The learning of perceptual-motor skills by men and machines and its relationship to training

Brian R. Gaines
Department of Electrical Engineering Science
University of Essex

Abstract

As part of a program of research on-the feasibility and utility of automated training devices, "teaching machines," for perceptual-motor skills, a comparative study has been made of human operators and computer-simulated learning-machines learning a high order tracking task under a variety of conditions.

One outcome of this study has been to suggest, through the similarity of learning phenomena shown by men and machines, that there are important characteristics of learning which may be divorced from the particular physical realisation of a learning device, and hence are entirely conditioned by the learning environment. This has practical implications for training and for the evaluation of training devices, since a general theory of learning phenomena may be developed and standard learning artifacts may be used in the optimization of training techniques.

This paper analyses training as a control problem in the state-space of the *adaption-automaton* of the trainee, and develops a strategy for training based upon the epistemological problems of the trainee. A specific *feedback training* controller for a class of tracking skills is then outlined and its behaviour investigated theoretically and experimentally. Finally, a major experiment involving the training of both humans and learning machines under a variety of conditions is described and analysed.

1 Introduction

The main objective of the research studies described in this paper was to investigate the feasibility and utility of *teaching machines,* automatic training devices, for perceptual-motor skills (Gaines, 1965; 1966; 1967; 1968a). The value of such devices in teaching cognitive skills, such as arithmetic and language, has been widely studied (Macdonald Ross, 1969), but the use of equivalent systems in teaching perceptual-motor skills, such as flying, driving and tracking, has been investigated by few workers with varied results (Ziegler *et al*., 1962; Kelley, 1962; Hudson, 1964; Lowes *et al*, 1968; Kelley and Kelley, 1970). It has been the aim of the current studies to develop a simple and automatic training system, to demonstrate clearly its utility in at least one fairly realistic training situation, and to set up a theoretical basis for the general use of such automatic training devices.

Work on advanced adaptive controllers, in the form of learning machines (Gaines and Andreae, 1966; Andreae and Cashin, 1969) had previously suggested that many of the phenomena of human learning, and in particular those which make training necessary and viable, would also be shown by adaptive controllers in similar situations to those encountered by the human operator. 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.

This further suggested that much could be learnt about training by using non-human subjects, adaptive controllers of a simpler nature than the human operator but still, perhaps, exhibiting the basic phenomena of learning. Human subjects are vagarious and so, it was discovered, are learning machines, but the vagaries of the latter are precisely replicable and hence capable of thorough investigation in a variety of situations; if the same were true of the human operator then many of the current problems of psychology and human engineering would long ago have been solved!

The outcome of these studies has been to demonstrate that the ability of human operators to learn a novel perceptual-motor skill (one dimensional, compensatory regulation of a third-order system) is profoundly affected by the sequence of tasks given to them in training—in particular that, whilst they may learn nothing when given the final task alone, they learn very well when given a sequence of tasks of ascending difficulty automatically generated according to their ability by the training device.

This result was as expected, and demonstrated the utility of an automatic teaching system in this situation. What was also interesting, however, with wider implications than the results of the human studies alone, was that the adaptive controllers used as simulated subjects were affected in the same manner by the variations of training procedure. That is, given a family of training procedures, if it is probable that the human will learn it is probable that the machine will learn, and *vice versa*. Hence if one training procedure is better than another for the machine then it is also better for the human.

The status of these adaptive controllers as *models* of human learning in perceptual-motor skills may best be seen by comparison with that of linear models of the human tracking strategy, nobody would suggest that the human operator actually implements a linear control policy—he is a data-sampling, discrete-action device—but linear models have a great importance and utility because they show similar behaviour in certain situations and are amenable to extensive theoretical study. Similarly no-one will suppose that the human operator learns in a manner identical to the devices here described, but these devices show comparable learning behaviour and their properties are well-known.

At the present time our knowledge of the human control strategy (Gaines, 1969) ranges from the linear describing function (Licklider, 1960; McRuer *et al,* 1965), accounting for a major part of the operator output in the frequency domain, through discontinuous, data-sampling models which account for much of the fine structure in the time domain (Bekey, 1965; Lange, 1965), to analyses of the basic neuromuscular components of the human controller (Bigland and Lippold, 1954; Young and Stark, 1965). These models are, however, based on the stationary characteristics of the input/output relationship after learning, and offer no account of learning itself or of the variables affecting it.

More recent experiments (Young *et al,* 1964; Weir and Phatak, 1967; Preyss and Meiry, 1968) on the variation of the human control strategy in response to changes in the parameters of the controlled element have focussed upon the fine structure of learning over short periods rather

than the massive, long term establishment of new control strategies by a naive operator. Any account of the latter would be of great utility in industrial and military training, and the equivalence between human and machine learning outlined above suggests that suitable models may be found in certain forms of adaptive controller.

The behavioural theory of adaptive control (Zadeh, 1963; Gaines, 1968b; Gaines, 1972), as opposed to the structural theory of adaptive controllers, is not widely documented, and the initial section of this paper contains an outline of this theory leading to a formal basis for training. The theory enables teaching-machines for both cognitive and perceptual motor skills to be considered together as examples of abstract training procedures, and provides a formal foundation for generalizing the results of experiments on particular skills. The control situation and associated training system used in the experimental studies are next described, and the results of a series of experiments using both human operators and adaptive controllers are reported.

2 Adaptive control theory

The three elements constituting an adaptive control situation are (Fig. 1):

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 *et al,* 1959), and important successes have been reported in the third situation (Weizenbaum, 1966; Winograd, 1971).

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.
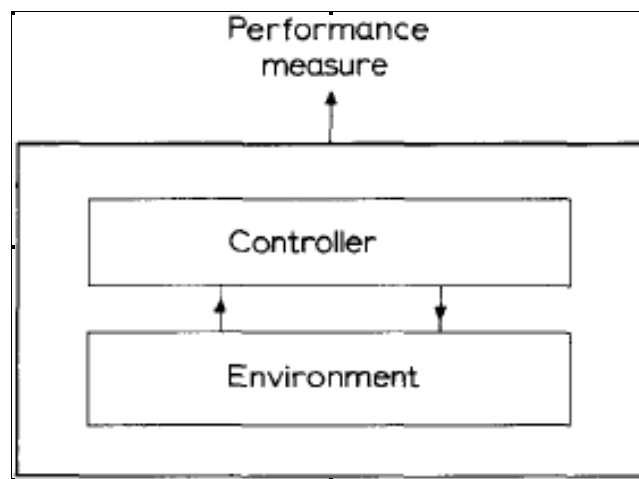

Fig. 1. Adaptive control situation

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.

