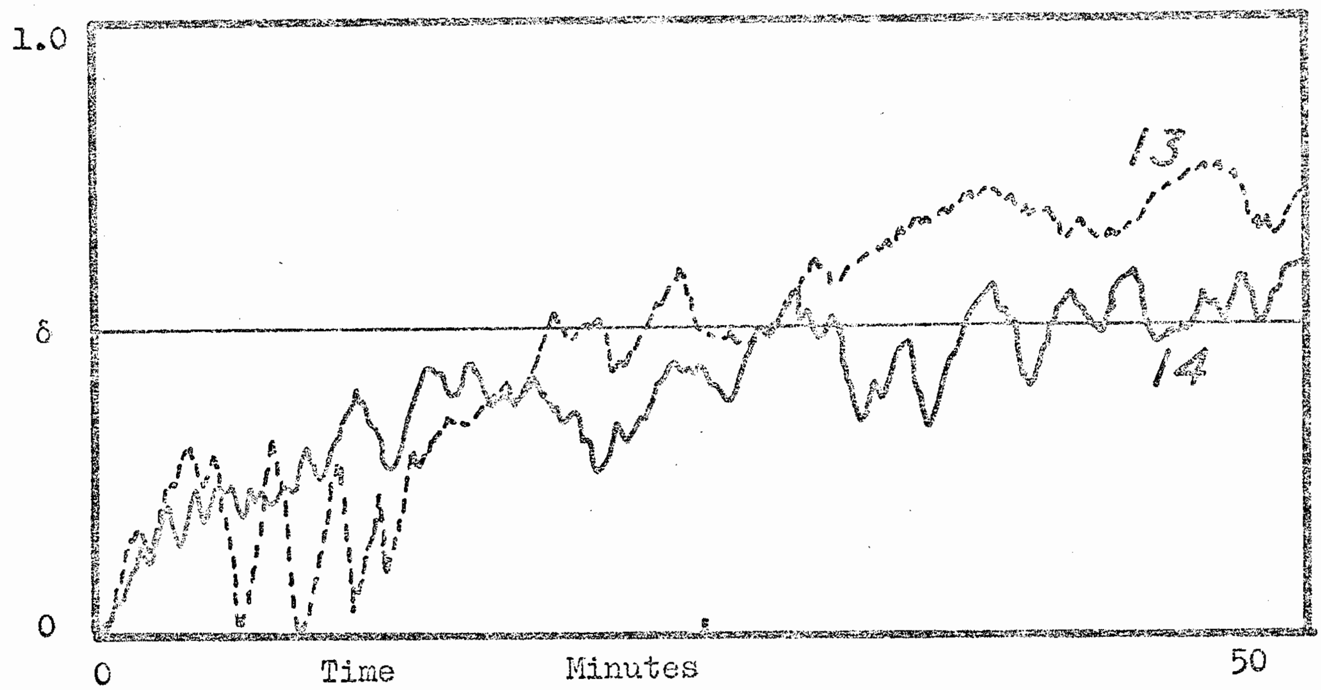


Figure A.-1 Trajectories of Difficulty



