AUTOMATED FEEDBACK TRAINERS FOR PERCEPTUAL-MOTOR SKILLS

FINAL REPORT

Ministry of Defence Contract:

'Servomechanisms in Operator Training for Tracking Tasks'

By: B. R. Gaines,
Department of Experimental Psychology,
University of Cambridge.

September 1967

Contents

Acknowledgements

I am grateful to Cambridge University for permission to undertake the work described in this report and to Prof. O.L. Zangwill for providing space in his department. I am also grateful to Mr. P.E.K. Donaldson, Mr. R.L. Gregory and Mr. H.C.W. Stockbridge, for initiating this research study and for their early advice and encouragement. The overall objectives and approach to this study owe much to long discussions with Dr. J.H. Andreae and Dr. J.L. Gedye. I am grateful to Dr. E.M. Hudson and Dr. G. Pask for providing information on their own studies of feedback training.

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 of human learning over long periods could not have been taken without access to the controlled environment set up by Dr. J.L. Gedye for his own work.

Finally, I have a personal debt to the Ministry of Defence for their ready co-operation and advice in all matters affecting the administration of this contract.

Section 1: Objectives, Summary of Results and Implications

1.1 Title of Investigation

Servo Mechanisms in Operator Training for Tracking Tasks.

1.2 Description of Work

An investigation to be carried out into the effect of varying the parameters of the display variables and the transfer function of the control system on human operator learning rates in tracking tasks. In particular to investigate the effect of varying the parameters and transfer-function automatically, so that the subject's mean tracking error tends to be constant.

1.3 Objectives of the Research

The basic objectives of the research are summarized in the title and description of work in the agreement between the Ministry and University (Paragraphs 1.1 and 1.2 above). The problem considered is that of training human operators in control skills, such as manoeuvring an aircraft, missile or submarine. It is assumed that the system to be controlled can be simulated in such a way that the difficulty of control may be varied. The training technique to be investigated involves the automatic variation of this difficulty according to the operator's state of learning.

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, 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 the operator then he will learn more rapidly with training at a higher level of difficulty - whilst if the required task is very difficult for the operator then he will learn better 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 best training technique should, therefore, involve feedback from the operator's state of learning to the level of difficulty of the task. A device which automatically adjusts the task difficulty as a function of the operator's skill in order to maintain the optimum conditions for learning is called "an automated feedback trainer".

Hence, the basic questions to be answered by the research may be summarized:-

(1) Given a control skill which is easy for the naive operator, compare learning at this easy level with transfer from learning at a more difficult level to testing at this easy level, and, in particular, investigate the situation in which the more difficult level is not fixed but varies with the operator's state of learning.

(2) Given a control skill which is very difficult for the naive operator, compare learning at this difficult level with transfer from learning at an easier level to testing at this difficult level, and, in particular investigate the situation in which the easier level is not fixed but varies with the operator's state of learning.

1.4 Further Objectives

Preliminary studies of the literature and initial exploratory experiments suggested that the basic objectives outlined above required further additions and qualifications. The automatic adjustment of task difficulty introduces problems of its own, and previous studies had either been unsuccessful in implementing the necessary feedback or suggested that it introduced artifacts through its instability. Hence both theoretical and practical studies of the operation and stability of feedback training loops were undertaken.

The learning of control skills and training techniques for them lack a theoretical foundation, and there is no standard terminology for learning phenomena or training techniques. A basic study of adaption and learning in both human and automatic control devices was undertaken with the objective of providing a firm foundation for the synthesis of feedback trainers and for the analysis of the results of experiments on human learning and training.

Experiments on human learning have inherent difficulties because it is virtually impossible to match operators in their characteristics and the learning process is irreversible. Hence the effects of different training regimes may be determined only by statistical comparison of large samples of trainees. The number of variables affecting training are very large, however, and an investigation was undertaken of the possibility of using advanced automatic controllers in the form of 'learning machines' as simulated operators in the preliminary evaluation of possible training regimes.

Although verbalization plays little part in the ultimate performance of a control skill, it may have a profound effect during the early stages of learning. In any practical training system verbal instruction would be used for maximum effect, and in a laboratory investigation verbalization requires control if artifacts are not to be introduced in the comparison of training techniques. The basic objectives, therefore, were extended to include examination of the interaction of instructions with training techniques, and the evaluation of the trainee's attitude to, and knowledge of, the training situation.

A detailed discussion of the objectives in the context of previous studies is given in Section 2.1 of this report.

1.5 Overall Summary of Results

A study has been made of human learning of a tracking skill under various training regimes, including automatic feedback training. This has involved a major experiment with 72 operators from a homogeneous population comprising RAF pilots at an advanced stage of selection and training, together with a number of minor experiments on a variety of operators. A stable feedback training system has been developed which is free from artifacts, and this has been investigated both theoretically and experimentally. Simulated 'learning machines' have been used as dummy

/dummy
operators under the same training conditions as those for human subjects. The interaction of information contained in instructions with the effect of different training techniques has been investigated both for human operators and learning machines. An extensive theory has been developed which accounts for many of the phenomena of human learning of control skills and the effects of different training techniques.

In the main experiments operators were trained under one of three conditions: high difficulty; low difficulty; or feedback, in which the level of difficulty was adjusted automatically to maintain their mean error constant. In each group, half the operators were given strong, or informative, instructions which explained to them the nature of the controls, whilst the others were given weak, or non-informative, instructions which told them only what they were required to achieve.

All the operators were tested (mean error measured) after training, first at the high level of difficulty and then at the low level. The test at the high level was in two stages, one incorporated at the end of the final training session without informing the operator, the other directly afterwards when the operator was warned of the test; this enabled the effect of instruction-induced stress to be measured. The operators also filled in questionnaires which enabled their stated interest and evaluation of the training situation to be measured, and which were open-ended as far as comments were concerned so that an estimate of the operator's verbalization could be obtained.

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 perform badly on all 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% of the feedback grouped reached it, or a much higher level, during their training, and hence could perform well at the high level.

(b) The operators trained at the low level of difficulty split clearly according to the instructions given - those with weak instructions do not show appreciably better performance than those trained at a high level, 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 in their performance from the group who 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 informative versus non-informative instructions is that strong instructions give significantly improved performance in all groups.

(e) The effect of instruction-induced stress is that operators trained at a high level of difficulty get worse, operators trained at a low level do not vary appreciably, whereas operators trained under feedback conditions get significantly better; there is no interaction with weak and strong instructions. This is the only difference in performance which differentiates the group trained at a low level with strong instructions from those trained under feedback conditions with weak instructions.

(f) The questionnaires show no appreciable difference in the interest expressed by the various groups. There is a marked difference in the variance of the estimates of the difficulty of the control task. The grouped trained at a high level with strong instructions show significantly more verbalization than the other groups.

(g) In terms of 'transfer of training' the results demonstrate clearly that the question as to whether easy-difficult or difficult-easy transfer is best is not meaningful. They indicate, however, that transfer from a difficult level to an easy one is better than learning solely on the easy one, provided that the difficult level is within the operator's ability to perform reasonably well.

(h) Since the operator's skill varies with learning, this also indicates that the optimum level of difficulty must be selected according to the operator's ability. The automatic feedback training loop used in these experiments is shown to be effective by the results, since, under the same conditions of instruction, the feedback group performed significantly better than the others under all test conditions.

(i) Informative instructions are shown by the results to be capable of compensating for poor training conditions, provided that these are not too poor. However, even though the feedback group with weak instructions could not be distinguished from the low level group with strong instructions on performance alone, they showed significantly less deterioration of performance under instruction-induced stress.

(j) The results with computer-simulated learning machines parallel those with human operators completely. This demonstrates not only the utility of these devices as dummy operators but also the generality of the learning and training phenomena investigated.

(k) The theoretical studies indicate that the effectiveness of feedback training in these experiments may be attributed to the interaction between learning to use the controls and learning to control the system dynamics. Training any task that involves learning at least two skills, of which one must be performed satisfactorily to provide a normal environment for the other, and vice versa, is liable to benefit from feedback.

1.6 Practical Implications and Suggestions for Further Research

(1) 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 on one creates a difficult situation in which to learn another.

In the laboratory this situation was established by giving the operators unusual controls and high-order dynamics in a one-dimensional tracking task. In practice the situation is more likely to be created by the control of a multi-dimensional system 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 learned and there are interactions between performance on one and learning on another. It is less likely to be of value in single-dimensional tasks, although they may be difficult, eg. high-order tracking in one axis; in multi-dimensional tasks where there is little interaction, eg. three-dimensional tracking with compatible controls/displays and no interactions; in multi-dimensional tasks where there is strong interaction but little opportunity for learning one of the interfering tasks, eg. tracking with the displayed signal immersed in random noise.

It is suggested 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 control/displays but differing dynamics in the axes and strong cross-coupling between them.

(2) Feedback training, in general, should also be of value where the interacting skills are not continuous, but a different form of training system will then be required. For example, in missile control the nature of the skill changes with the position of the missile along its trajectory, and the skill in attaining a correct position and orientation during early flight interacts strongly with the skill in locating the target during later flight. Equally, the trajectory in early flight cannot be evaluated by the naive operator, except with reference to verbal instructions, without awareness of the requirements in the final phase of flight. Hence, feedback techniques should be of value, but because of the discontinuous nature of the skill some memory and decision-making capability is required in the feedback trainer. This would best be provided by using a hybrid analog/digital computer rather than the conventional analog computer of the present studies. Similarly, skills in which discreet actions are required, such as the setting of switches and the sampling of displays, are also amenable to feedback training using a hybrid configuration.

It is suggested that further studies encompass skills with both continuous and discreet components, and that the feedback training loop of future systems consist of a small digital computer; normal analog and logical techniques being used in the main simulator.

(3) The results of the present study emphasize the importance of verbal instruction in the teaching of perceptual-motor skills and the importance of evaluating and controlling verbal effects in laboratory studies of skill learning.

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 subject the 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, eg. if the level of difficulty in the feedback training loop does not rise to a criterion level after a certain time then the instructions would be repeated in an alternative form. The same system could be used to administer the questionnaires and evaluate the operator's response to, and knowledge of, the training situation.

It is suggested that further studies incorporate a programmable audio/visual display (film-strip/audio-tape system) in the feedback trainer in order to control fully the verbal instructions to the operator, and evaluate his verbal behaviour.

(4) The theoretical studies, together with the experimental results using human operators and learning machines, have demonstrated that it is possible to establish a general theory of adaption and learning which may be used as a basis for synthesizing feedback trainers and also leads to predictive models of human learning of control skills.

It is suggested that further studies should extend the theoretical basis for controlling and analysing human learning behaviour, and that some particular control system should be investigated in detail for the effects it induces in the state of adaption of both human operators and learning machines.

(5) 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 ideal for the evaluation of feedback training, but unsuitable for the validation of the trainer as a test.

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

Section 2: Previous Work Relevant to Automated Feedback Training

Documented research on automated feedback trainers for tracking skills has been very slight, comprising one major theoretical study, one major experimental study, and a number of minor apparatus studies. The earliest mention of the possibility appears to be that of Stockbridge and Siddall (1956), who suggested the use of a guided weapons tracking trainer in which 'the difficulty of the task is proportional to the success of the subject'. In a number of papers Pask (1960, 1961, 1964, 1965) has made available a very deep and comprehensive discussion of automated training, and has placed it in the general context of interactions between self-organizing systems. The only major experiment reported to date is that of Hudson (1964) who trained over seventy subjects for ten hours each on a third-order, two-dimensional tracking task.

A short exposition by Senders (1961) of the principles of "adaptive teaching machines" is a good example of the many studies of feedback training which have been reported informally. Hudson's work was supported by the U. S. Navy and several workers associated with the U. S. Naval Training Devices Centre have proposed feedback training systems: Ziegler, Birmingham and Chernikoff (1962 have described a 'teaching machine for the selection and training of operators of higher order vehicles' which removes 'quickening' as the operator's mean error modules decreases; Chernikoff's (1962) report on this machine and the ensuing discussion are particularly interesting. Bowen, Hale and Kelly (1962) have described a "general vehicular research tool"; Briggs (1962) has described experiments on scheduling augmented feedback according to the operator's performance which might be automated in a feedback training.

Since the level of difficulty of the control task is varied according to the operator's ability, a feedback trainer may be used to test this ability and measure it in terms of highest tolerable task difficulty. Jex, McDonnell and Phatak (1966) have carried out a comprehensive program of research for NASA on the use of control systems with varying dynamics to measure some parameters of the human operator describing function. This is the only published work which discusses the viability of different feedback loops from error functionals to task dynamics.

Feedback training may be regarded as a sophisticated technique for inducing and utilizing 'transfer' of training from one task to another. It takes a continuum of tasks which includes the required one and utilizes performance feedback to cause the operator to progress through these tasks at a uniform rate in terms of his own ability to transfer. Conventional 'transfer' techniques, in which the operator is trained in one environment and tested in another, may be seen as step transitions in difficulty, made 'open loop' without regard to the operator's individual ability. There is a vast literature on transfer but the relevant parts of it are summarized in an excellent paper by Holding (1962).

Although performance on a tracking task in terms of error functionals provides a coarse criterion of the operator's ability which may prove adequate for the operation of practical feedback trainers, one would hope to use rather finer estimates of the operator's actual control strategy for research purposes. Until recently, virtually all work on human strategies in manual control was in terms of the linear describing function, a device which is useful in defining the stability boundaries of high performance aircraft but has little relevance to the operator's actual strategy and its variation during learning. Historically this work is well reviewed by Licklider (1960) and a more recent report by McRuer, Graham, Krendel and Reisener (1965) covers the latest work in this field. DeLessio and Palin (1961) have proposed parameters in the describing function to account for learning, but this has not been followed up experimentally.

During the last five years there has been a great upsurge in research upon nonlinear models of the human operator, both in terms of optimal switching mode control of high-order systems and in terms of adaptive parameter changes in responses to changes in the controlled element. Much of this work, together with that on multivariable control strategies, has been associated with NASA contracts handled by Bolt, Beranek and Newman. The impetus for this work arises through a advances in control theory in which the analysis of data-sampling, optimal switching, and adaptive, control systems has attained a level comparable to that of linear systems.

Definitive experiments on data sampling models of the human operator in manual control have come out of work at Imperial College by Wilde and Wescott (1962) and more recently by Lange (1965). Switching mode models now proliferate (Li (1965), Bekey (1965)) and have become accepted as the best approximation to the human control strategy with high-order systems. Since the strategies involved give the fastest response for a force-limited controller this is a reasonable conslusion. The models of human adaption to step changes in the parameters of the controlled system have been largely in terms of describing functions, (Young (1964)), but recently variation in the parameters of a switching-mode model have been considered (Weir (1967)). A rambling review of much of this work, including studies of the muscle servos, has been prepared by Young and Stark (1965).