2.1. 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; and 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 a wide range 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. 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.

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 adaptation-automaton as the controller adapts.

The state-space of the automaton is shown as a rectangular region in Fig. 2, 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.
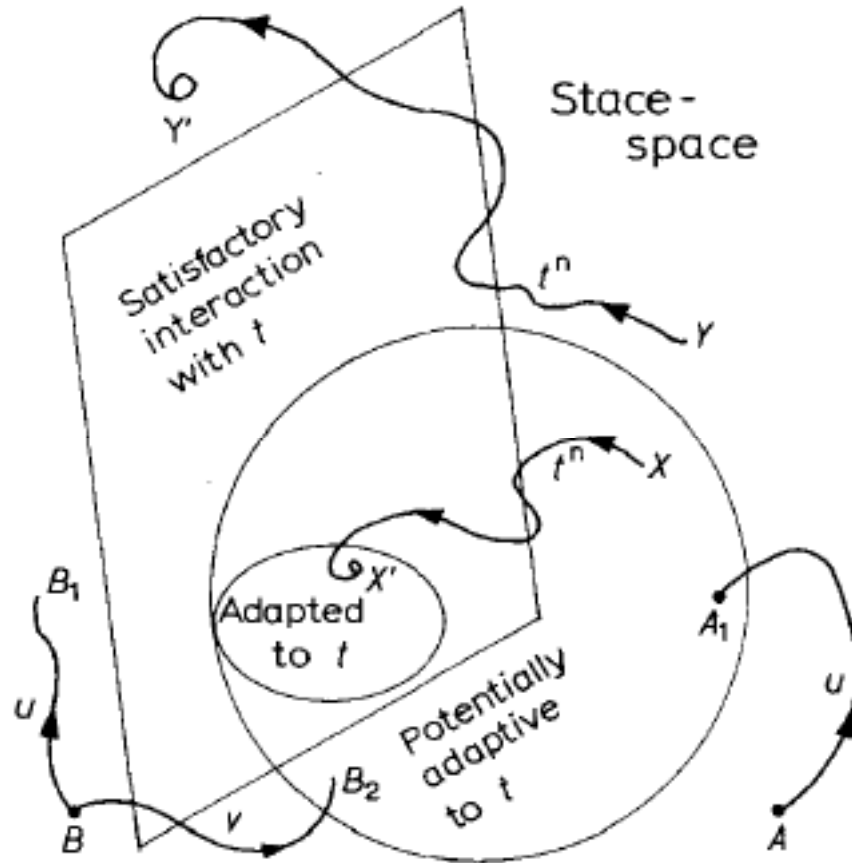

Fig. 2. Adaption to a single task

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

> started outside the potentially adaptive region, at Y. it may enter the region of satisfactory interaction but must eventually 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. Fig. 3 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.
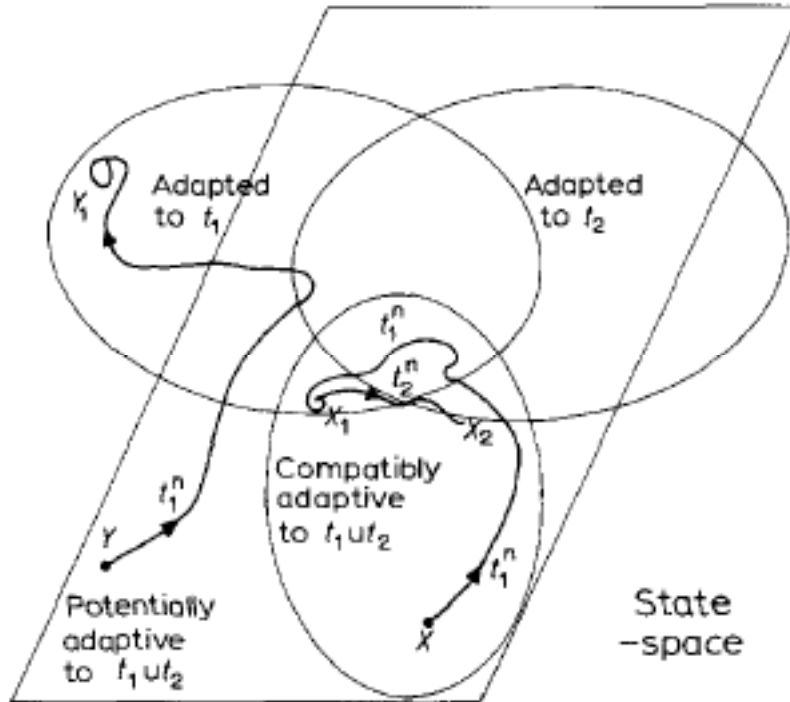
Fig. 3. Adaption to a pair of tasks

A trajectory generated by to 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 so doing 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 *jointly adapted* 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 phenomenon of *set* 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 Fig. 2 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. Fig. 13 and 14, 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 curie arise which are of main interest.

2.2 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 *trained* it for the original task.

In Fig. 2, 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, a, however, gives rise to a trajectory which terminates within the region where the controller is potentially adaptive to *t*. If the training sequence, a, 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, a, will not necessarily be suitable for all initial conditions of the controller; e.g. the trajectory induced by u from the point *B* in Fig. 2 terminates at B_1 which is outside the region of potential adaption. Fig. 4 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 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.
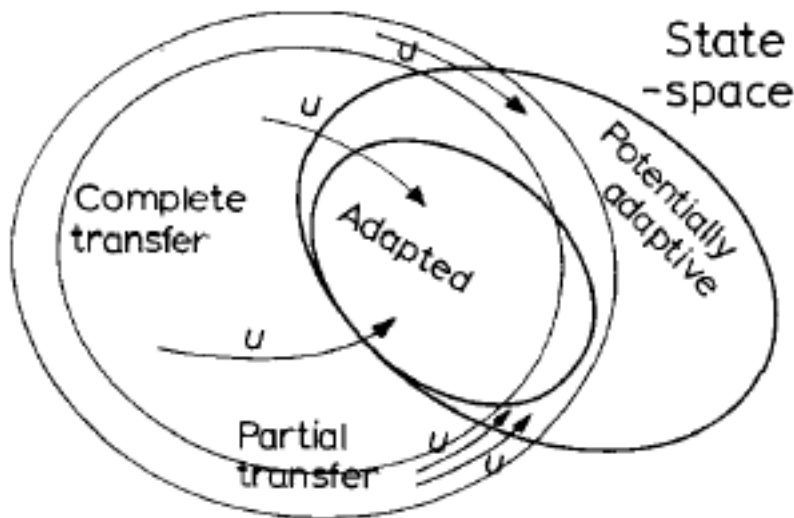

Fig. 4. Open-loop training

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, a, 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).