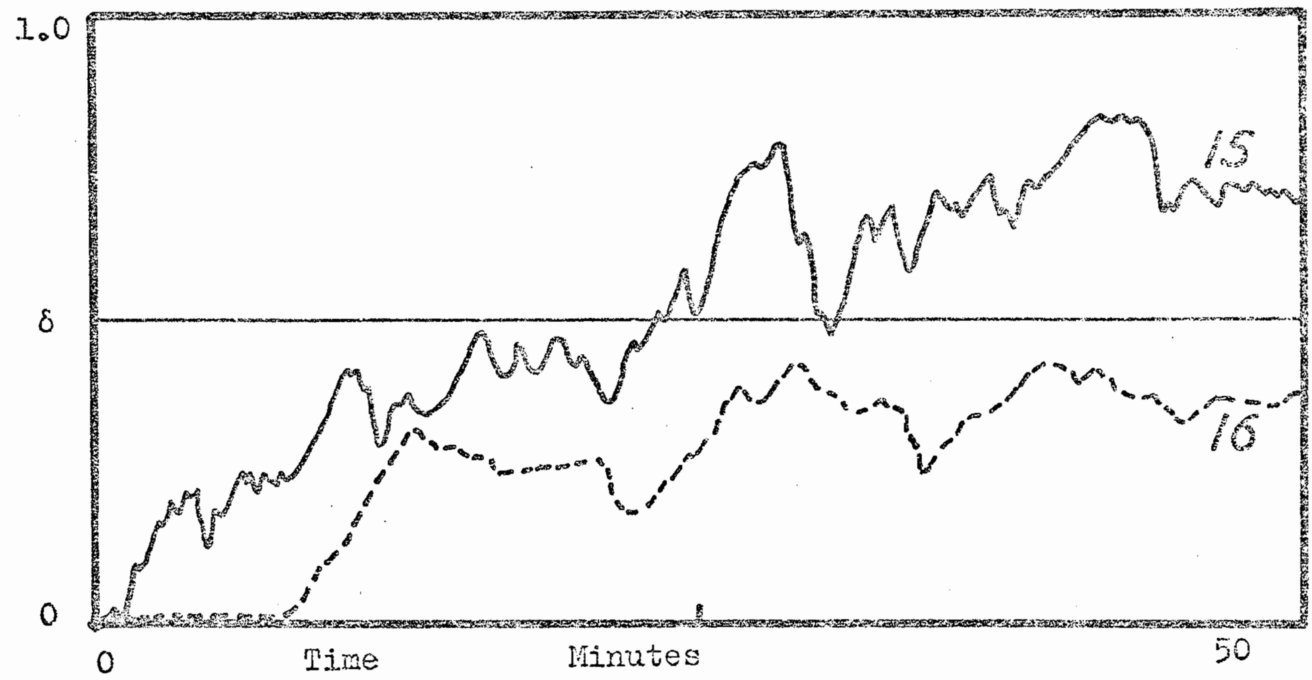


Figure A.-1 Trajectories of Difficulty

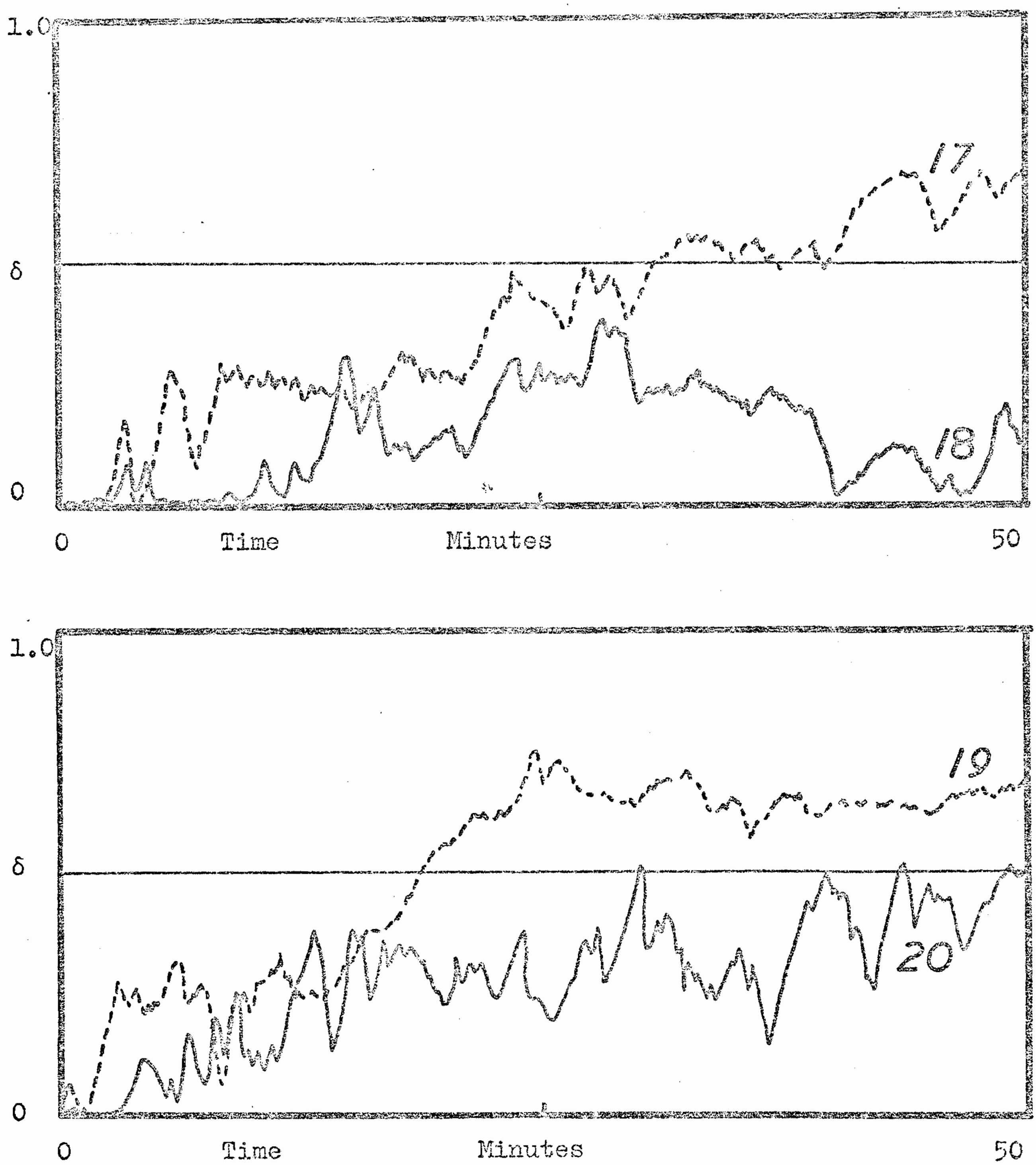