None of the work on human operator dynamics leads to a model of human learning which will predict the variation in the control strategy and performance of a naive operator coupled to a control system of known state and dynamics. In particular they do not predict under what circumstances the operator will learn to perform a control task and under what circumstances he will not. Failure in learning, 'psychological fatigue', the effects of instructions, are not taken into account, but these are vital to establishing an optimum training regime. Fortunately, again during recent years, there has been an increasing effort devoted to problems of adaptive control and machine learning, and much experience has been gained which is relevant to human learning training (Andreae (1967))

A very much neglected topic is the influence of verbal instruction and operator verbalization upon the learning of perceptual-motor skills. Apart from the outstanding work of Luria (1961) in developmental studies of the influence of language, and the theoretical structures involving a hierarchy of 'languages' proposed by Pask, there is virtually no literature on this topic. One has to turn to studies of psychopathology, particularly aphasia, or to philosophical discourses, to find even vague models of the interaction between cognitive and perceptual-motor activities (Critchley (1953)).

2.1 Summary of Previous Work in Relationship to the Present Study

This summary is intended to be fairly parochial, bringing out those aspects of previous work relevant to the present study even though they may not be the major consideration of the authors. This gives rise to five requirements which determined the practical objectives of this study and were largely satisfied by it.

2.11 Stability of Feedback Trainers

There have been no studies of the realization of stable automatic feedback training loops.

Although Jex has mechanized similar feedback loops to those of the present study, he has done so for purpose of testing, not training, and has found them unsuccessful. The reasons for this failure, compared with their proven utility in the present study, may be found in a theoretical analysis of feedback training loop dynamics with various operator strategies and controlled systems (Section 4.4).

Hudson's feedback training technique was not successful in operations because it related the mean error to the absolute dynamics of the controlled element. He suggests that a loop be mechanized instead to 'keep the error level at some desired point', and this is what has been done in the present study. This may be roughly paraphrased as 'maintaining the difficulty constant for the operator', which may by taken as a basic concept in feedback training.

The abvious empirical test of the dynamics and stability of a feedback training system is to use simulated operators, automatic controllers, of different but non-time-varying 'ability', as experimental subjects. This does not appear to have been done in any previous work on feedback training, although Taylor and Birmingham (1959) have plotted the 'learning' curve for a variable gain element on various control systems in order to deride too close an analysis of the shape of learning curves.

Requirement (a) The stability of feedback training loops should be investigated both theoretically by analysis of the loop dynamics, and practically by experiments with simulated operators.

2.12 Utility of Feedback Trainers

Only one previous study has reported any advantage to be gained from feedback training.

Although Hudson's automatic adjustment of controlled element dynamics was not successful, he did experiment with manual control of the dynamics in order to maintain the operator's performance at a fixed level. He compares operators trained under this manually effected feedback training regime at two levels of difficulty with operators who train on the required system throughout (a third-order transfer function). His results show quite clearly that there is an optimum level of difficulty for the operator, defined in terms of his performance, which induces the most rapid learning. Training at levels either below or above this gives a slower rate of learning. One feature of particular interest in Hudson's experiments is that he used a variety of plant parameters in order to maintain the difficulty constant for the subject, but these variations seemed to have little effect on the overall result.

The lack of advantage suggested by Chernikoff may be ascribed to the same defect which caused Hudson to abandon his automatic feedback trainer experiments. The discussion at the end of Chernikoff's paper is illuminating in showing the attitudes adopted towards the feedback trainer. It is interesting to note that the other studies at Dunlap and Associates and the Naval Training Devices Center have not led to the publishing of any definitive results during the intervening five years. Poulton's comment to the effect that there is no guarantee that an

/that an
arbitrary feedback trainer will be the optimum one is probably the most relevant. Hudson, in evaluating a number of ways of varying the tasks dynamics in order to maintain performance at one of two levels, has, in fact, taken account of this.

Requirement (b) An automated feedback trainer, shown to be stable and free from artifacts, should be evaluated in an experimental situation such that a clear distinction is possible between feedback training, fixed training (training on criterion system) and open-loop training (transfer from training on easier or more difficult tasks to testing on criterion). Ideally there should be no learning in fixed training on the criterion task, but high performance on it after feedback training. Such an experimental result would provide a basis for further studies, applications and theoretical developments.

2.13 Theoretical Basis for Feedback Training

Theoretical studies of adaption and training are inadequate for the design and analysis of feedback trainers.

Despite considerable effort by both biologists and control engineers there is no generally accepted definition of adaptive structures and adaptive behaviour, let alone any theoretical basis for the analysis of human learning and the development of optimal training techniques. Zadeh has proposed a definition of adaptivity which breaks new ground in the importance of behavioural rather than structural aspects of adaption in providing explicata for the concept of 'adaption'. What is lacking is the development of this form of definition into a comprehensive treatment of the full range of phenomena of adaptive behaviour. In so far as an adaptive system may be regarded as a hierarchial structure, the phenomena of learning may be ascribed to the dynamics, and particularly the stability, of the higher levels, but unfortunately the theory of the stability of nonlinear systems is also very little developed.

A theory of adaptive behaviour is necessary to the analysis of the phenomena of learning, but it is not adequate for the synthesis of optimal training algorithms. This requires some account of the problems involved in learning to control environments with different characteristics. It may also involve analysis of the structure of the learning system, but this is not necessarily so, and one essential feature of the concept of 'training' is insensitivity to the nature of the learning system.

There are two sources of guidance on the design of automated feedback trainers, and especially those to satisfy Requirement (b) of this study. Pask, in treating teaching as a control problem, has developed a theory of adaption and training which emphasizes the importance of interaction between sub-skills in causing difficulty in learning which may be overcome by feedback training. Previous studies on feedback trainers for perceptual-motor skills have generally used controlled elements with the same high-order dynamics in both axes and no interaction between them. This involves the minimum establishment of interacting sub-skills, and, from Pask's account, is not a situation in which feedback training would be necessarily advantageous.

The other source of information about the problems of learning comes from recent work on the design of intelligent artifacts. In attempting to synthesize a controller which will learn to perform well in a new environment, and adapt itself to changes in that environment, it soon becomes obvious that there are problems in the act of learning itself, basic epistemological problems independent of the learning system and confronting any controller, man or artifact, which attempts to gain information about its environment in order to improve its control over it. These problems arise because the initial control strategy of a naive controller may create a situation in the environment entirely different from that which would arise of a strategy giving good performance were implemented. The learning which takes place in this initial situation may be irrelevant, or even deleterious, to the establishing of a useful control strategy. The interaction between sub-skills, instanced by Pask, is one example of this phenomenom - poor performance on one task may create an environment for another which makes learning impossible.

Requirement (c) A theory of adaptive behaviour should be developed to provide a basis for the analysis of adaption and training. A theory of learning should be developed which clarifies the problems involved in learning about an environment whilst controlling it, and enables training algorithms to be established for particular environments.

2.14 Use of Learning Machines to Evaluate Feedback Trainers

The effects of various training procedures on human learning are difficult to evaluate because of individual differences and vagaries.

Experiments on human learning have inherent difficulties over and above those of other psychological studies. Gross individual differences, and the lack of techniques to measure them, make it impossible to set up matched groups of subjects for comparison of different training regimes. Since learning is essentially an irreversible process it is impossible to use an operator as his own control and compare his rate of learning under different conditions. In practice it is necessary to use large numbers of subjects with as uniform a past history as possible and compare the overall behaviour of groups of operators under different regimes.

The availability of computer programs simulating intelligent behaviour makes it possible to overcome some of the problems of evaluating training techniques with human operators by utilizing these 'learning machines' as simulated trainees. The machines are precisely replicable and individually matched machines may be compared in their rate of learning under different training regimes. The information obtained in this way would be valueless if the structure of the machines, which probably differs greatly from that of the human operator, were to affect the relative outcomes under different training machines to a large extent. However, as noted in the previous section, there are basic problems in learning which are independent of the learning system, and these may be expected to affect men and machines in a similar fashion. Hunt, Marin and Stone (1966) have compared the behaviour of men and machines in learning cognitive skills, but this has not been done for perceptual-motor skills.

Requirement (d) An appropriate computer-simulated learning machine should be developed and any experiments comparing the effects of different training regimes on learning by human operators should be replicated using the learning machine as a subject.

2.15 Control of Verbalization in Feedback Training

The influence of verbal instruction and verbalization by the operator on the learning of perceptual-motor skills has been neglected.

Although eminent psychologists (Bartlett (1958), Fitts (1964)) have long pointed to the similarities between linguistic and perceptual-motor skills, the emphasis has been on analogies between them rather than interactions between verbally correlated cognitive behaviour and supposedly non-verbally organized perceptual-motor skills. There are very few studies of the operational use of language whereby patterns of movement, relationships between stimuli and responses, and evaluative criteria over and above those inherent in the objective of an action, can be directly programmed into the human operator by natural language communication, rather than through a process of learning by experience. Luria has published some very interesting studies of the development in children of verbal control over actions. Pask has highlighted the hierarchical structure of communication structures between self-organizing systems, and pointed to the problems which may arise when a feedback trainer 'communicates' only through the immediate environment in which a skill is to be learnt.

Although some recent work has demonstrated the effects of instructions on tracking skills (Miller (1965), Weir & Phatak (1967), the evidence for a strong verbal content in the early stages of learning a perceptual-motor skill is largely informal and anecdotal. Operators learning a new skill do comment, often quite extensively, on their problems and techniques to overcome them. At a later stage of learning they report that the best way to perform a demanding skill is not to think but to relax and let the perceptual-motor system get on with it. A good example of the possible importance of directed verbal interaction is in driving instruction. During the early stages the trainee will be told to handle the clutch gently and slowly to avoid stalling. At a later stage he will reach a point where he is letting in the clutch so slowly that the engine races and he will be told to let in the clutch as quickly as possible.

There is not adequate experimental evidence, either positive or negative, of the degree to which verbal instruction affects the learning of perceptual-motor skills. If there is a profound effect then it is important that automated feedback trainers either incorporate appropriate instruction or may be shown to obviate the need for it.

Requirement (e) That the interaction of verbal instruction with feedback training be investigated.

Section 3: The Theoretical Basis for Feedback Training

3.0 Summary of Learning and Training Theory

The theoretical studies demonstrate that a unified and universal theory of adaption and learning behaviour is possible, independent of the structure of the learning device or the nature of the task which it is learning. The nature of learning as a change with time is encompassed by postulating that the interaction between the learning device and its environment is divided into tasks. The purposefulness of learning is taken into account by postulating that the performance of a task is satisfactory or unsatisfactory. These are the fundamental, undefined, elements of a learning situation, and the theory of learning behaviour, and hence of training can be built on these.

A state-description of the learning device, based on the satisfactoriness of its performance of various tasks, leads to the definition of its adaption-automaton, whose inputs are tasks and whose output is the satisfactoriness of the controller in a given state performing a given task. The behaviour of this automaton for various forms of input sequence defines a variety of possible modes of adaption. In particular, it leads to the definition of training procedures, in which initial task-sequences are given to the learning device in order to improve its inherent adaptability.

Three modes of training are distinguished:-

(i) Fixed training - in which the trainee is immediately presented with the final task for learning.

(ii) Open-loop training - in which some preliminary sequence of tasks is given to the trainee.

(iii) Feedback training - in which the sequence of tasks given to the trainee is varied according to information about the state of his adaption-automaton.

The theory of learning behaviour does not suggest, in itself, the means for implementing open-loop or feedback training - except by experimentally determining the structure of a controller's adaption-automaton. A study of the basic epistemological problems involved in attempting to control an environment whilst at the same time learning about it in order to improve the control policy, the dual control problem, does, however, indicate the causes of failure in learning.

A given control policy restricts the states and state-transitions of the environment to some sub-set of the total possible behaviour, a sub-environment. The desired sub-environment, generated by a satisfactory control policy, may be entirely different from that generated by the initial policy of a naive controller, and the adaption which takes place in it may be irrelevant or even deleterious. The aim of the training system should thus be to maintain the desired sub-environment regardless of the control policy of the trainee.

The additional control loop around the environment, necessary to maintain the desired sub-environment, may be said to be closed by a training controller, the selection of which may be ascribed to a trainer having access to information about the state of the trainee. If a channel of direct communication, such as one based on natural language, exists between trainer and trainee, then it may be used to prime the trainee with an initial control policy or adaptive strategy. The possibilities and limitations of direct communication in teaching perceptual-motor skills have not yet been investigated in any depth, but may be expected to be of great importance in any practical training system.

These theoretical considerations have been applied to the design of a feedback trainer for human operators and adaptive controllers learning a novel tracking skill. The environment chosen is a third-order transfer function, consisting of an integrator following a stable second-order term of variable natural-frequency and damping ratio. The inputs to this are a disturbance of variable amplitude and impulses from two push-buttons acting as manual controls. The output is displayed on a cathode-ray tube, and the operator's task is to regulate the system so that its output is zero. The push-buttons incorporate memory so that the operator not only has to learn to control a high-order system, but also to use the manual inputs correctly.

The desired sub-environment is a region about zero in the state-space of this control system, and within it the system behaves in a linear manner. A control policy which makes the overall loop unstable generates a sub-environment around the boundaries of the state-space, where the system is non-linear. Failure to maintain the desired sub-environment may be detected by measuring the mean error modulus, and the trainer bases its strategy upon this.

When the error modulus is above a given tolerance, the trainer decreases the difficulty of the control task - otherwise it increases the difficulty. This is a stable strategy for adjusting the training controller, provided certain reasonable conditions are met, and the difficulty is rapidly adjusted to an asymptotic value with both nonadaptive human operators and automatic controllers. The feedback trainer used in this way forms the basis for a very sensitive test of control ability.

3.1 Adaptive Control Theory

The three elements constituting an adaptive control situation are:-

an environment - which, in abstract, is a black box with inputs and outputs, and, in reality, may be a vehicle under manual control to be driven from one location to another - or, alternatively, a geometrical figure subject to the transformations of Euclidean geometry in which two angles are to be proved equal - or a person who must be persuaded, using natural language, that some proposition is true or false;

a controller, coupled to the environment for the purpose of regulating it to the desired end specified above - the human being is an obvious controller for each of the three tasks outlined here - automatic controllers are available which will perform the first task very well and the second task in a reasonable way (Newell (1959)), and some success has been reported recently in the third situation (Weizenbaum (1966)), although the level of conversation was rather mundane! ;

a performance measure, which indicates the extent to which the controller is achieving its purpose - this will often be a numerical measure or rank-ordering of results, but the minimum requirement is that it should be possible to establish whether or not the interaction between controller and environment is satisfactory (Zadeh (1963), Gaines (1967)).