2.3 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 led 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 *dual control problem* (Feldbaum, 1963).
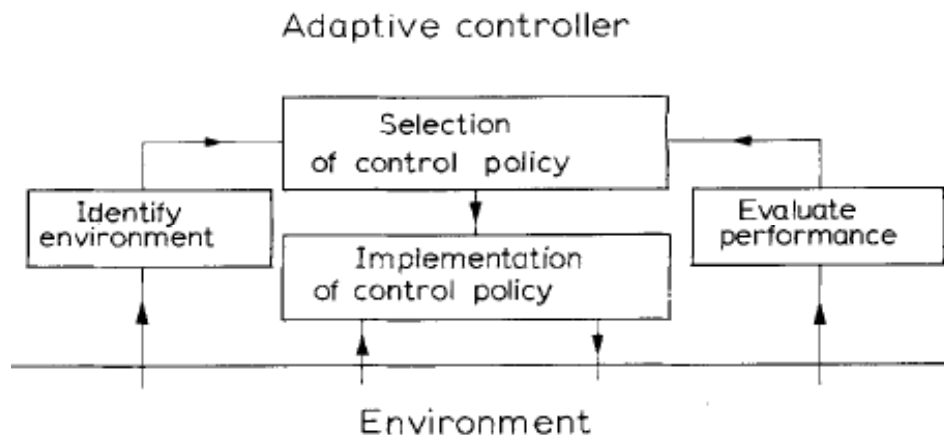

Fig. 5. Adaptive control structure

8

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, Fig. 5. This definition emphasizes the relativity of the term *adaptive,* 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 *adaptive controllers* of one era to become the *control policies* 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 developed 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 probably 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.

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, Fig. 5. 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 occur because any given control policy will generate some sub-environment, that is, it will restrict the states and state-transitions of the environment to some sub-set of the total possible behaviour. 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*

9

policy, the initial sub-environment may be so extensive that learning about it would take an unacceptably long time.

Although the concept of the sub-environment is an abstract one, it has the advantage inherent in modem "state-space" theory that a vast range of phenomena are subsumed under a single theoretical construct enabling theoretical developments to be widely applied, and providing a formal basis for the transfer of results in one field to analogies in another. In many situations the sub-environment has a direct geometrical interpretation—for example, in maze learning a trainee who uses a particular determinate search strategy (e.g. follow left hand wall) will generally restrict himself to some sub-region of the maze. For continuous control tasks where motions are considered as trajectories in a phase-plane the sub-environment is again a sub-region of the phase plane—for an unstable controller, generally a sub-region of non-linear behaviour not containing the desired end-state. The phenomenon of psychological set is a less obvious instance where an attempt to apply a strategy which has succeeded in the past leads to the subject being trapped in a fruitless approach not leading to a solution

The sub-environment phenomenon 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 an 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 phenomenon 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.

2.4 Training strategies

The obvious training strategy to alleviate the difficulties caused by the sub-environment phenomenon is to force the initial sub-environment of the controller to be the desired sub-environment. This may be represented, Fig. 6, 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 manoeuvering a simulated vehicle the training control might apply auxiliary feedback to maintain the overall control loop marginally stable.
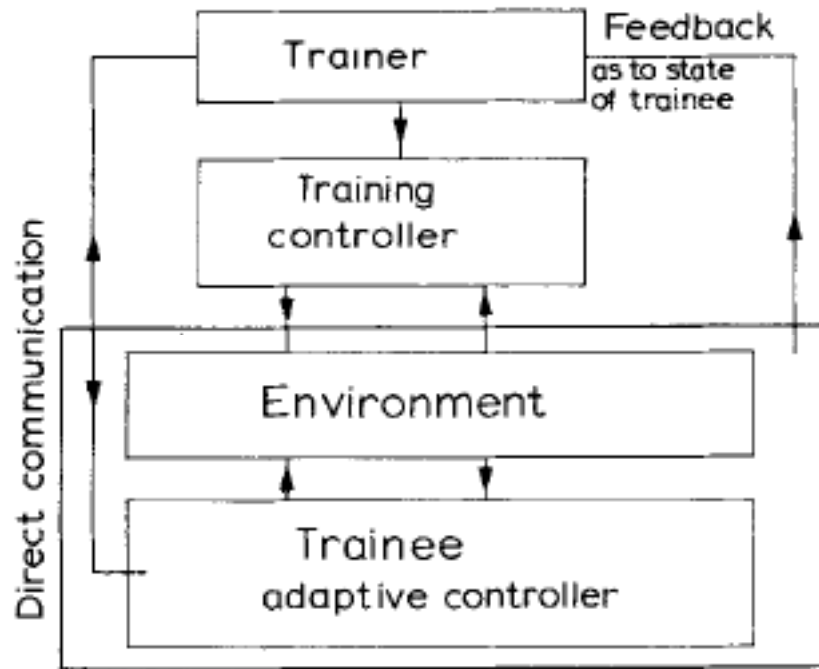

Fig. 6. Feedback training system

Since the trainee in Fig. 6 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 2.2) or according to information about the state of the trainee (feedback training, section 2.2). 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 (Luria, 1961), to allow a thorough analysis of the use of a direct communication channel between trainer and trainee. In the context of identification and performance evaluation (section 2.3), 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 does manifest itself 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.