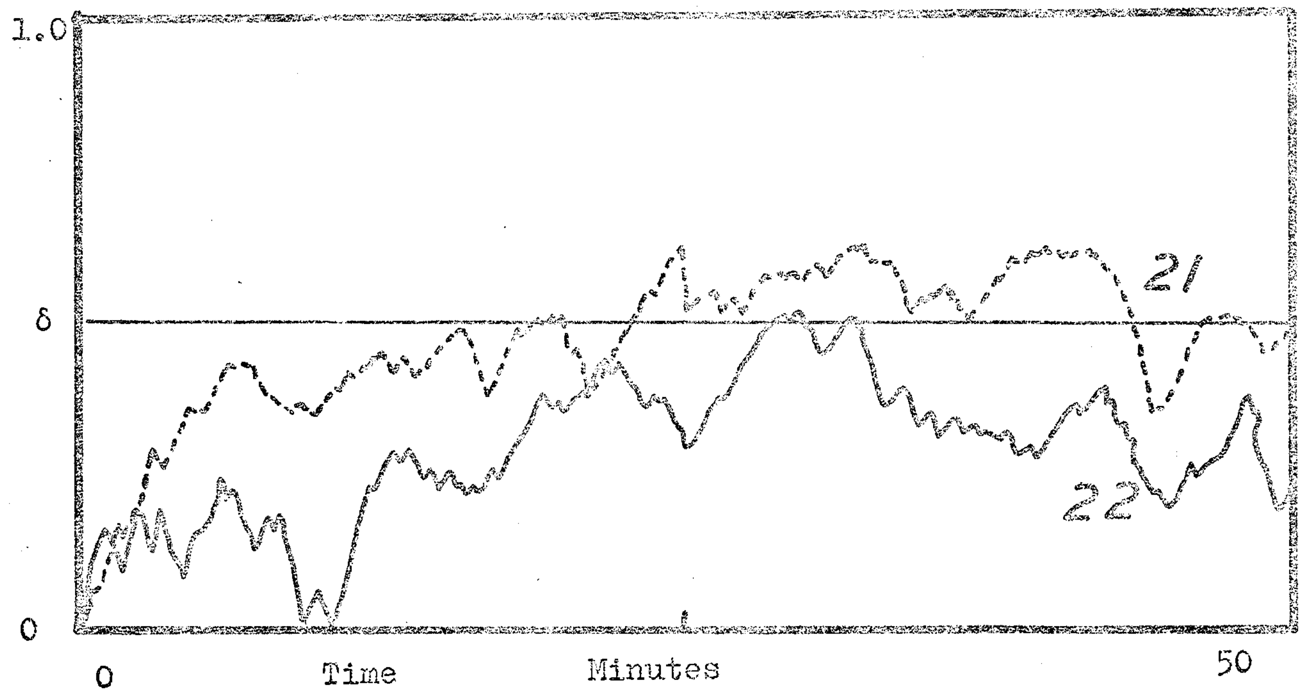


Figure A -1 Trajectories of Difficulty