The expected behaviour of an adaptive controller when coupled to an environment is that, if its control policy is not satisfactory for that environment, then it will eventually become so. Thus it must be possible to segment the interaction of the controller with its environment into at least two phases, in the first of which it is not satisfactory and in the last of which it has become so.

This segmentation, which is inherent in the basic concept of adaption, may be extended to form a description of the full range of possible adaptive behaviour. A task is defined to be a segment of the interaction between controller and environment for which it is possible to say whether or not the controller has performed satisfactorily. At the beginning of a task the controller will be in some state which causes it to implement a particular control policy. At the end of a task it will be in some other state and will implement a different control policy. If the controller is deterministic and the task is reasonably defined, then the final state of the controller will be uniquely determined by its initial state and the task given to it.

Hence the adaptive behaviour of a controller may be ascribed to an automaton whose inputs are tasks, whose states are controller states, and whose outputs are on one hand control policies, and on the other the satisfactoriness of the control policy for a given task. This is the adaption-automaton of the controller relative to the set of tasks, and training may be shown to be a problem of controlling this automaton.

3.2 Modes of Adaption

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. This is equivalent to the adaption-automaton being given a sequence of inputs consisting of the same task repeated indefinitely. An interaction between controller and environment consisting of the repetition of a single task is defined to be 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. Eventually, however, it must become satisfactory and remain so; an acceptable interaction is one which reaches a stable condition of satisfactoriness. In this stable condition the controller is said to be adapted to the task.

These definitions, based on the behaviour of the adaption-automaton, may be extended to account for all the tremendous variety of adaptive phenomena. Only two further modes of adaption, potential adaption and compatible adaption, will be considered here. A controller in such a state that it will have an acceptable interaction with any one of a set of tasks is defined to be potentially adaptive to that set of tasks. 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.

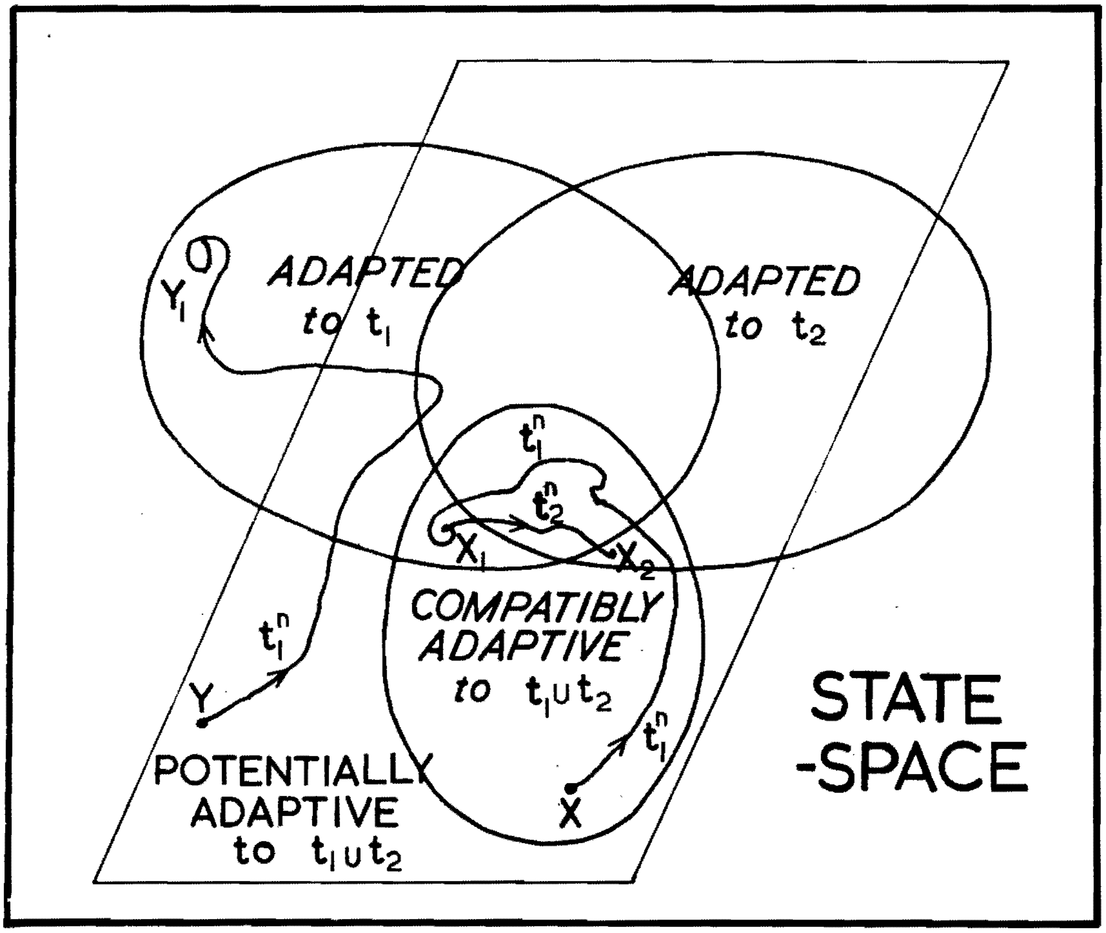

Figure 1 Adaption to a single task

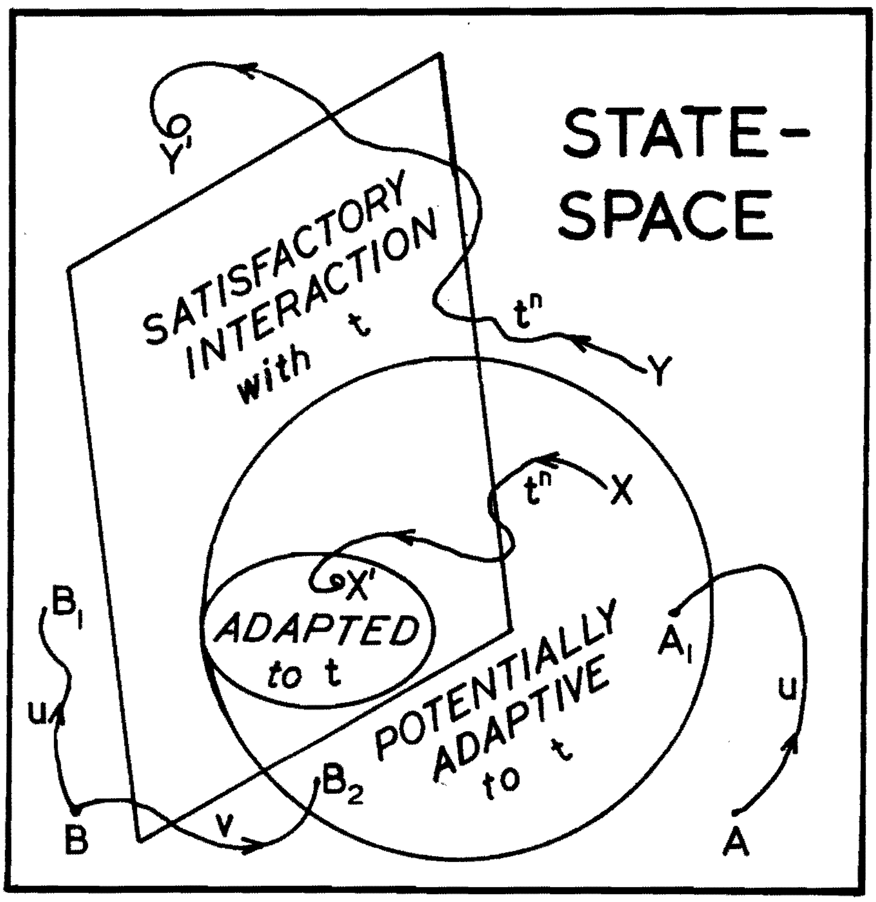

Figure 2 Adaption to a pair of tasks

Thus a controller which is potentially adaptive to a set of tasks is able to learn to perform any one of them satisfactorily. If it is also compatibly adaptive to the set, then having adapted to one, it is able to re-adjust its control policy to become adapted to another. These definitions are best illustrated by diagrams showing the state-trajectories of the adaption-automaton as the controller adapts.

The state-space of the automaton is shown as a rectangular region in Figure 1, and within it are delimited those states for which the controller is satisfactory given the task, t. The states for which the controller is adapted to t form a sub-set of these, since a trajectory starting in the adapted region must always remain satisfactory. The states for which the controller is potentially adaptive to t form another region enclosing the adapted one.

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 leave it;

started within the potentially adaptive region, at X, it will remain within that region, eventually entering the adapted region and never leaving it.

Compatible adaption involves a wider variety of possible behaviour. Figure 2 illustrates the adapted regions for two tasks, t_1 and t_2, and the potentially adaptive and compatibly adaptive regions for the pair of tasks. A trajectory generated by t_1 from Y, which is within the potentially adaptive region but outside the compatibly adaptive region, must eventually enter the adapted region for t_1, but in doing so it must also leave the potentially adaptive region for t_2.

Only trajectories from within the compatibly adaptive region are such as to always remain within the potentially adaptive region for both tasks. A trajectory from the point X, generated by repetition of t_1, arrives at X_2 where the controller is not only adapted to t_1 but also remains potentially adaptive to t_2. Hence the change from t_1 to t_2 at X_1 causes the trajectory to cross over into the adapted region for t_2. Note that the trajectory between X_1 and X_2 crosses a region in which the controller is adapted to both t_1 and t_2. If a sub-set of this region exists such that the controller is not only adapted to both tasks but also remains so when given any sequence of the two tasks, then the controller is defined to be <u>jointly adapted</u> to the tasks within the sub-region.

Discussion of adaptive behaviour in these abstract terms has the advantage of bringing together phenomena from human and animal psychology, automatic control and general systems theory, which go under different names, have different explanations, and yet are essentially examples of the same system relationships. In particular, the behavioural theory of adaption makes no distinction between the learning of a cognitive skill and the learning of a perceptual-motor skill. The phenomenom of <u>set</u> in problem-solving, for example, whereby a person given problems of type A or type B learns to solve them readily, but having learned to solve type A then finds it impossible to learn to solve type B is an example of potential but not compatible adaption.

In the experiments on control skills reported here, all the abstract phenomena described have been encountered. Learning may well be unstable, and a trajectory of the type shown in Figure 1 starting from Y, where the performance improves to a high level and then deteriorates again is not uncommon for both human operators and learning machines. Figures 10, 11 and 12, for example, show the learning curves for different operators and machines performing a control skill, and each gives an example of zero learning, unstable learning and stable learning. In the study of control strategies the first two examples would be regarded as artifacts or failures to be discarded from the experimental material. In the study of learning and training it is the situations in which each of these forms of learning curve arise which are of main interest.

3.3 Modes of Training

When the controller's initial state is outside the region of potential adaption to a task, 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 has <u>trained</u> it for the original task.

In Figure 3, for example, the point A is outside the region of potential adaption to the task, t, and repetition of this task will not lead to stable satisfactory performance. The sequence of tasks, u, however, gives rise to a trajectory which terminates within the region where the controller is potentially adaptive to t. If the training sequence, u, consisted of another task, t', repeated, then we would say that there had been transfer of learning from the task t' to the task t.

The training sequence, u, will not necessarily be suitable for all initial conditions of the controller; e.g. the trajectory induced by u from the point B in Figure 3 terminates at B_1 which is outside the region of potential adaption. Figure 3 illustrates the regions in which the training sequence u gives complete transfer (trajectories enter adapted region) and partial transfer (trajectories enter potentially adaptive region) to the task, t. Outside these regions, although u is not a suitable training

/training
sequence, it may be possible to find an alternative sequence, v, which again causes the controller to become potentially adaptive or adapted. To choose between u and v, however, it is necessary to have some information about the initial condition of the controller.

These considerations lead to the definition of three types of training:-

(i) Fixed training - to 'train' the controller for the task, t, it is given a sequence consisting of this task repeated a number of times. This involves no training effort and relies on the controller being potentially adaptive to the task, t.

(ii) Open-loop training - the trainer is given some training sequence, u, before proceeding to learn t. The training sequence is not adjusted for differences in controllers, but will obviously be chosen to maximize the region in state-space from which it gives partial or complete transfer.

(iii) Feedback training - the controller is given a sequence of tasks selected according to information about its state. Training has now become a problem of controlling the adaption-automaton by varying its input according to some output received from it.

Examples of the various types of training may be drawn from everyday experience - riding a bicycle is quite a difficult feat which benefits from training. In case (i) we would give the would-be rider a machine and let him attempt to ride it. If, instead, we set him off on a downward slope for the first few attempts, and then, regardless of his proficiency, expected him to learn to ride under normal conditions, we would be practising open-loop training, case (ii). If, further, we noted his performance or degree of confidence at each stage of learning, and selected the conditions accordingly, then we would have added an element of feedback to the training, case (iii).

3.4 Problems of Learning

Although adaptive control theory enables the rationale behind various forms of training to be defined and analysed, it does not itself indicate the techniques for setting up an open-loop training sequence or realizing a feedback trainer. Indeed it may appear that these are highly dependent on the structure of the specific controller to be trained, and cannot be considered in general. For example, the information flow from most natural environments is highly redundant, and different animals use different senses to identify, for instance, the route through a forest. Factors which increased the difficulty for an animal relying on its sense of smell would probably be irrelevant to one relying on its sense of sight.

These specific properties of the learning system and its environment are not the fundamental causes of difficulty in learning, however, and their importance lies in their influence on more basic and profound problems, inherent in the act of learning itself, which may be analysed in general. The study of adaptive artifacts, in particular, has lead to an understanding of the basic epistemological problems involved in attempting to control an environment whilst at the same time learning about it in order to improve the control strategy (the so-called <u>dual control problem</u>, Feldbaum (1963)).

The fundamental structure of all forms of adaptive controller is a two-level hierarchy in which the lower level implements one of a class of possible control policies, whilst the upper level selects the policy to be implemented, Figure 4. This definition emphasizes the relativity of the term <u>adaptive</u>, since any particular controller may be split up in many ways according to the definition of the class of possible control policies. Indeed, as control science progresses, we may expect the <u>adaptive controllers</u> of one era to become the <u>control policies</u> of the next; adaption, in engineering, is essentially a means of increasing the power of the present generation of controllers, whatever that may be.