3 An automatic feedback trainer for a tracking skill

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 the simulated aircraft dynamics used by Hall (1963) in his extensive studies of the human control strategy indicated that tracking through a second-order, stable transfer-function, with undamped natural frequency in the range, $0 \leq w_n \leq 5$ radians/sec (0.8 Hz), and a damping ratio in the range, $0 \leq k \leq 1$, was most suitable. However, the human operator is capable of compensating such a system fairly easily, and to increase the difficulty the dynamics were increased to third-order by the addition of a rate control. The overall transfer-function was thus of the form

$$G(s) = L/s(w_n^2 + 2kw_n s + s^2) \qquad [1]$$

which, by variation of w_n and k, 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 w_n 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 seconds. 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 phenomenon of fatigue and its effect 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 2.3, 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 that 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).

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.

3.1 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.

Fig. 7 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 naturalfrequency and damping ratio—that is, the transfer-function described in section 3.
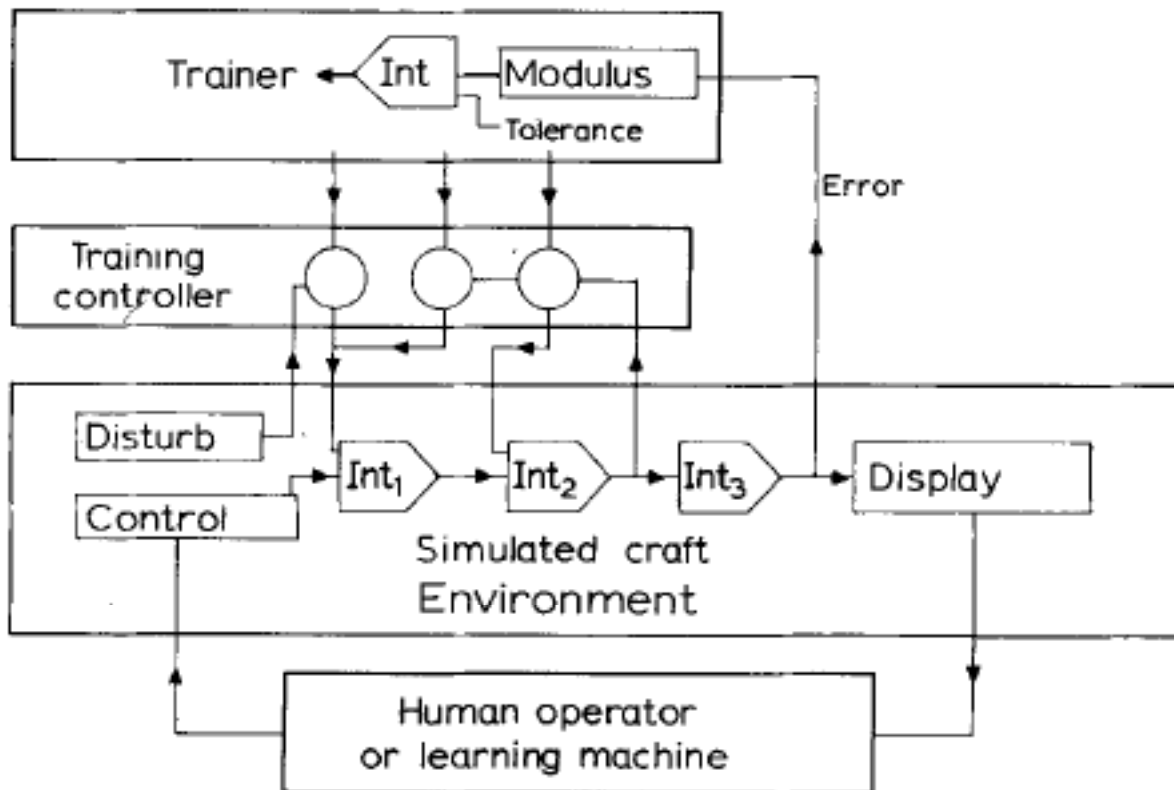

Fig. 7. Feedback trainer for a control skill

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

14

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 the modulus of the error, subtracting a 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 disturbance 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. This could clearly be a source of artifacts in the study of the trainer, and both theoretical and experimental studies were made of the dynamic behaviour of the feedback training system. The overall system of Fig. 7 is complex, since many feedback loops are operative and a major part is nonlinear, but an approximate analysis is possible 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.

3.2 Dynamic behaviour of the training system

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 non-linear so that a limit-cycle forms, then, under the experimental conditions, the error modulus increases uniformly in amplitude for decreasing naturalfrequency and damping-ratio. 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.

The relation to be expressed approximately in linear terms is that the mean error modulus and its rate of change are together linearly dependent on the difficulty of the task for the operator. Since the error modulus cannot be less than zero, for the linearization it must be expressed as a deviation from some positive value, and this is conveniently taken as the tolerated level, e_0. It is clear that the error must increase with the difficulty of the task and decrease with the operator's ability, but of these only the task difficulty is independently measurable and it is convenient to

relate the operator's ability to this. Let the task difficulty increase monotonically with the increase of some parameter, δ, and let the operator's ability be defined in related units as a, such that when α=δ the mean error modulus, $\bar{e}$, is at the tolerated level, e_0.

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

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

where s is the time differentiation operator. The constant, f, *will* be large relative to g for switching mode controllers, whilst f will be a function of the disturbance and b large for linear controllers. It is clear that f and g are functions of $\bar{e}$, δ and α—however, for small deviations from a possible stable point, Equation 2 will be valid.

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

$$s\delta = -h^2 (\bar{e} - e_o) \qquad [3]$$

where l/h^2 is the time constant of the integrator. Combining Equations 2 and 3, we obtain:

$$\delta + (f/h^2)\, s\delta + (g^2/h^2)\, s^2\, \delta = \alpha \qquad [4]$$

This is the overall equation for the training loop dynamics, and it may be seen that 6 follows a through a second-order transfer-function with undamped natural frequency of h/g radians/second, and a damping ratio of $f/2hg$.