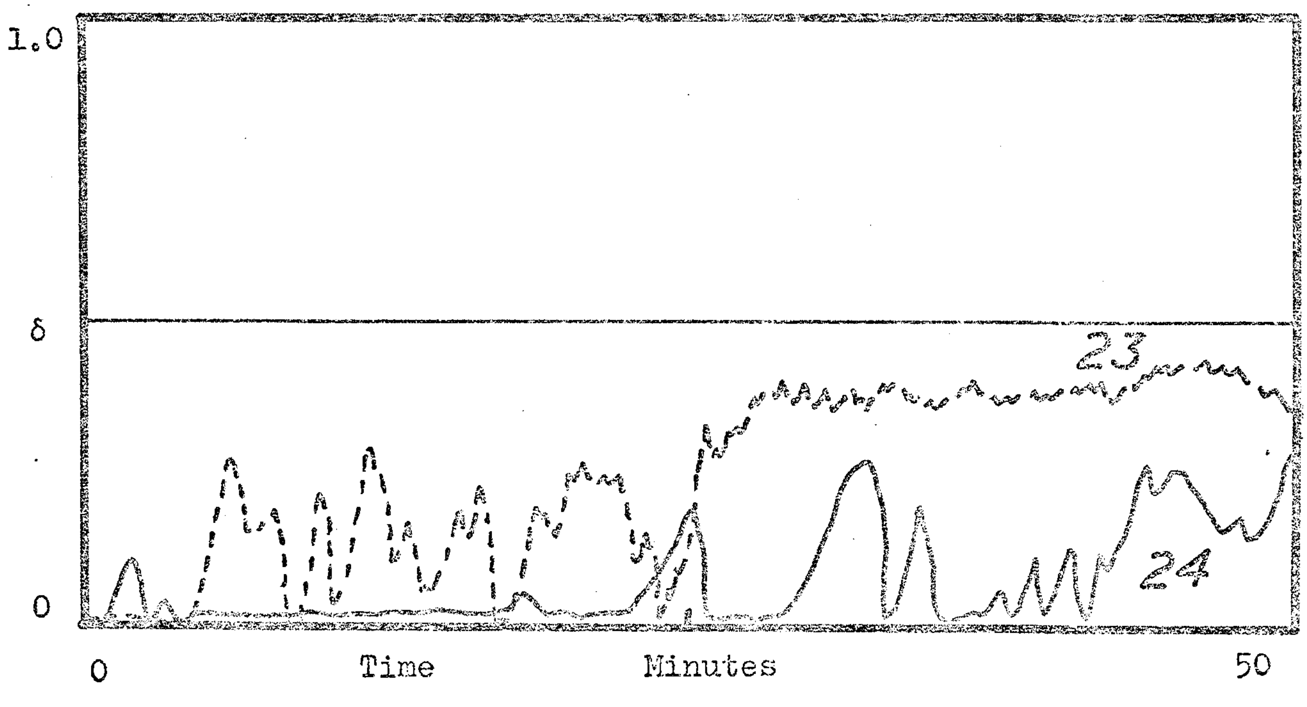
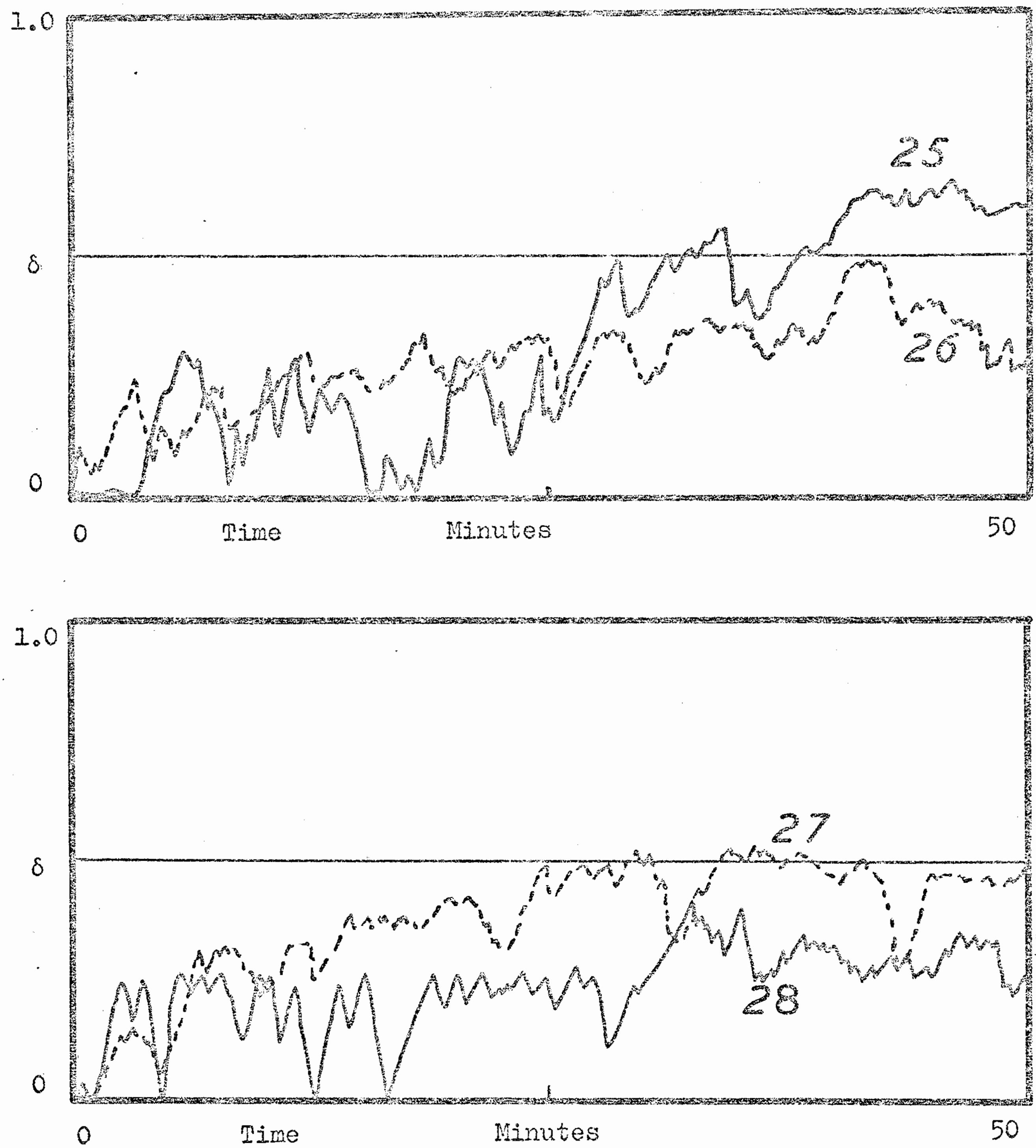