The importance of this relativity, so far as training is concerned, is that, in training, we are attempting to control the upper level of the adaptive hierarchy. If, in synthesizing a trainer, we conceptually split the controller at too low a level, and develop training strategies for manipulating the control policies thus defined, we may find that the upper level of the adaptive controller has become a complex structure very difficult to control. For example, a simple learning machine might operate on the assumption that its environment has fixed dynamics, measure these and implement a control policy accordingly. A training sequence for such a machine would probsbly consist of some variation in the dynamics of its environment. The human operator, or a more complex learning machine, might well detect the variation in plant dynamics and adapt to follow the expected changes in dynamics.

Thus, the training system for something as complex as the human operator may well consist of a multi-loop controller, with each loop corresponding to a different split of the trainee into upper and lower levels; that is, the learning system has to be regarded as a multi-level hierarchy. In the case of the human controller, the most important upper level, which must not be neglected even in teaching perceptual-motor skills, is that with which verbal communication is possible. The implications of this are discussed in a later section, and the present discussion assumes that the two parts of the controller are well-defined.

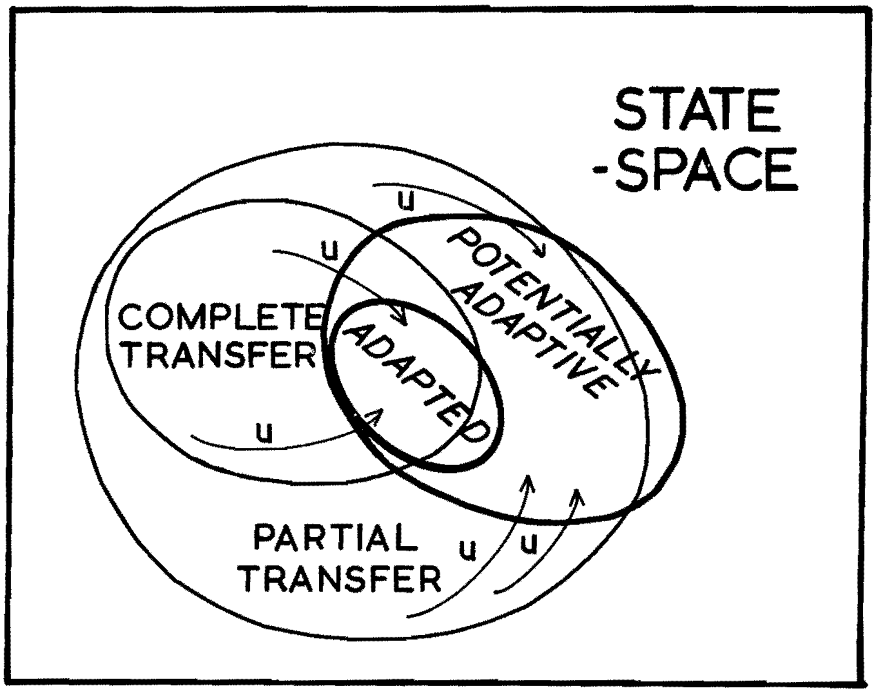

Figure 3 Open-loop training

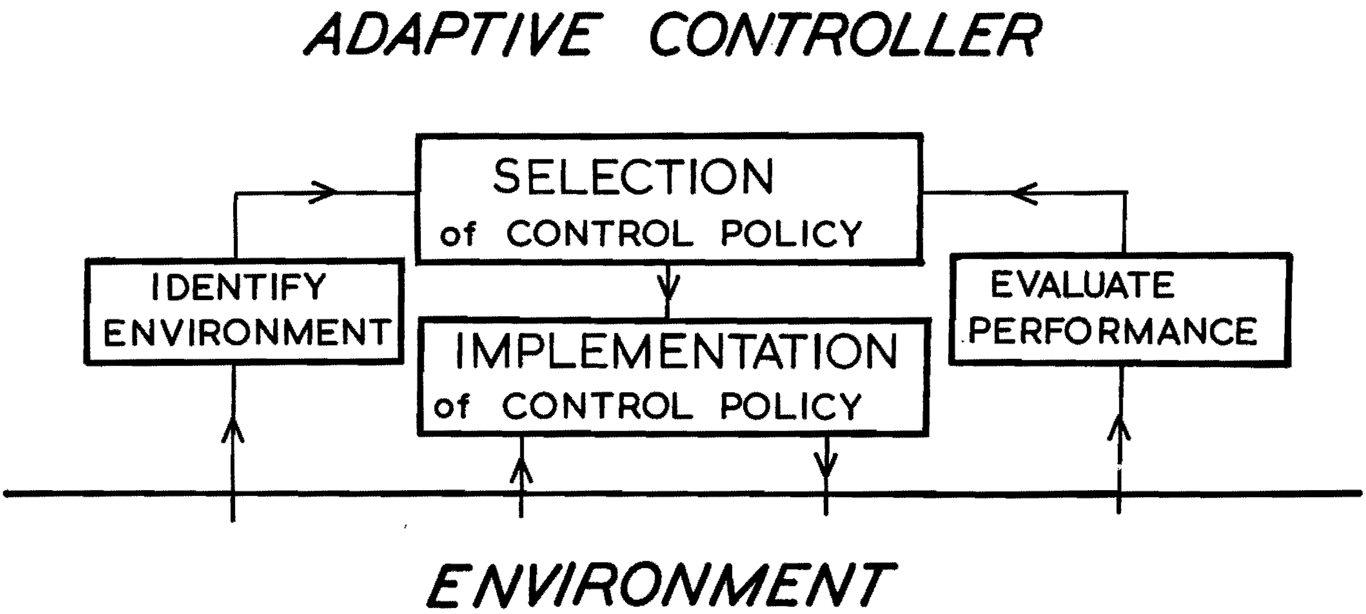

Figure 4 Adaptive control structure

To select a control policy appropriate to the environment and its goal in controlling it, the upper level requires information about the environment. There are two distinct classes of information relevant to the selection of a control policy: the nature of the environment itself - i.e. identification of the input/state/output relationship for the environment; and performance measures for various possible control policies operating upon the environment, Figure 4. Obviously either type of information forms the basis for selection of a control policy: if the controller is able to identify the environment completely and correctly, and has access to an optimum mapping of environments to policies, then it can implement the best policy available; equally, if the controller is able to establish the performance of every possible control policy for the environment, it can select the best policy 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 difficulties arise 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. 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 deleterious to performance in the desired sub-environment. Alternatively, and especially if the adaptive controller adopts a deliberate search policy, the initial sub-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. 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 would be correlated with these and could be tacitly deduced from them. If the initial control policy generates an abnormal, or unexpected, environment, then the measured parameters might carry no inference about other characteristics of the environment, and the selected control policy based on them would be invalid. If this policy also generated in abnormal sub-environment then mal-adaption would continue, so that a normal sub-environment and an optimal control policy were never 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, this is impossible, and controllers of this type assume that there is some topology on their control policies such that, given the evaluation of a number of policies, they are able to select a new policy near to the best and away from the worst - that is, they utilize an incremental modification of their policies. 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 controller will be trapped and never attain an optimum control policy, this difficulty is called one of false-peaks in hill-climbing systems.

3.5 Training Strategies

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-environment. This may be represented, Figure 5, as the addition of a training controller to the environment, but the strategy of this controller in maintaining the desired sub-environment is dependent on the particular form of environment. In 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 control might apply auxiliary feedback to maintain the overall control loop marginally stable.

Since the trainee in Figure 5 is assumed to be adaptive, the training controller need exert less and less control to maintain the desired sub-environment as time passes. Hence there may be a trainer which selects a suitable training controller either as a function of time (open-loop training, section 3.3) or according to information about the state of the trainee (feedback training, section 3.3). It will be noted that the structure of the training system is an exact image of the structure of the trainee, regarded as an adaptive controller. With the human controller, and with recent learning machines, direct verbal communication is possible with a higher level, and the trainer may have access to this channel for priming the trainee with control policies or adaptive strategies.

The nature of verbal communication is not sufficiently understood, especially in its effects on the learning of perceptual-motor skills, to allow a thorough analysis of the use of a direct channel between trainer and trainee. In the context of identification and performance evaluation (section 3.4), however, it is obvious that it may be possible for the trainer to pass information about the identity of the environment or the optimality of various control policies to the trainee, and hence eliminate the major source of difficulty in the dual control problem.

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 identity or optimality available, but must be able to communicate this in a form assimilable by the trainee. In practice the trainer's knowledge about the environment may be such as to make it possible to select a suitable training controller, but not allow any useful advice to be given to the trainee. The optimum trainer may be expected to combine verbal instruction with feedback training, and the experiments reported here have investigated the interaction between these two techniques.

The discussion to this point has been maintained at a high level of abstraction to demonstrate the possibility of treating adaption and learning as universal phenomena, and developing a theory of training independent of the skill for which training is required. It remains to be demonstrated that this theory can be applied in practical situations, that the variety of possible learning behaviours do manifest themselves in reality, and that feedback trainers of value are feasible within the bounds of present technology. The following sections describe the control situation chosen for an investigation of feedback training, the strategies used in the training controller and trainer, and the results of some experiments on training human operators and adaptive controllers using this system.

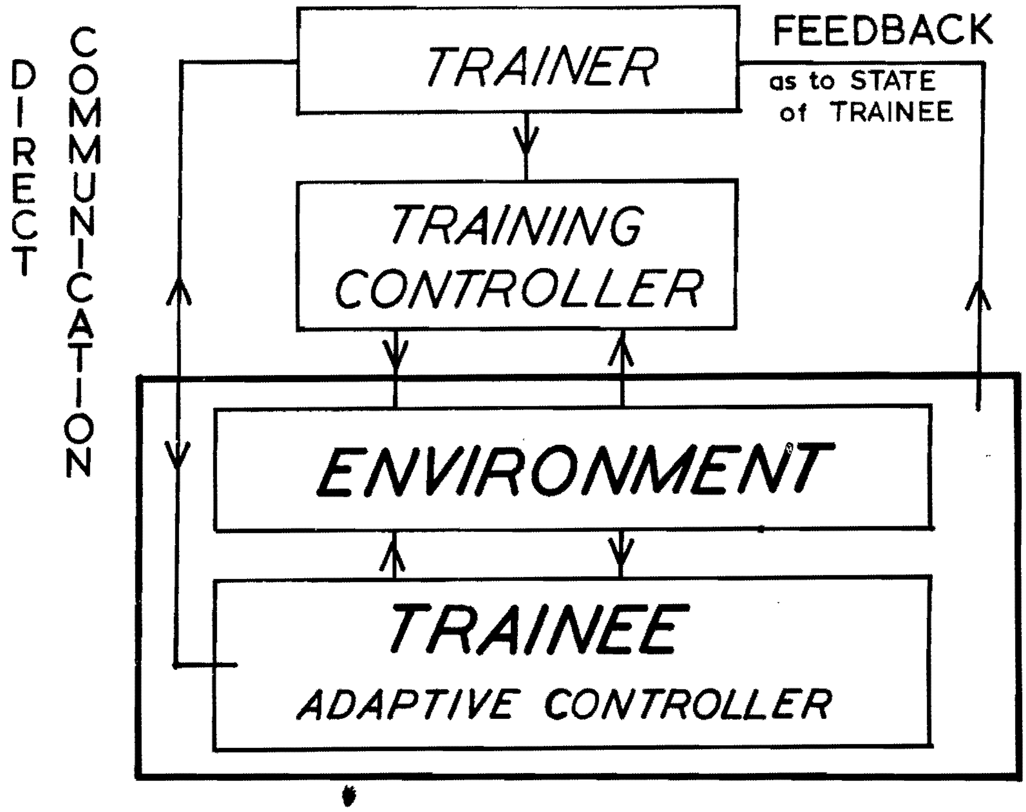

Figure 5   Feedback training system

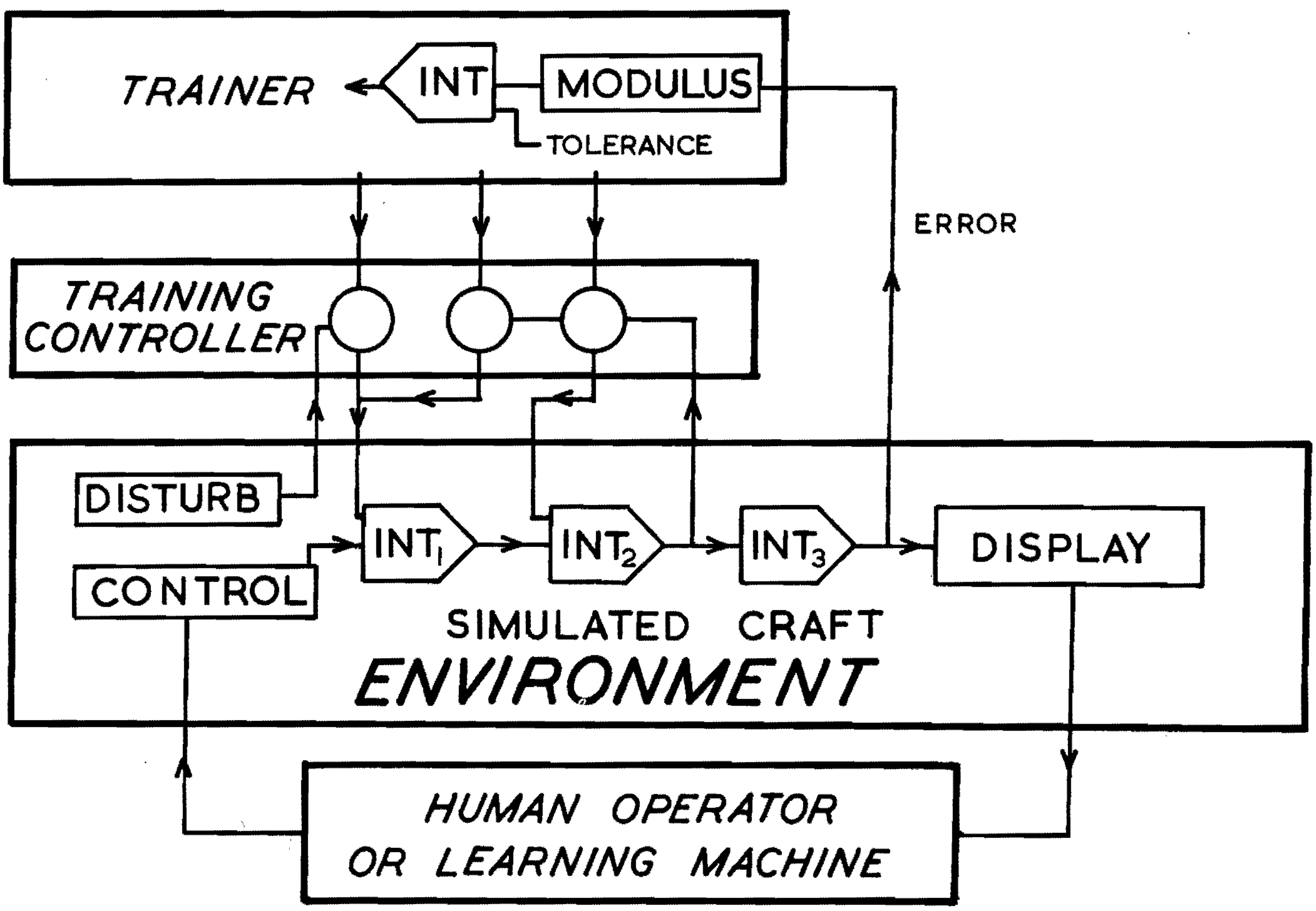

Figure 6   Feedback trainer for a control skill

Section 4: An Automated Feedback Trainer for a Tracking Skill

4.1 Choice of control task

In choosing a control situation in which to investigate the learning of a perceptual-motor skill many factors were taken into account. It was required that the task be related to practical situations in which training was already employed, and the regulation of high-order dynamics, such as those of the longitudinal motion of an aircraft (Blakelock (1965)) or submarine, was selected as being both realistic and of fundamental interest in manual and automatic control.