If there is no true limit cycle and f is zero then so is the damping ratio and the training loop becomes oscillatory. It was found experimentally that this did occur when a linear controller was used as the "operator," and the difficulty oscillated widely. However, this has no practical effect since the human operator's control policy is sufficiently nonlinear to cause the value of f to dominate over that of g^2. When this is so, and g^2 can be neglected, Equation 4 reduces to:

$$\delta + (f/h^2)\, s\delta = \alpha \qquad [5]$$

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

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

$$e = (1 + sin(wt))e_o \qquad [6]$$

that is, an oscillatory signal, always positive, with a mean equal to e_0 and a frequency equal to that of the oscillations in the lower control loop (given by Equation 4.10 for the relay controller). Equations 5 and 6 then give the steady-state solution for δ as:

$$\delta = \alpha + (h^2/w)e_o cos(wt) \qquad [7]$$

Equations 4 and 7 indicate that, if the time-constant of the training loop integrator is sufficiently long so that *h is* small, then the difficulty adjustment is well-damped and little of the oscillation in the tracking task control loop appears in the difficulty variation.

This approximate analysis 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, the nature of the manual controls, and the parameters of difficulty that were used, it was reasonable to expect that this condition would hold. Experiments were carried out with simple relay servos simulating the operator to test the theory, and with human operators to check its applicability.

Fig. 8 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 midrange 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 overall non-linearity of the system, this has little effect on the response-time and only varies the rapidity of turn-over to the final, stable value of difficulty.
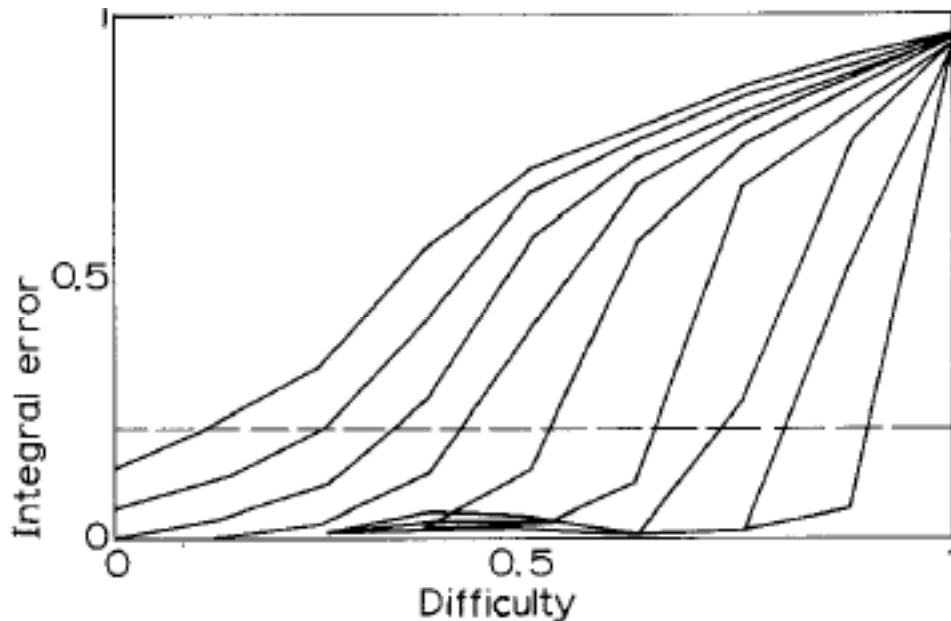

Fig. 8. Variation of performance with difficulty for relay controllers

Four examples of the variation of difficulty with time for nonadaptive controllers are given in Fig. 9. 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 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.
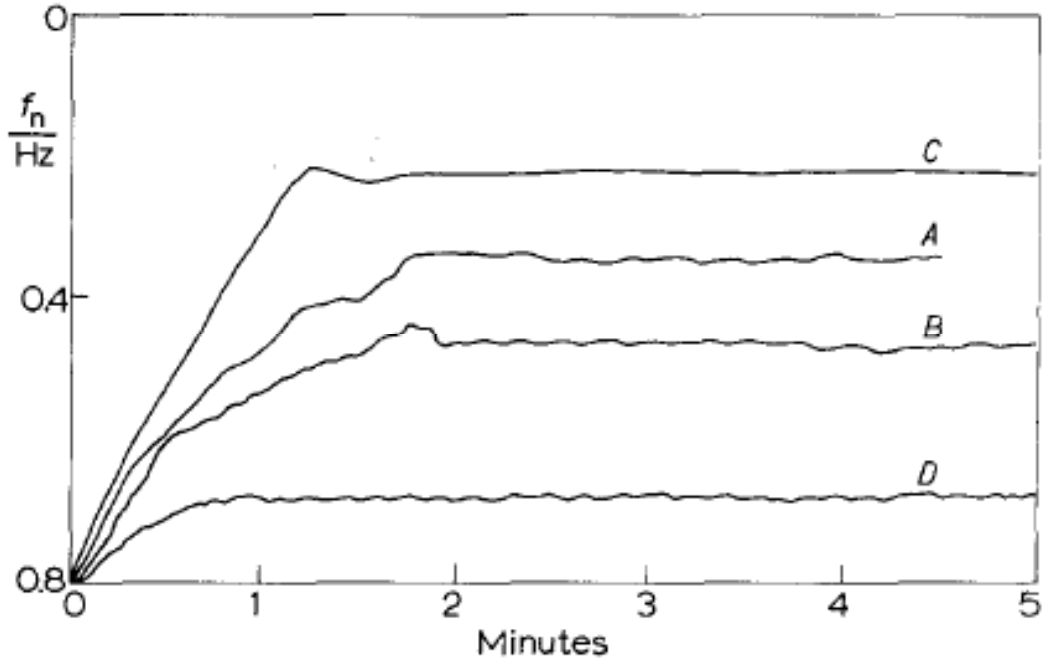
Fig. 9. Use of feedback trainer in testing