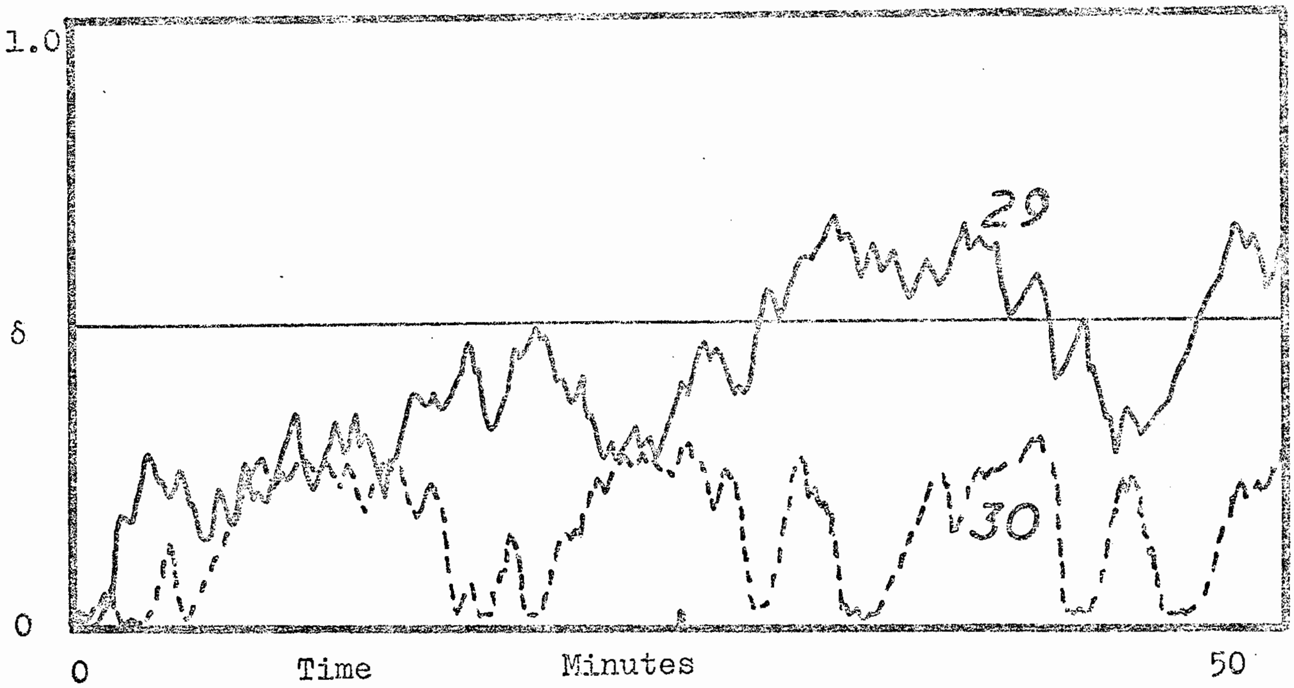


Figure A.-1 Trajectories of Difficulty



A2-1 Trajectories of