Preliminary experiments and comparison with aircraft dynamics indicated that a second-order, stable transfer-function, with undamped natural frequency in the range between 0 and .8Hz, and a damping ratio in the range between 0 and 1, was most suitable. However, the human operator is capable of compensating such a system fairly easily, and the dynamics were increased to third-order by addition of a rate control. The overall transfer-function was thus of the form:-

$$1/s(a^2s^2 + 2kabs + b^2),$$

which, by variation of k and a, may be swept from virtually first-order to pure third-order in a variety of trajectories through the natural-frequency/damping-ratio plane. Variation of k and a thus constitutes a means of changing the degree of compensation required, and hence the difficulty of the task for the human operator.

The operator was provided with an input to the above transfer-function by means of a manual control, and a second disturbing input was provided within the system. The error in maintaining the output of the transfer-function at zero was shown to the operator on a cathode-ray tube display.

This control problem is similar to that used in some previous studies of human perceptual-motor skills, but, as commonly used, it suffers from two major defects which make it difficult to obtain meaningful results in experiments on learning. The first problem is operator fatigue which may affect tracking performance after intervals of as little as 90 sec. This is a minor nuisance in studies of control policies since short tracking runs have to be used, but in the study of learning it causes artifacts and difficulties in experimental control which are virtually insuperable (artifacts, that is, when we are not trying to investigate the phenomenom of fatigue and its effects on learning).

Fortunately, preliminary experiments had shown that the use of discrete push-button controls, rather than a conventional joy-stick, greatly reduced operator fatigue so that runs of 15 to 20 minutes became acceptable. This effect is interesting in its own right, since the use of discrete controls not only improved tracking performance, as would be expected in theory and has previously been found in practice, but also reduced fatigue at the same level of performance (that is, when the task difficulty was increased until the performance had deteriorated to its previous level). Push-button controls were used, therefore, in the experiments on training, whilst a conventional joystick was used in some experiments on the application of the same apparatus to testing.

The second criticism of the basic control system, even with discrete controls, is that there is an unrealistic emphasis on the particular task of compensatory tracking. In most real-life situations where manual tracking is required, the difficulties which the operator must learn to overcome are not based on requirements for a very high level of skill in a single task, but rather for competent performance of each of a number of interacting tasks. In terms of the discussion of section 3.4, the effect of interactions between tasks is that the sub-environment provided for the learning of one task by poor performance on another may be very different from the desired sub-environment when that task is being performed correctly. For a skilled operator the apparent interaction between tasks may be very slight - in terms of multivariable control theory (Mesarovic (1960)) we would say that he has de-coupled the control loops - and in the measurement of final control policies the interaction terms may usually be neglected. During learning, however, the interaction may be the predominant variable affecting adaption, and should not be omitted from laboratory simulators for experiments on training.

Task interaction was introduced into the control system already described by the simple device of incorporating memory in the push-button controls. The operator had two buttons built into the arms of his chair, one for each thumb, and at any instant pressing one of the buttons would give a positive impulse at the input to the transfer-function, or pressing the other would give a negative impulse. At each push, however, the sense of the push-buttons reversed, so that the one which had previously been positive was next negative and vice versa. The chief problem introduced by this reversal was than an operator, on pushing a button and finding that it increased the error, had an innate tendency to then push the other button - increasing the error still further. A description of the various stages of learning to use this form of control is given in Gaines (1966(2)).

Finally, the disturbance at the input to the transfer function was chosen to be a square wave of 20 second period, so that its characteristics could also be learned. The overall control system is obviously very different from that required in work on operator dynamics, where sources of non-stationarity are to be avoided. The experimental paradigm is, however, one in which the phenomena of learning have every opportunity to appear.

4.2 Feedback Training Strategy

The third-order transfer-function described above is that of a linear system with three state-variables, the position, velocity and acceleration of a spot on the CRT display. The desired sub-environment of a regulatory controller is a region about zero in the 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), 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 it is predicted, cannot be cancelled instantaneously. The maximum value of the disturbance was chosen so that a skilled operator using the push-button controls could maintain the system in its linear region.

The control policy of a naive operator attempting to control a 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 non-linear, rather than linear, system. A suitable training controller will then be one which attempts to maintain the desired sub-environment by cancelling the disturbance and stabilizing the control-loop.

Figure 6 illustrates a training controller for a pure third-order system which has two feedback paths to stabilize the control loop and a third to vary the magnitude of the disturbance. The particular feedback paths chosen are, in fact, those which reduce the system to a pure integrator following a stable second-order transfer-function of variable natural-frequency and damping ratio - that is, the transfer-function described in section 4.1.

There are thus three parameters of the training controller which affect the difficulty of regulating the environment. Within the range of values used one may say that the difficulty increases as:-

the disturbance is increased from zero to its maximum value;
the undamped natural frequency is decreased from its maximum value to zero;
the damping ratio is decreased from its maximum value to zero. In the experiments this three-dimensional space was reduced to a single dimension by fixing the values of one or more of the parameters and locking the others to a single continuum of difficulty. In fixed training and open-loop training the difficulty could be set at one or more levels during the training period, whereas in feedback training it had to be made to co-vary with state of the operator.

Since the desired sub-environment is a region about zero in the state-space of the environment, 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 position of the spot were far more stringent than those on its velocity or acceleration, and hence the positional error 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 to decrease it otherwise.

This feedback training strategy was realized in practice by taking modulus of the error, subtracting the tolerance from it, and feeding the result to an integrator. The output of the integrator drove a servo multiplier whose potentiometers set the magnitudes of disturbances and feedback around the integrators in the environment. Thus, when the mean error modulus was above tolerance, the output of the integrator tended to rise and decrease the difficulty of control, whilst when it was below tolerance the output of the integrator would fall and increase the difficulty of control.

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. The overall system is complex, since many feedback loops are operative and a major part is nonlinear, but a simple analysis may be based on linearization of the outer, parameter-adjusting loop (section 4.4). This demonstrates that the system is stable for an operator whose control strategy gives rise to a limit-cycle of monotonically increasing amplitude with difficulty. Because of the high-order of the control system, and the nature of the manual controls, this condition was satisfied in the experiments reported here.

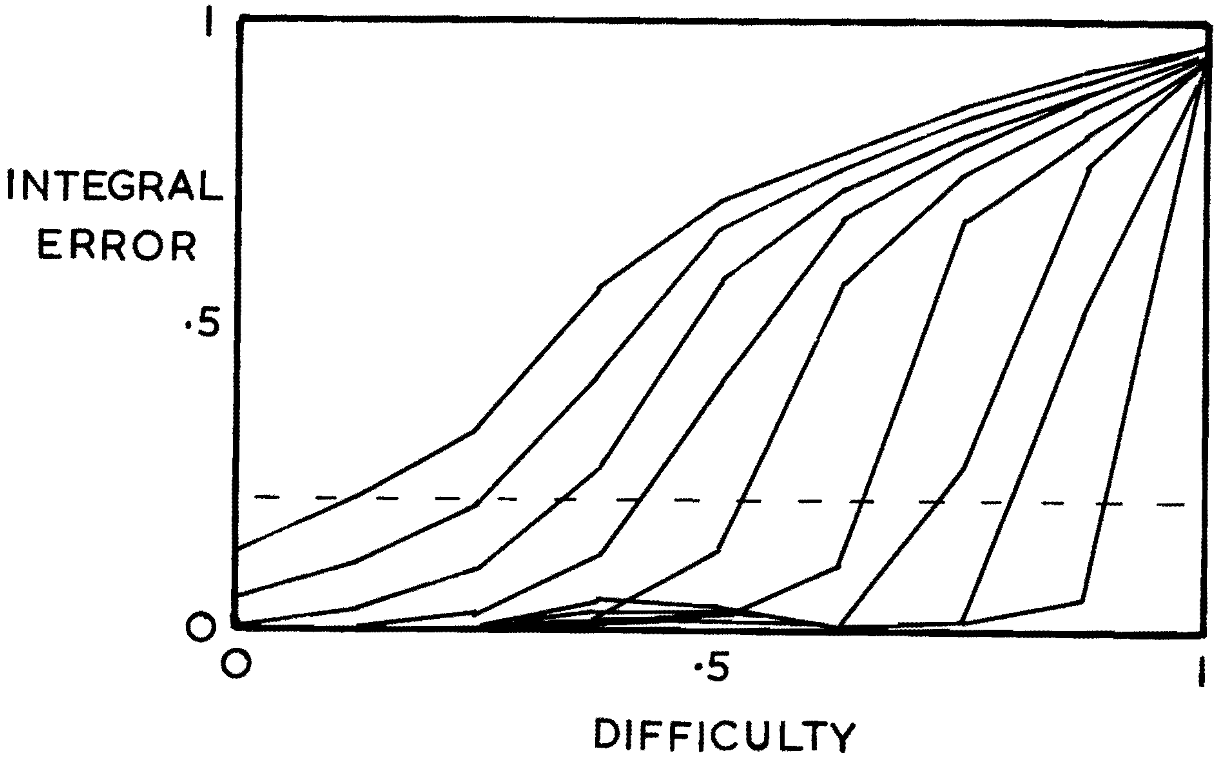

Figure 7 Variation of performance with difficulty

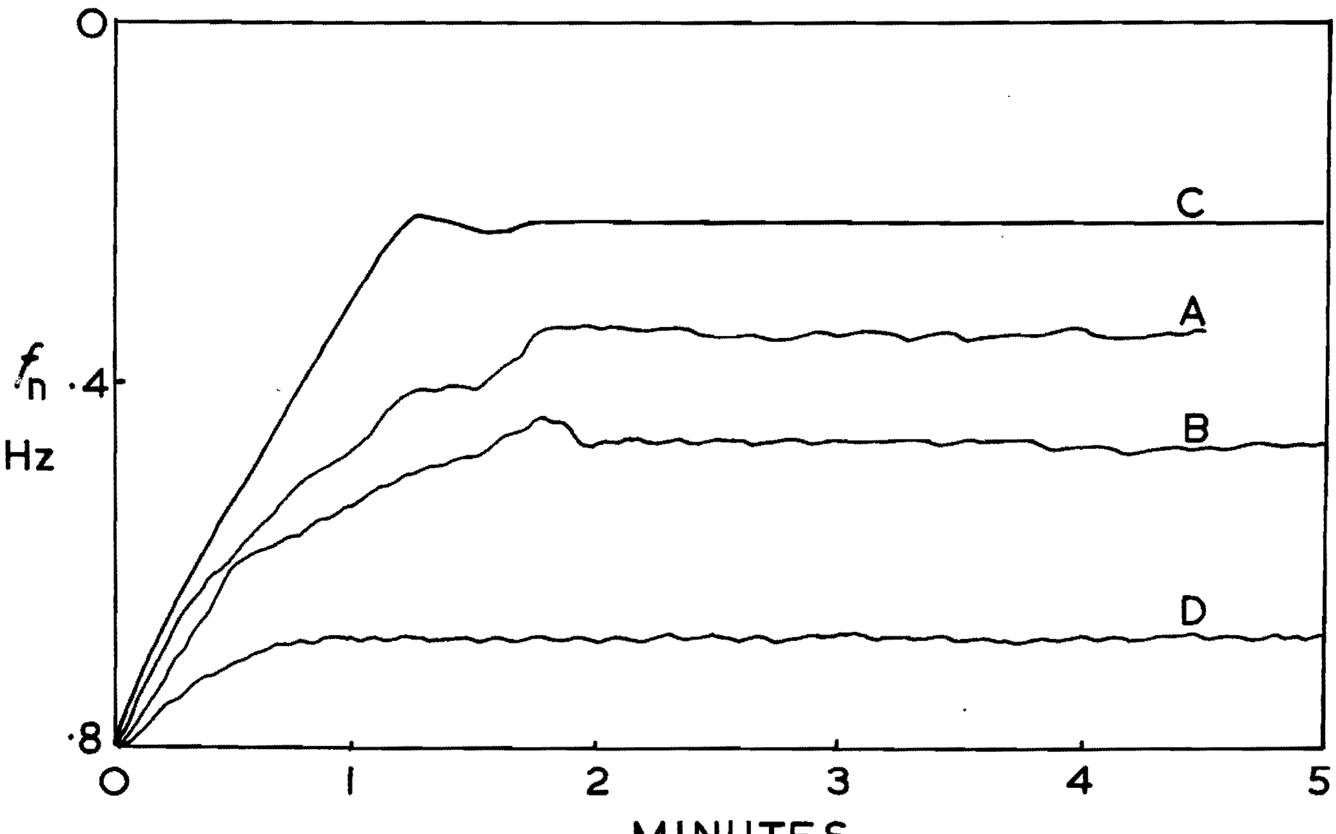

Figure 8 Response of a feedback trainer

The linearized characteristics of the outer loop depend on the form of control policy used by the operator - in particular, whether it induces a limit-cycle and at what rate this cycle is established. The experimental conditions were such that both the human operators and automatic controllers induced limit-cycles whose change in amplitude, with changing environment, was rapid compared with the change of parameters in the training controller. The outer loop then behaves as a first-order system in which the difficulty approaches its asymptotic, stable value as a decaying exponential, whose time constant decreases with the gain of the integrator in the outer loop, and decreases with the rate of change of the amplitude of limit-cycle relative to the difficulty.