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 plane may be measured. Fig. 10 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 *feel* of the simulated vehicle (Hall, 1963).
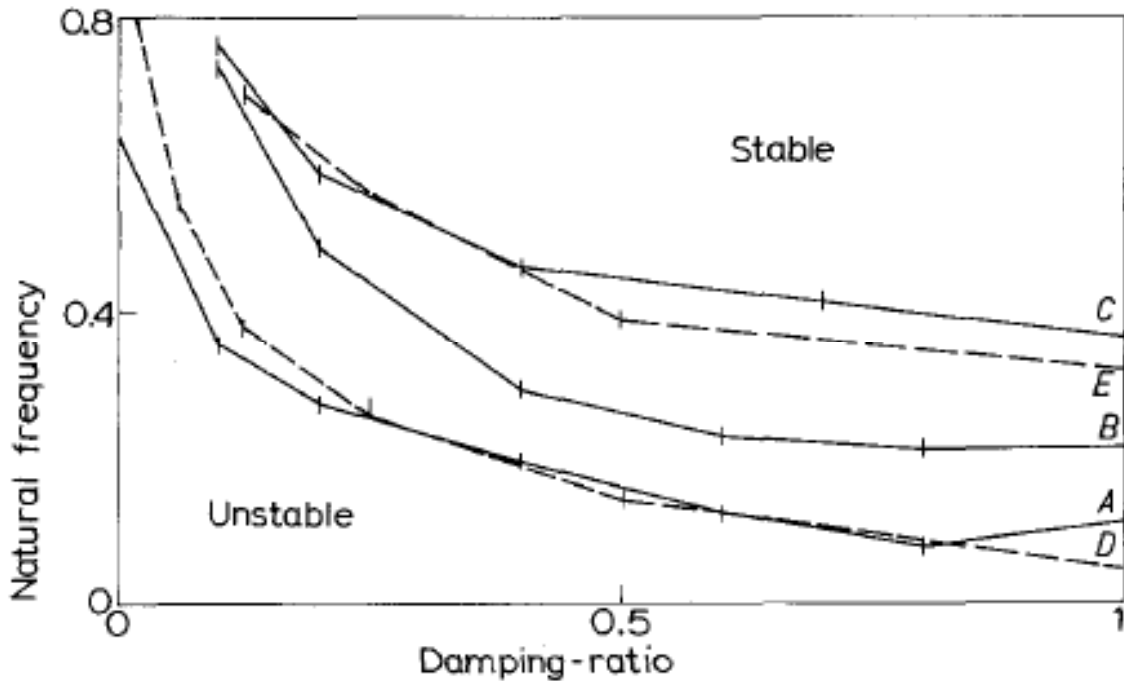

Fig. 10. Stability boundaries for human and automatic controllers

Used in this way, the feedback trainer may be regarded as a device for *testing* the skill of the human operator, and its advantage in testing will be apparent from the curves of Fig. 8. Any attempt to separate the controllers of Fig. 8 by giving them a test of fixed difficulty and measuring the integrated error, leads wither 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 Fig. 8) which uniformly separates them according to ability throughout the whole range. It is also possible to apply a test of this type without feedback, by increasing the difficulty with time and noting the level at which the smoothed error attained a prescribed value (Jex *et al, 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 shown in Fig. 9. Instead, the difficulty will rise rapidly at first until the mean error modulus is at the prescribed level, and will thereafter 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 are described in Section 4. In the following section the learning controller used in the experiments as an alternative to the human operator is described.

3.3 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 parameter-adjustment 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 Amith, *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; Andreae and Cashin, 1969), 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 by Widrow (1966). Whenever it is possible to say that a *sequence* of behaviour is a good or bad, then the probability of occurrence of every decision in the 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 in the system described here has, however, some similarity to bootstrap learning—a decision is good if the error modulus at some given future time decreases. 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, to behave as an adaptive controller, and hence as a potential model of human learning.

In the experimental system, the inputs and outputs of the adaptive threshold logic element (ATLE) were constrained to information rates roughly equivalent to those of the human operator. The position and velocity of the spot on the CRT were coarsely quantized and sampled at 200 millisecond intervals—a positive or negative impulse was given at the output 100 milliseconds after the corresponding input had been received.

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

The impulse at the output, θ, was determined by:

$$\theta = sgn(\Sigma\ w_i y_i) \qquad [8]$$

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).

$$w_i\ (n+k) = \beta\ w_i\ (n+k-1) + (1-\beta)\ y_i\ (n)\ \theta(n)\ \delta \qquad [9]$$

which is a conventional TLE convergence procedure, in which δ is plus one (reward) if the error modulus at time, *n+k, is* less than the error modulus at time, *n,* and δ is minus one (punishment) otherwise.

Adjustment of k and β enabled a family of ATLE controllers to be generated, some of which showed learning to a level of performance comparable to some human operators in the

experimental task, and others of which showed a variety of behaviour between total maladaption and stable learning.

4 Experiments on the utility of feedback training

The affect 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 2.2), whether fixed, open-loop, or adaptive.

(b) The form of direct communication between trainer and trainee— 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.

The system to be controlled was the third-order transfer-function, as described in section 3 and illustrated in Fig. 7, with push-button inputs incorporating memory as controls for the human operator, and memoryless, impulsive 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 (δ=0), taking zero and maximum values respectively at maximum difficulty (δ =1) .

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

In order to provide an adequate evaluation of the operators' capabilities after training it was necessary to test their performance at several levels of difficulty, and it was convenient to choose these also as the levels for open-loop training, since the same experimental results could then be used as a basis for the evaluation of fixed training and of open-loop training at higher, or lower, levels of difficulty than the required task. Three levels of difficulty were selected δ = 0.*25, 0.50, 0.70,* and their relative levels can be appreciated from approximate descriptions of the appearance of the system to the operator:

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

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

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

Although the circumstances were such that no measurement of the long term retention of learning was possible, it was felt desirable to obtain some measure of the "robustness" of the learning, and a form of "instruction-induced stress" was introduced by informing the operators that they were being tested and measuring the effect on performance. Questionnaires were also adminstered to give some quantitative measure of the operators' verbal reactions to, and knowledge of, the training situation, and as shown in Fig. 11 and 12.

Name . . . .

Rank . . . .

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

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

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

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

Very Boring |_____| Very Interesting

ANY COMMENTS:

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

Far too Difficult |_____| Far too Easy

Just Right

ANY COMMENTS:

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

Complete Failure |_____| Perfectly

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

Complete Failure Perfectly

Fig. 11. Questionnaire part I

How many further practice runs would you need to perform the task — ade-
quately? . . . . Perfectly? . . . .