Difficulty



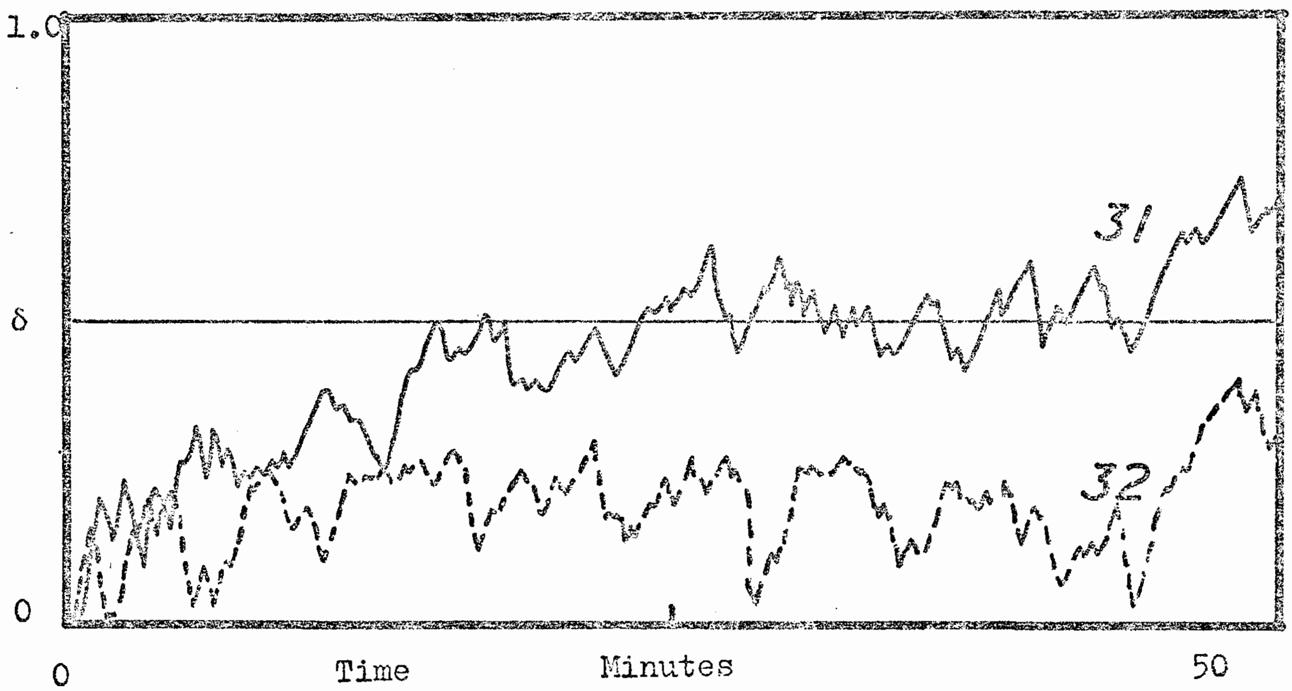


Figure A2-1 Trajectories of Difficulty