Figure 7 shows the variation in the limit-cycle amplitude with respect to the difficulty of the control task, for nine non-adaptive controllers (simple relay controllers with positional and velocity feedback) of varying ability. The curves are sigmoidal, limiting at low and high error amplitudes, but have an extensive linear region centred about the mid-range of error; for training, the desired sub-environment was defined to be one in which the mean positional error modulus was .34 (on the integral error scale shown). The mid-range slope of the curves increases with increasing performance, but, due to the global non-linearity of the system, this has little effect on the response-time and only varies the rapid of turn-over to the final, stable value of difficulty.

Four examples of the variation of difficulty with time for non-adaptive controllers are given in Figure 8. A and B were generated by human operators (skilled pilots using a joystick control), whilst C and D were generated by simple relay controllers. The parameter of difficulty, in this case, is the undamped natural frequency, with the damping-ratio fixed at unity and the disturbance at a low level. The asymptotic value of natural-frequency is a measure of the skill of the pilots, or the goodness of the automatic controllers. It may also be thought of as defining the stability boundary of the pilots and controllers with respect to the natural frequency of the controlled element. For frequencies much below the asymptotic value the loop is unstable, whilst above it the loop is stable.

By measuring the asymptotic value of natural frequency for different values of damping ratio, the stability boundary for a controller in the natural frequency/damping ratio may be measured. Figure 9 shows such boundaries for three human operators (A, B, C), and two relay controllers (D, E). These have the same form as those obtained in the study of human reactions to aircraft dynamics, the linearity or non-linearity of the control policy adopted and the subjective <u>feel</u> of the simulated vehicle (Hall (1963)).

Used in this way, the feedback trainer may be regarded as a device for <u>testing</u> the skill of the human operator, and its advantage in testing will be apparent from the curves of Figure 7. Any attempt to separate the controllers of Figure 7 by giving them a test of fixed difficulty and measuring the integrated error, leads either to all the good controllers having small errors so that they cannot be distinguished, or to all the poor controllers having maximum error so that they cannot be distinguished. Thus a vertical load-line, corresponding to constant difficulty, gives a test which is insensitive over all but a narrow range. Testing the controllers by adjusting the difficulty until the error attains a standard value, however, gives a horizontal load-line (dashed line in Figure 7) which uniformly separates them according to ability throughout the whole range. It would be equally possible to apply a test of this type without using feedback, by increasing the difficulty with time and noting the level at which the smoothed error attained a prescribed value (Jex (1966)).

When the operator attached to the feedback trainer is adaptive and improves his control policy with experience, then the asymptotic value of difficulty will not be reached as rapidly, if it exists, as those of Figure 8. Instead, the difficulty will initially rise rapidly until the mean error modulus is as the prescribed level, and will thereafter vary to follow any changes in the operator's ability. Most importantly, it will maintain the desired sub-environment, whilst all the time minimizing the influence of the training controller and hence maximizing the extent to which the operator is performing the required task. Experiments comparing this feedback training strategy with both fixed and open-loop training strategies are described in the section 5.

4.3 <u>Technical Description of the Feedback Trainer</u>

The feedback trainer used in the experiments described in section 5 had three parameters of difficulty set up in the following way:-

the undamped natural frequency was constant at 2.5 radians/second, approximately .4Hz;
the damping ratio was .5 at zero difficulty, ranging to 0 at maximum difficulty;
the disturbance was zero at zero difficulty, ranging to maximum at maximum difficulty.

This particular form of feedback trainer may be defined rigorously by the mathematical relationship between its variables:-

Let E be the error in volts displayed to the operator on the CRT. This was limited smoothly such that it ranged between plus and minus 9 volts which corresponded to plus and minus 2.25 inches on the face of the CRT, or plus and minus 4 degrees angular deviation to an operator at the preferred distance of slightly below 3 feet. Marks were placed on the screen at the zero error position and at the plus and minus 2 volts position; the operator was requested to maintain the CRT beam within these outer marks. The mean error modulus maintained by the machine was 3 volts, so that an operator who kept the beam within the marks was performing above the required level and the difficulty was adjusted upwards accordingly.

Let D be the level of difficulty of the control task such that D is zero for minimum difficulty and unity for maximum difficulty. For the operators undergoing feedback training D started at zero and was adjusted according to their performance, whilst for the other groups it was fixed at .25 and .5 respectively.

Let N be the periodic disturbance in volts applied to the control system weighted by D. This was a square wave with a period of 20 seconds:

$$N = \mathrm{sign}(\sin(.3146t)).$$

Let U be the impulsive control signal resulting from pushing a push-button. This was in fact a 10 millisecond wide pulse, but in the equations it may be taken as having the value of plus or minus the dirac delta function. The sign of U is determined by the push-button pressed, the button previously pressed, and the previous value of U, such that the relationship between the button pressed and the sign of U reverses each time a button is pressed.

The controlled system equation then becomes:-

$$(s^2 + 2.5(1-D)s + 6.25)sE = 11.25U + 3 + 30\,ND,$$

where s is the differential operator, d/dt. This is, a second-order transfer function with an undamped natural frequency of 2.5 radians/second, and a damping ratio of .5(1-D), followed by a pure integration, and excited by an impulsive term, a unilateral 'drift' and a periodic disturbance of maximum amplitude, D volts.

It can be seen from the coefficients of U and ND that even at the maximum difficulty the operator only has to press the buttons at a rate of 3 per second to match the magnitude of the disturbance. Whilst the relationship between the drift term and the coefficient of U shows that a rate of about one push every 4 seconds is sufficient to compensate for the drift.

The feedback trainer equation is:-

$$SD = (3 - \mathrm{MOD}(E))/640 \quad ,$$

where MOD(E) is the modulus of the error. Since this is bounded below by zero and above by 9 volts, the minimum time for D to go from unity to zero is about 100 seconds, and from zero to unity about 200 seconds.

These equations were realized in practice using chopper-stabilized, transistorized operational amplifiers, with a DC gain of over one million and full output swing above 1 KHz. These were virtually drift-free so that the overall accuracy was set by passive components which were one percent tolerance. A motorized potentiometer with relay servo was used to adjust the level of difficulty. The repeatability of readings of D for dummy operators with a fixed control strategy was found to be plus or minus .1% in the short term, and plus or minus .5% over the course of the experiments.

4.4 Analysis of the Stability of the Feedback Trainer

The feedback training system of Figure 6 contains many feedback paths, some of which are definitely non-linear. An approximate analysis is possible, however, by linearizing the training loop and assuming that its time-constants are rather longer than those of the third-order transfer function in the main control loop.

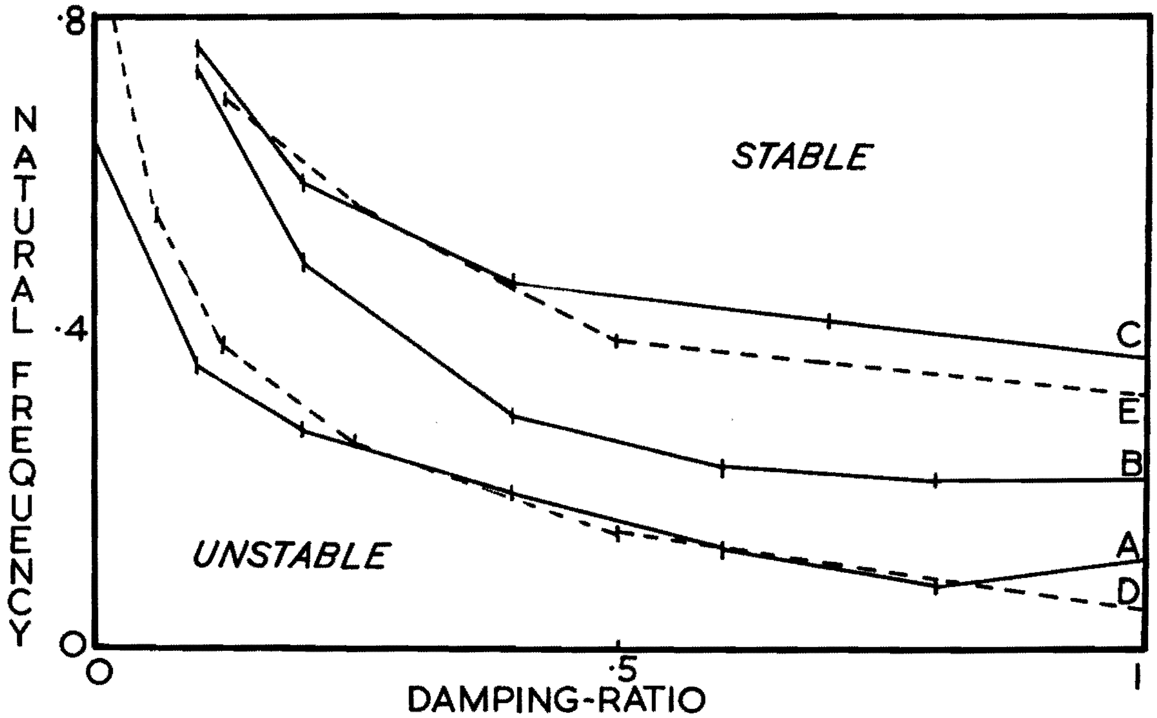

Figure 9 Stability boundaries for human and automatic control

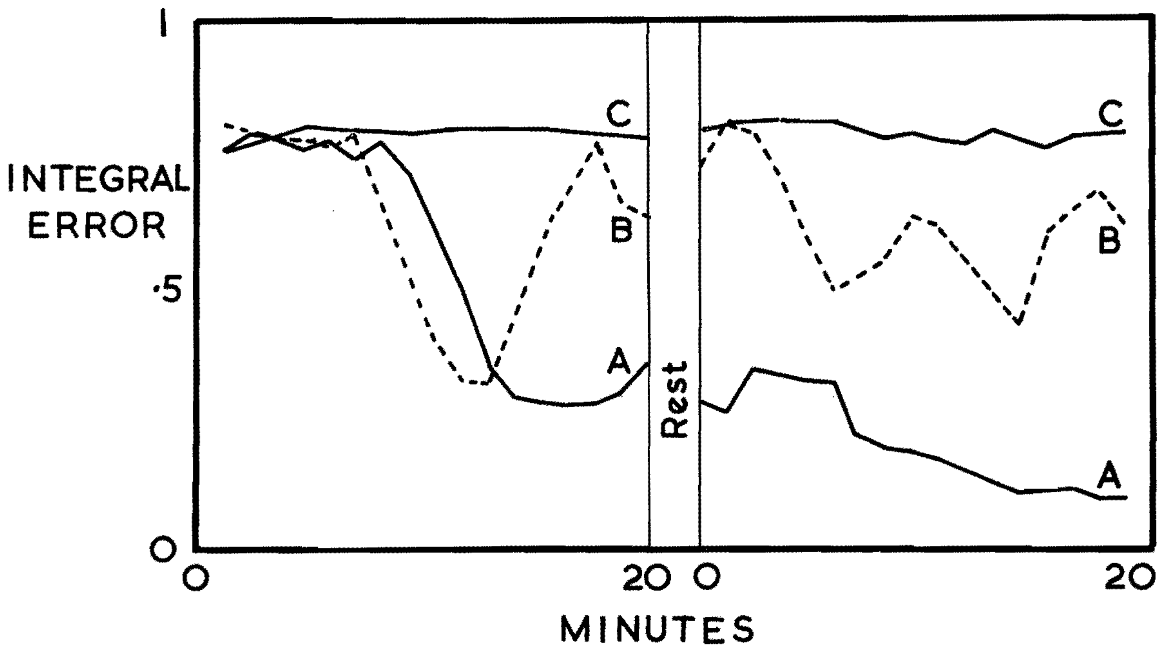

Figure 10 Open-loop training of human operators

Consider first the variation of error modulus with change of difficulty, i.e. natural-frequency, damping-ratio, disturbance, in the main control loop. For an operator with a fixed control policy, at zero disturbance, there will be two main factors, the amplitude and time-dependence, of this variation. If the control policy is nonlinear so that a limit-cycle forms, then, under the experimental conditions, the error modulus increases uniformly in amplitude for decreasing natural-frequency and damping-ratio; Figure 7 illustrates this amplitude variation. The limit cycle takes time to build up and decay as the task difficulty changes, and this time variation may be approximated by an exponential lag. If the control policy is linear, however, there is no true limit-cycle, and the error modulus rises exponentially to its maximum possible value on one side of the stability boundary, and decays exponentially to zero on the other side.

In both cases the behaviour of the error modulus (more properly the envelope or smoothed error modulus) can be approximated by the relationship:-

$$a(M-M_o) + b^2 sM = D-A ,$$

where M is the smoothed error modulus, M_o is a constant to account for M being greater than zero, D is the difficulty of the task, and A is the ability of the operator. The constant, a, will be large and b small for switching mode controllers, whilst a will be a function of the disturbance and b large for linear controllers.

The training feedback is of the form:-

$$sD = -c^2(M-M_o)$$

so that, combining the two equations,

$$D + \frac{a}{c^2} sD + \frac{b^2}{c^2} s^2 D = A ,$$

hence, for non-zero b, D follows A through a second-order transfer function with an undamped natural frequency of c/b radians/second, and a damping ratio of a/2bc.

If a is zero, then so is the damping ratio, and the system becomes oscillatory - this does occur if a linear controller is attached to the system at zero disturbance, but has no practical effect since the human operator's control policy is sufficiently non-linear to generate its own disturbance (that is, in terms of the describing function, a large remnant term).

In the experimental situation, b was small in its effect compared with a, and equation (3) then reduces to:-

$$D + \frac{a}{c^2} sD = A ,$$

so that, again, D follows A, but this time through a simple exponential lag of time-constant, a/c^2. Figure 8 illustrates the variation of D with time in these circumstances - the initial rise is rate-limited because M is greater than zero, and an exponential climb is seen only at the turn-over. The slope at turn-over varies with the asymptotic value of D since, as may be seen in Figure 9, the constant, a, is a function of the performance of the controller.

4.5 Technical Description of Adaptive Controller Used as a Simulated Operator

Although much research effort has been devoted to the study of learning-machines, this has largely concentrated on adaptive pattern-recognizers rather than adaptive controllers. Work on adaptive control itself has been mainly concerned with extending present controllers by adding some form of parameterladjustment with changing circumstances, rather than with the problem of providing a universal learning machine.