ANY COMMENTS:

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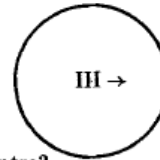

The spot is stationary as shown: III.

which push-button would you press?   LEFT   RIGHT

ANY COMMENTS:

Having pushed the button the spot moves as shown: III →

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

ANY FINAL COMMENTS:

Fig. 12. Questionnaire part II

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

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

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

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

Two forms of instruction were used in the experiments, and are recorded in Table 1. The *weak* instructions (w) tell the operator nothing about the task he is to perform, except the performance criterion and the controls to be used. The *strong* instructions (s) tell him, in addition, the nature of the push-button controls. Thus one group of operators was given the opportunity to set up a reasonable control policy which would probably be sufficient to keep them in the desired sub-environment at the low *(L: $\delta=0.25$)* level of difficulty, but not at the high (H: $\delta=0.5$) level, and

the other group was given no information so that their initial sub-environment was bound to be outside their desired one. Half the operators in each experimental group were given each form of instruction, and the overall conditions for the groups are shown in Table 2.

The experimental procedure is shown in Table 3—all subjects were given the instructions to read initially, and then trained for 20 minutes under high (H), low *(L)* or feedback *(F)* conditions. Without the knowledge of the operator, at the end of this period the difficulty was set to the high (H: $\delta=0.5$*)* level and the integrated error modulus over 4 minutes was recorded (a test at H-level). The operator was then given the two-page questionnaire shown in Figs. 11 and 12 to fill in, which established his attitude to the task and his knowledge of the control policy—it also asked him to read the instructions again. This was followed by another 20 minutes training under the same conditions (the feedback *(F)* group were re-started at the level of difficulty they had finally achieved), together with another unannounced test at the H-level. A one-page questionnaire similar to the first was then administered, and the operators were informed that they were to take three 4-minute tests—these were administered at the H-level, L-level and V-level. The difference between the final unannounced and announced tests at the H-level was taken as a measure of the effect of instruction-induced stress.

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. 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.

4.1 Experimental results

The learning behaviour of the various groups and types of subject is illustrated in the accompanying figures. Fig. 13 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 H-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 36 operators, 16 reach the H-level within 20 minutes, and 23 reach it within 40 minutes; no operators attained a level of difficulty less than the L-level *($\delta=0.25$)*.
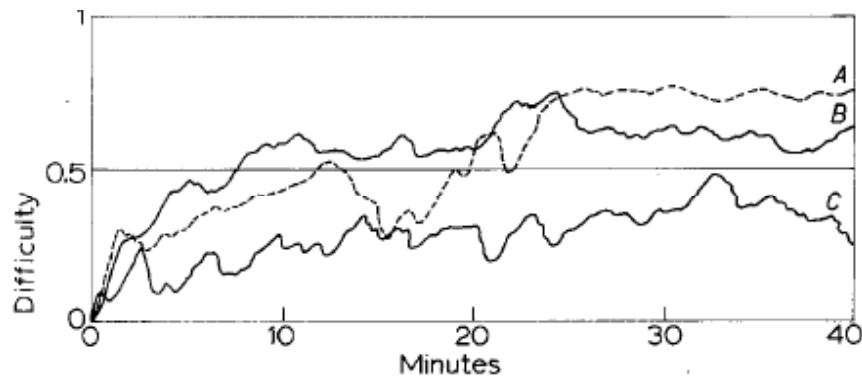

Fig. 13. Feedback training of human operators

Fig. 14 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 Fig. 13. 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.
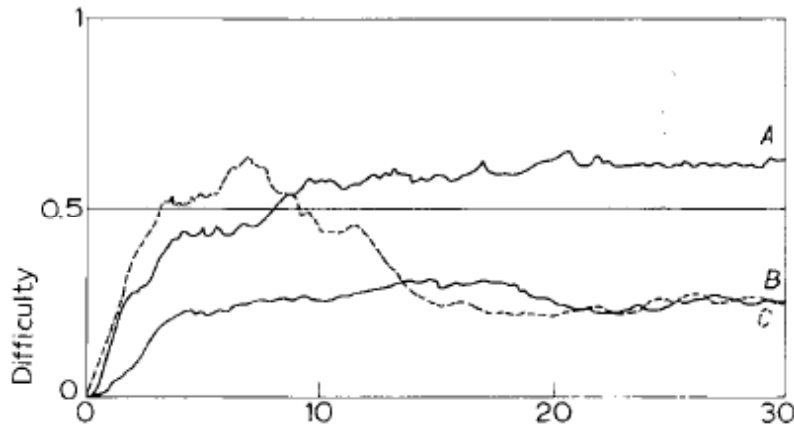

Fig. 14. Feedback training of learning controllers

Instability of adaption was also shown by the human operators, especially in the early stages of learning. As radical as example as that of Fig. 14 however, was obtained only once, during some preliminary experiments. It was ascribed at the time to fatigue, boredom, or some much 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 phenomenon 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. Fig. 15 shows the variation of integrated error with time for three subjects training at the L-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 phenomenon 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 Fig. 15. Hence operator B in this figure is temporarily satisfactory at the L-level of difficulty. These graphs are typical for operators in the group training at the L-level—those for operators training at the H-level are not shown because they are all similar to Fig. 15 C.
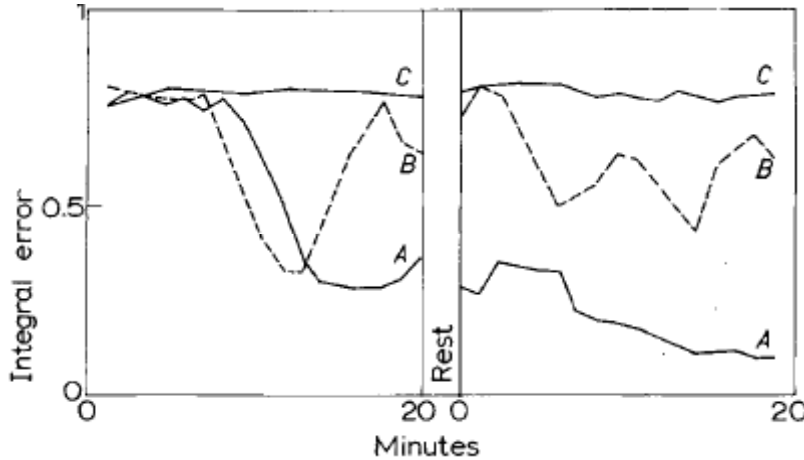
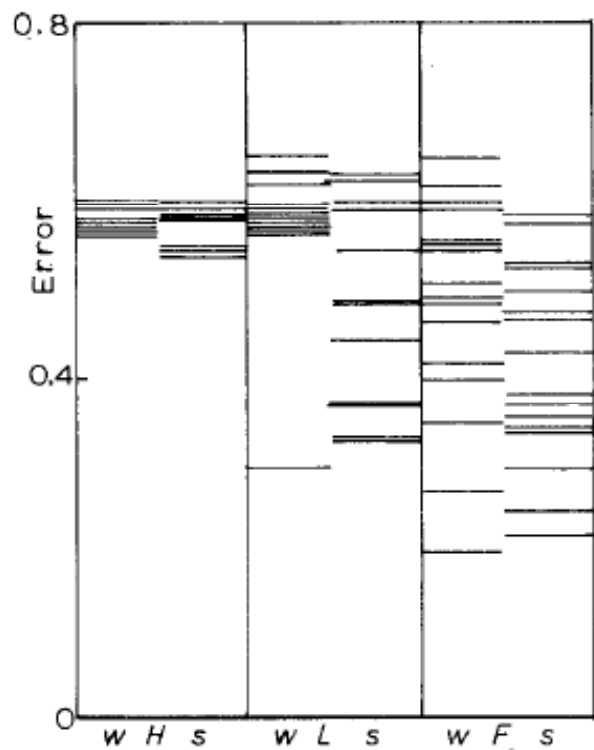**Fig. 15. Open-loop training of human operators**