Although pattern-recognizers have been applied to control problems (Widrow and Smith (1964)), they generally require immediate feedback as to whether their response to an input was correct or incorrect (that is, whether they had classified the pattern correctly). In control problems, information as to performance is global, rather than local, and obtaining an immediate evaluation of individual responses is a major problem in its own right. A minimum mean-square error criterion, for example, may be applied to long sequences of control behaviour, but it is very difficult to use it to evaluate the individual decisions which generated that behaviour. It is not just that the overall effect is generated by a mixture of right and wrong decisions, but also that such an evaluation may not be possible since the effect of individual elements of a policy is relative to the remainder of the policy.

The complexity of structure required in a general-purpose learning machine is well-represented by that of STeLLA (Gaines and Andreae (1966)), a machine which has been simulated in a wide variety of environments, including a control task similar to that described here. This machine is potentially very useful as a model of human learning, since it implements a range of inter-dependent adaptive strategies including both open-loop and closed-loop adaption. However, the computer-time required to simulate STeLLA in even short training runs is exorbitant on standard digital computers, and it will become feasible to use machines of this complexity as models only when they become available at reasonable cost, on-line if possible.

A technique for using adaptive threshold logic pattern classifiers as adaptive controllers when only global performance criteria are available, called bootstrap learning, has been proposed recently (Widrow (1966)). Whenever it is possible to say that a sequence of behaviour is good or bad, then the probability of occurrence of every decision in that sequence is increased or decreased respectively. The evaluation of individual decisions may be correct or incorrect, their probability may be increased when it should be decreased, or vice versa, but in the long run, under certain conditions, this procedure may be shown to lead to optimal convergence.

In control systems with an error-functional performance-criterion, even this technique cannot be applied since it is impossible to say that a sequence of behaviour is good or bad, only that it is better or worse than some past sequence. The technique adopted to individualize performance feedback had, however, some similarity to bootstrap learning - a decision was good if the error modulus at some given future time decreased. This does not guarantee the satisfaction of a minimum mean error modulus criterion, but it makes some attempt to do so - a better criterion could be based on some weighted mean future error compared with some past weighted mean error. However, this simple criterion was sufficient for convergence in the control problem of interest, and enables a general-purpose adaptive threshold logic device, whose behaviour has been widely studied (Rosenblatt (1962)), to behave as an adaptive controller, and hence as a potential model of human learning.

The inputs and outputs of the adaptive threshold logic element are constrained to information rates roughly equivalent to those of the human operator. The position and velocity of the spot on the CRT are coarsely quantized and sampled at 200 milli-second intervals - a positive or negative impulse is given at the output 100 milliseconds after the corresponding input has been received.

A fifteen-bit binary pattern, $\underline{Y}$ ($y_i = \pm 1$) is generated at the input to the threshold logic element (TLE) by thresholding the position and velocity, each at seven levels - the remaining bit is permanently set. The threshold levels cover the ranges of position and velocity, nominally ± 10, with maximum discrimination in the region about zero; they were ± 6.0, ± 3.5, ± 1.5, 0.0.

The impulse at the output, I, was determined by:-

$$I = \mathrm{sgn}(w_1 y_1 + w_2 y_2 + \ldots\ldots + w_n y_n) \quad ,$$

where w_i are the weights within the TLE. These weights were adjusted according to the relationship between the error-modulus at the time the output was selected and that k sampling instants later (k = 4, say):-

$$w_i(n+k) = v w_i(n+k-1) + (1-v)\; y_i(n)\; I(n)\; r,$$

which is a conventional TLE convergence procedure, in which r is plus one (reward) if the error modulus at time, n+k, is less than the error modulus at time, n, and r is minus q (punishment) otherwise, where q is some number between zero and one (q = .25, say). The time constant of learning is of order 1/v and was about 3,000 sampling instants in the simulated operator used in the experiments.

Section 5: Experimental Evaluation of a Feedback Trainer

The effect of three major variables on learning behaviour, and the efficacy of training, were investigated in the experimental trials. These were:-

(a) The mode of training (Section 3.3), whether fixed, open-loop or feedback.

(b) The direct communication between trainer and trainee - this was not investigated in depth, but two forms of instruction, differing in the amount of information they contained about the task, were used.

(c) The type of trainee - human operators and adaptive automatic controllers were used as subjects.

5.1 Experimental Procedure

The system to be controlled was the third-order transfer-function, as described in section 4 and illustrated in Figure 6, with push-button inputs incorporating memory as controls for the human operator, and memoryless impulse inputs for the computer-simulated learning machines. The natural frequency was locked at .4 Hz, whilst the damping-ratio and disturbance co-varied as the difficulty; the damping-ratio was .5 and the disturbance zero at minimum difficulty, taking zero and maximum values respectively at maximum difficulty. The mathematical equations defining this system are detailed in section 4.3, where the parameter of difficulty is defined to be D.

Three values of difficulty were defined to be desired levels, potentially those at which the operator was to perform satisfactorily. These were:- D = .25 (low difficulty); D = .5 (high difficulty); D = .70 (very high difficulty). Subjectively, these levels of difficulty may be described as:- D = .25: very easy for the skilled operator - spot can be maintained in centre 10 per cent of CRT face. With naive subject, spot moves lackadaisically, traversing full width of screen slowly and regularly.
D =.5: More demanding for the skilled operator but still within his capability to maintain spot in centre 50 per cent and never let it get to edge of screen. With naive subject, the spot moves rapidly from one side to the other and remains for a time at each edge.
D =.7: The skilled operator can keep the spot under control but finds it difficult not to hit the edge occasionally. With a naive subject, the spot races about, both system oscillation and the disturbance affecting its movements.

Three levels of difficulty were used in testing the subject after training, rather than one defined as the desired level, because this enabled the same data to be interpreted in different ways - in particular as fixed training at different levels, and as open-loop training with transfer in difficulty either up or down.

Three modes of training were used:-

(i) High difficulty - H - the 16 operators trained under this condition had the level of difficulty set at D = .5, throughout the experiment except for the final tests. It was predicted that this group would show little learning and perform badly on all tests.

(ii) Low difficulty - L - the 24 operators trained under this condition had the level of difficulty set at D = .25, throughout the experiment except during certain test periods. 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 had the feedback training loop in throughout the experiment except during the test periods, and D varied to maintain their mean error modulus constant. It was predicted that all members of this group would learn to a high standard.

The experimental procedure was that subjects entered a soundproof room, 9' x 14', dimly but comfortably lit and free of experimental apparatus except for some tables, a typist's chair with push-buttons mounted in the arms, and a 5" screen oscilloscope mounted 3' from the ground about 3' from the chair. They were given a foolscap sheet of instructions to read initially, and then trained for 20 minutes under their appropriate training regime (for group F the difficulty was initially D = 0). Without the knowledge of the operator, and whatever the condition, at the end of 20 minutes the difficulty was put at D = .5 and the integrated error modulus over four minutes was recorded; this last procedure will be termed testing at D = .5. The operator was then given a two page questionnaire to fill in, which established his attitude to the task and his knowledge of his control policy - it also asked him to read the instructions again. This was followed by another twenty minutes training under the same conditions (the group F were re-started at the level of difficulty they had finally achieved), together with another unannounced test at D = .5. A one-page questionnaire, similar to the first, was then administered, and the operators were informed that they were to take three, four-minute tests with a one-minute break in between - these were given at D = .5, D = .25 and D = .70 respectively.

The learning-machines, simulated in a digital computer, went through a similar procedure, but the training period was thirty minutes without a break and no questionnaires were administered; a description of these machines is given in section 4.5. The effect of verbal instructions was simulated by having the machine capable of interpreting statements such as, 'When the spot is on the left and moving to the right, press the right-hand button'. It imagined the input it would receive from the environment under these conditions and the action it was told to perform, and then it rewarded itself for doing this action under these conditions. This effectively gave it an initial control policy dependent on the instructions.

The human operators in these experiments were 72 RAF pilots, all with a uniform and extensive background of flying experience. They provided a well-documented, homogeneous group, without impediments to the learning of new perceptual-motor skills, and with a high level of motivation. They were divided amongst the groups at random, on a day by day basis in which the training regime was changed at least once a day to guard against artifacts due to temperature, communication amongst subjects, and so on.

5.2 Form of Instructions

Two forms of instruction were used in the experiments:-

the weak instructions tell the operator nothing about the task he is to perform, except the performance criterion and the controls to be used; the strong instructions tell him, in addition, the nature of the coding of the push-button controls. Thus one group of operators was given the opportunity to set up a reasonable control policy which would be sufficient to keep them in the desired sub-environment at a low level of difficulty, and the other group was given no information so that their initial sub-environment was bound to be outside the desired one.

The actual instructions used were as follows:-

R.A.F., Oakington

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 control region of a display, using either a rolling-ball joystick or a pair of push-buttons.

That 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 is 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 still 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 the 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 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 at 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 then 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, were subdivided into two groups with weak or strong instructions, labelled w and s respectively. Thus there were six experimental groups altogether, with 8 operators in each of the groups Hw and Hs, 12 operators in each of the groups Lw and Ls, and 16 operators in each of the groups Fw and Fs.

5.3 Form of Questionnaires

The prime objective of the questionnaires was to provide some measure of the individual operator's attitude towards 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 the operator an interesting but relaxing task in between experimental trials.

Figures 13 and 14 show the two page questionnaire administered at the end of the first twenty five minute train/test period. The operator is first asked a question about the instructions which requires him to read them again. After estimating the duration of the experiment, he is requested to mark a number of lines, 10 cm in length according to his response to a question. All operators marked the lines without query and without apparent difficulty, so that this form of answer appears to be acceptable. The questions in Figure 14 attempt to evaluate the degree to which the operator is able to solve the control problem in verbal form.

Figure 15 shows the one page questionnaire administered at the end of the second twenty five minute train/test period. This is very similar to the first questionnaire and has the same function.

5.4 Experimental Results (Learning Curves)

The learning behaviour of the various groups and types of subject is illustrated in the accompanying figures. Figure 11 shows the variation of difficulty with time for human operators in the feedback training group, F. Within 10 minutes B has adapted to a criterion of satisfactoriness corresponding to the D = .5 level of difficulty, whereas A takes over 20 minutes to become adapted. C never adapts to this criterion, although some learning can clearly be seen. These are typical behaviours for this group - out of 32 operators, 16 reach the D = .5 level within 20 minutes, and 23 reach it within 40 minutes; no operators attained a level of difficulty less than D = .25

Figure 12 shows equivalent variations of difficulty with time for learning machines undergoing feedback training. A and B show learning to a high level and to a medium level, similar to that of human operators B and C in Figure 11. Machine C, however, shows (broken line) a new variety of learning behaviour, in that its control policy rapidly becomes satisfactory for a high level of difficulty, but then declines slowly in its effectiveness.

Instability of adaption was also shown by the human operators, especially in the early stages of learning. As radical as example as that of Figure 12 however, was obtained only once, during some preliminary experiments. It was ascribed at the time to fatigue, boredom, or some such other 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 nor 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 as to boredom or fatigue, but these do not explain the mal-adaption.

The integrated error modulus throughout each minute of training was measured for the operators who were not undergoing feedback training; this was smoothed by summing for each successive period of four minutes. Figure 10 shows the variation of integrated error with time for three subjects training at the D = .25 level of difficulty (the interval for filling in a questionnaire is shown explicitly since there was no possibility of re-starting at the same level of integrated error). A and C are normal variations, one showing learning, the other none, but B once again exhibits the phenomenom of unstable adaption. The level of error maintained by the feedback trainer for the operators under conditions of varying difficulty is .34 on the scale of Figure 10. Hence operator B in this figure is temporarily satisfactory at the D = .25 level of difficulty. These graphs are typical for operators in the group, L, training at D = .25 - those for operators training at D = .5 (H) are not shown because they were all similar to Figure 10 C.

5.5 Experimental Results (Performance Difference Between Groups)

The sequence of training and testing has been described in Section 5.1 - it is, briefly:-

	Read Instructions.
	Train under conditions Lw, Ls, Hw, Hs, Fw or Fs.
$Test_1$	At D = .5, without warning.
	Complete questionnaire and read instructions again.
	Train under same condition as before.
$Test_2$	At D = .5, without warning.
	Complete questionnaire and read test instructions.
$Test_3$	at D = .5
$Test_4$	at D = .25
$Test_5$	at D = .70

The tests lasted five minutes and the integrated error modulus during the last four minutes was measured for performance evaluation.

Figures 16 and 17 show bar graphs in which the performance of each of the 72 operators on a particular test is represented by a horizontal bar at the appropriate ordinate. The bars are grouped in six columns corresponding to the various training conditions. These charts make clear the performance differences between groups induced by different training conditions. The significance of these differences may be determined from Figure 18 which gives the mean and variance for each group plus t-statistics and variance ratios for the differences berween means and variances. Values which attain a 1% level of significance are bracketed for ease of interpretation.

Figure 16(a) shows that, at the end of the first training 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 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 each of the others however.

An analysis of this nature may be made for each of the charts of Figures 16 and 17 using the tables of Figure 18. Further discussion will concentrate on each training situation individually, however.

At the end of the first training period, the group, H, trained at D = .5 show virtually no learning ($Test_1$), but at the end of the second training session the sub-group with strong instructions is almost significantly better; however, their learning is slight. In the easiest test, $Test_4$ at D = .25, this group shows a very wide spread of performance - those who did well showed appreciable learning during the test. Thus the group, H, trained at D = .5, performed poorly in all the tests, and the difference between sub-groups induced by the instructions is slight.

The group, L, trained at D = .25, split clearly according to the instructions given - those with weak instructions do not show appreciably better performance than the group H, whilst those with strong instructions show a spread in learning from very high performance to very low throughout the tests.

The group, F, under feedback training, show a spread of learning from very high to fairly low on all the tests, except $test_4$ (at D = .25) where they are uniformly high performers. Even though both sub-groups perform well, however, there is a significant improvement with strong instructions.

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 group, H, trained at D = .5 show little learning and hence little transfer to tasks either easier or more difficult, whereas the group, L, trained at D = .25 transfer well to the more difficult tasks. Hence, as has been noted previously, open-loop training is not sufficiently defined by concepts of relative ease of transfer from easy to difficult or vice versa. The concept of an optimum level of difficulty for learning, and the possibility of maintaining this level by feedback training, are both shown to be very useful in the experimental situation investigated.

The effect of strong, or informative, instructions, and weak, or non-informative instructions, is very appreciable. It appears less with the group under difficult conditions who learnt little anyway, and with the group under optimal feedback training conditions, all of whom learnt well. It is most pronounced with the group training at D = .25, where a control policy sufficient to maintain the desired sub-environment could be set up and applied verbally - the operator had time to think. One particularly interesting feature is that the effect of instructions is more pronounced on the group under feedback training 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 greatest benefit to the naive operator. 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 is, doubtless, partially due to poor instructions, but is also an indication that an optimum interplay between direct communication and feedback training is required; the instructions should be under the control of the automatic training system.

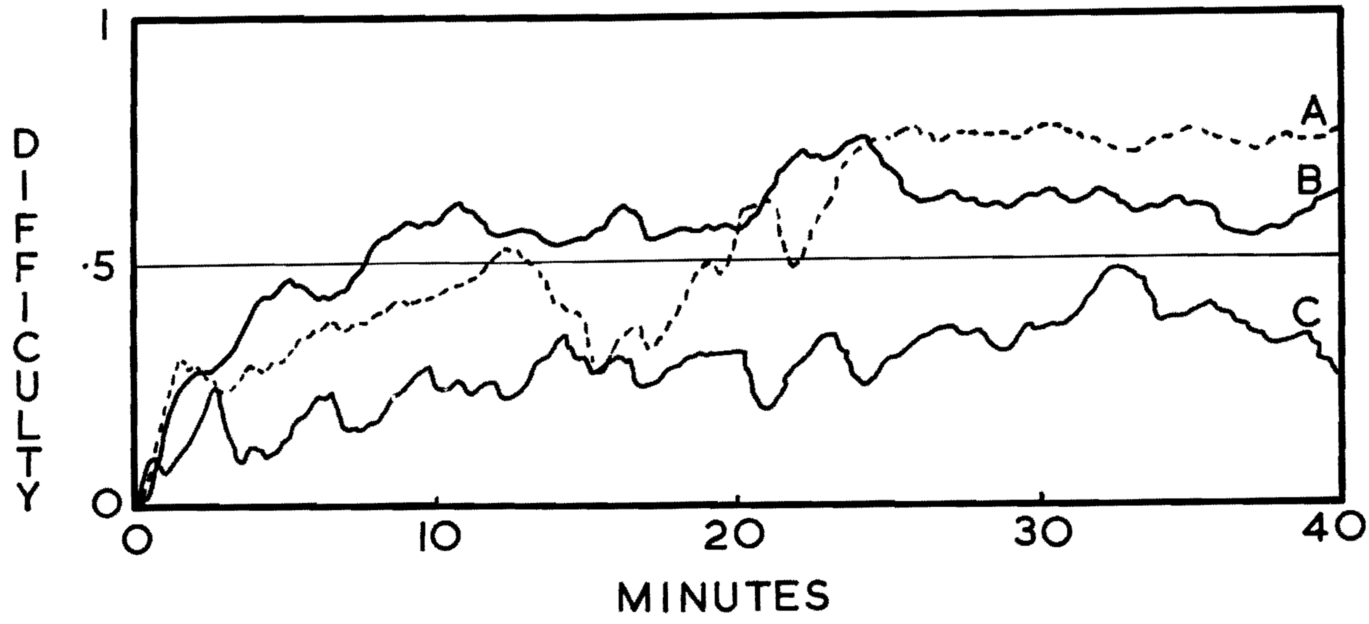

Figure 11 Feedback training of human operators

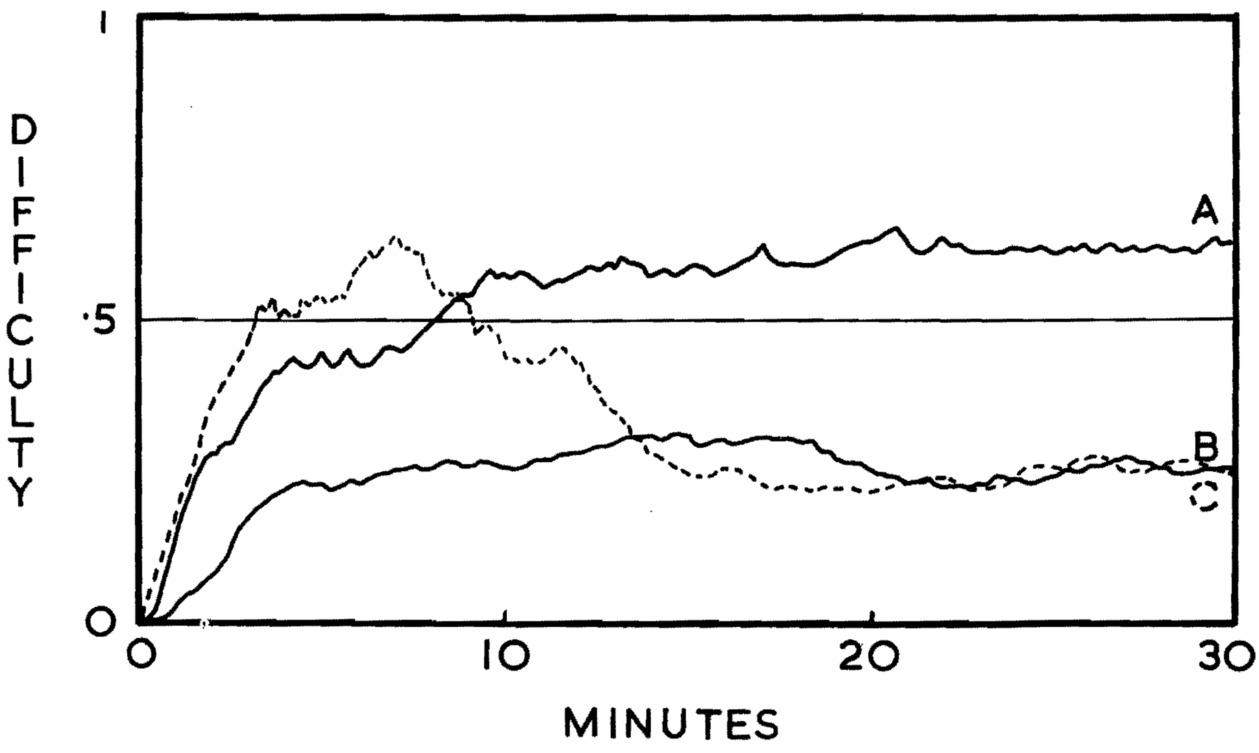

Figure 12 Feedback training of learning machines

At the end of the second training period the operators were tested ($test_2$) at D = .5, given a questionnaire to fill in, told they were to be tested, and tested again at D = .5 ($test_3$). This procedure was intended to be such as to induce stress in the operators so that the effect on the performance of the groups under the different training conditions could be evaluated. It was expected that feedback training, even if it did not improve ultimate performance, would make it more robust to the effects of stress.

Figure 17(b) is a chart of the decrease in error between $test_2$ and $test_3$, assumed to be due to the effect of knowledge about the test. In the group trained at a high level of difficulty 4 improve and 12 deteriorate; in the group trained at a low level of difficulty 12 improve and 12 deteriorate; in the group under feedback training 28 improve and 4 deteriorate. Figure 18 that the improvement in the feedback training group is significantly better than that in any of the other groups. This result is particularly interesting because it is the only performance measure where a significant difference appears between groups Ls and Fw, the groups trained at a low level with strong instructions and the group trained under feedback with weak instructions.

Figures 17(c) and 17(d) illustrate some differences in the responses of the various groups to the questionnaires. 17(c) shows the total number of words written on the three questionnaire forms. The difference between the groups Hw and Hs is particularly interesting, and may be ascribed to increased verbalization by operators who are told how to perform the task and then find they are unable to do so. 17(d) shows the position marked on the scale for evaluating the difficulty of the control task. Here the group Ls shows an interesting tendency to cluster around the central, 'Just right', position.

An overall summary of the results is given in Section 1.5, together with their implications.

5.6 Results with Learning Machines

The results with learning machines may be presented in a different form to those for human operators - it was unnecessary to use large ensembles of machines, since several replicas of the same machine could be trained under the various conditions; it was only necessary to select a small number of machines giving typical results, that is learning well or badly under each of the conditions.

Machine A in Figure 12, as can be seen, attained a high level under feedback training. The same machine trained at D = .5 showed little learning, since its final control policy contained no velocity-dependent terms. When trained at D = .25, however, it learned a very good control policy, similar to that learnt under feedback training. To investigate this phenomenom further, the same machine was trained at intermediate levels between D = .25 and D = .5, and was found to adapt up to levels of about D = .33, mal-adapting above this; the permissable open-loop training level varied according to the initial instructions given to the machine, i.e. the initial policy with which it was primed.

None of the machines tested was able to learn at the D = .5 level without prior instruction - also, all of the machines which were able to adapt above this level during feedback training showed good learning when trained at the D = .25 level. Thus, the experiments with adaptive controllers bear out the results obtained with human operators. Indeed, it was the adaption by machines when trained at the D = .25 level which suggested to the experimenter that this would lead to interesting effects with a control group of human operators.

Section 6: References

ANDREAE, J.H, (1967) Learning Machines in Encyclopaedia of Information, Linguistics and Control, Pergamon Press.

BARTLETT, F.C, (1958) Thinking, Allen and Unwin.

BEKEY, G.A, (1965) Description of the Human Operator in Control Systems in Modern Control Systems Theory, McGraw Hill.

BLAKELOCK, J.H, (1965) Automatic Control of Aircraft and Missiles, Wiley.

BOWEN, H.M, HALE, A., KELLEY, C.R, (1962) Tracking Training V: Field Study of the Effectiveness of the General Vehicular Research Tool, NAVTRADEVCEN TR 955-1

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

CHERNIKOFF, R, (1962) Human Equalization: Concept, Machine and Measurement in Complex Vehicular Controls I.A.M., Farnborough

CRITCHLEY, M, (1953) The Parietal Lobes, Edward Arnold.

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

FELDBAUM, A.A, (1963) Dual Control Theory Problems, Proc. 2nd Int. Congr. IFAC

FITTS, P.M, (1964) Perceptual-Motor Skill Learning in Categories of Human Learning, Academic Press.

GAINES, B.R, (1965) Automated Feedback-Trainers for Perceptual-Motor Skills, Tech. Rep. 1.
(1966(1)) The Evaluation and Control of Human Skill, IEE Conf. Proc., Colloquium on Human Operator Studies.
(1966(2)) Teaching Machines for Perceptual-Motor Skills, in Aspects of Educational Technology, Methuen.
(1967) Adaptive Control Theory, in Encyclopaedia of Information, Linguistics and Control, Pergamon Press.

GAINES, B.R, ANDREAE, J.H, (1966) A Learning Machine in the Context of the General Control Problem, Proc. 3rd Int. Congr.

HALL, I.A.M, (1963) Study of the Human Pilot as a Servo Element, J. Roy. Aero. Soc. 67(630) 351-360

HOLDING, D.H, (1962) Transfer Between Difficult and Easy Tasks, Brit. J. Psychol. 53(4) 397-407

HUNT, E.B, MARIN, J, STONE, P.J, (1966) Experiments in Induction, Academic Press.

HUDSON, E.M, (1964) Adaptive Training and Nonverbal Behaviour, NAVTRADEVCEN TR 1395-1

JEX, H.R, McDONNELL, J.D, PHATAK, A.V, (1966) A Critical Tracking Task for Manual Control Research, IEEE- HFE7(4) 138-145

LANGE, G.W, (1965) Synthesis of a Model of the Human Operator Engaged in a Tracking Task, Doctoral Thesis, London.

LI, Y.T, YOUNG, L.R, MEIRY, J.L, (1965) Adaptive Functions of Man in Vehicle Control Systems, IFAC Symposium on the Theory of Self-Adaptive Control Systems, Teddington, UK.

LICKLIDER, J.C.R, (1960) Quasi-linear Operator Models in the Study of Manual Tracking in Developments in Mathematical Psychology, Glencoe Free Press, Ill.

LURIA, A, (1961) The Role of Speech in the Regulation of Normal and Abnormal Behaviour, Pergamon Press.

McRUER, D, GRAHAM, D, KRENDEL, E, REISENER, W, (1965) Human Pilot Dynamics in Compensatory Systems, AFFDL-TR-65-15.

MESAROVIC, M.D, (1960) The Control of Multivariable Systems, M.I.T. Press.

MILLER, D. C, (1965) The Effects of Performance Scoring Criteria on Compensatory Tracking Behaviour, IEEE HFE 6(1).

NEWELL, A, SHAW, J. C, SIMON, H. A, (1959) Report on the General Problem Solving Problem, Proc. Int. Conf. Inf. Process., Paris.

PASK, G, (1960) The Teaching Machine as a Control Mechanism, Trans. Soc. Instr. Tech. 12(2) 72-82
(1961) An Approach to Cybernetics, Hutchison
(1964) Adaptive Teaching Systems in Teaching Machines, Pergamon Press
(1965) Comments on the Cybernetics of Ethical, Sociological and Psychological Systems in Progress in Biocybernetics 3, Elsevier.
(1965) Teaching as a Control Process, Control.

ROSENBLATT, F, (1962) Principles of Neurodynamics, Spartan Books.

SENDERS, J. W, (1961) Adaptive Teaching Machines in Programmed Learning and Computer-Based Instruction, Wiley.

STOCKBRIDGE, H. C. W, SIDDALL, G. J, (1956) Memorandum to Commonwealth Advisory Committee on Defence Science.

TAYLOR, F. V, BIRMINGHAM, H. P, (1959) That Confounded System Performance Measure - A Demonstration, Psych. Rev. 66(3)

YOUNG, L. R, STARK, L, (1965) Biological Control Systems - A Critical Review and Evaluation, NASA CR190

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 HFE 5(1) 6-13.

WARD, J. L, SENDERS, J. W, (1966) Methodological Studies of Tracking Behaviour: the Effect of Various Supplemental Information Feedbacks, Bolt, Beranek and Newman, Rep. 1458.

WEIR, D. H, PHATAK, A. V, (1967) Model of Human Operator 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(1) 36-45.

WIDROW, B, (1966) Bootstrap Learning in Threshold Logic Systems, Proc. 3rd Int. IFAC Congr.

WIDROW, B, SMITH, F. W, (1964) Pattern-Recognizing Control Systems in Computer and Information Sciences, Spartan Books.

WILDE, R. W, WESCOTT, J. H, (1962) The Characteristics of the Human Operator Engaged in a Tracking Task, Automatica 1(1) 5-19.

ZADEH, L. A, (1963) On the Definition of Adaptivity, Proc. IEEE 469-470.

ZIEGLER, P. N, BIRMINGHAM, H. P, CHERNIKOFF, R, (1962) An Equalization Teaching Machine, USNRL Rep. 5855.

Polyhedron Printers Ltd. Cambridge.