4.2 Performance on the tests

Figs. 16 and 17 show bar graphs in which the performance of each of the 72 operators on a particular test is shown as a horizontal line at the appropriate ordinate. The bars are grouped into six columns according to the training conditions, and the charts clearly illustrate the differences between the experimental groups; these were also evaluated statistically (Gaines, 1967).
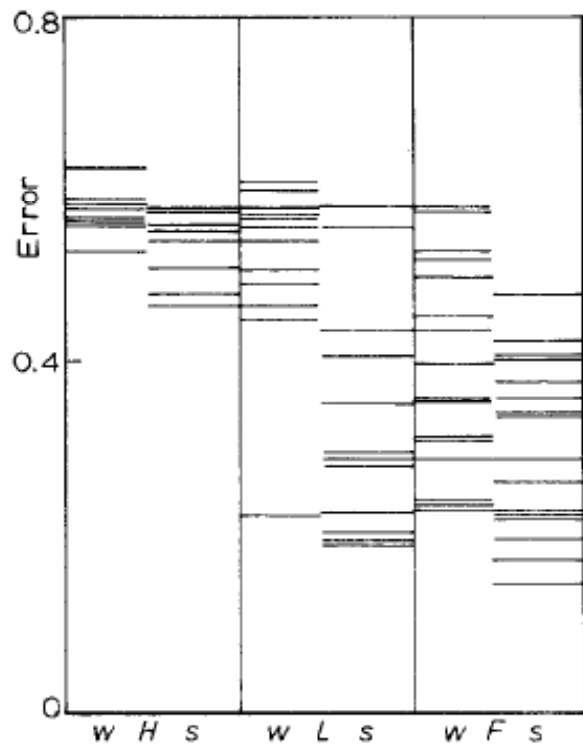
Fig. 16(a) shows that, on tests at the H-level of difficulty (H: $\delta=0.5$) session, the *Hw*, *Hs*, and *Lw* groups show a uniformly low level of performance, whilst the *Ls*, *Fw* and *Fs* groups show a spread of performance ranging from the very low to very high; only the *Fs* group is significantly better than the first three groups, however. The spread of the two groups, *Hw* and as is significantly less than all the other groups, and this may be related to the sigmoidal nature of plots of performance against difficulty.

From Fig. 16(b) it appears that, at the end of the second training session, these differences have been enhanced, and all three groups, *Ls*, *Fw* and *Fs*, are now better than *Hw*, *Hs*, and *Lw*, but there are no significant differences within each set of three groups. However, the graph is suggestive of an overall difference between groups with strong and groups with weak instructions. Fig. 16(c), which shows the performance on a test at the same level of difficulty after the operators were informed of the test, has the same interpretation as Fig. 16(b). However, the variable of greatest interest is the relationship between these figures and the effect of instructions on performance—this is shown by the plot of performance differences in Fig. 17(b) and analysed in Section 4.4.
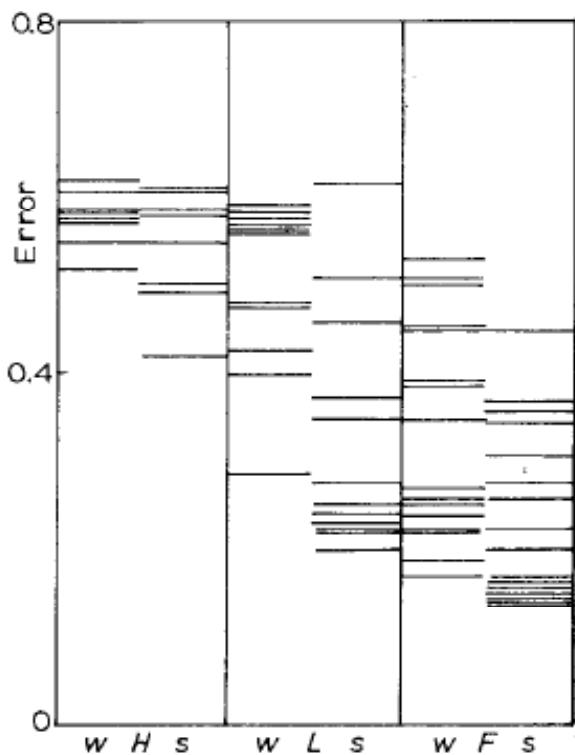
Fig. 16(d) shows the performances on a test of lower difficulty *(L. $\delta = 0.25$)*, and in this the *Hw*, *as*, and *Lw* groups again appear as not significantly different, the *Ls* and *Fw* groups are significantly better than these three, and the *Fs* group is significantly better than the other five. It is interesting to note the wide spread in learning of the *Hw*, *Hs* and *Lw* groups, particularly since the *Lw* group is being tested at the same level as that at which it trained. The high performance of some members of the H group on this test is due, in some part, to learning during the five minute test period. The test results at a higher level of difficulty *(V: $\delta=0.7$)*, shown in Fig. 17(a), demonstrate the spread in abilities which still exists in the better groups.
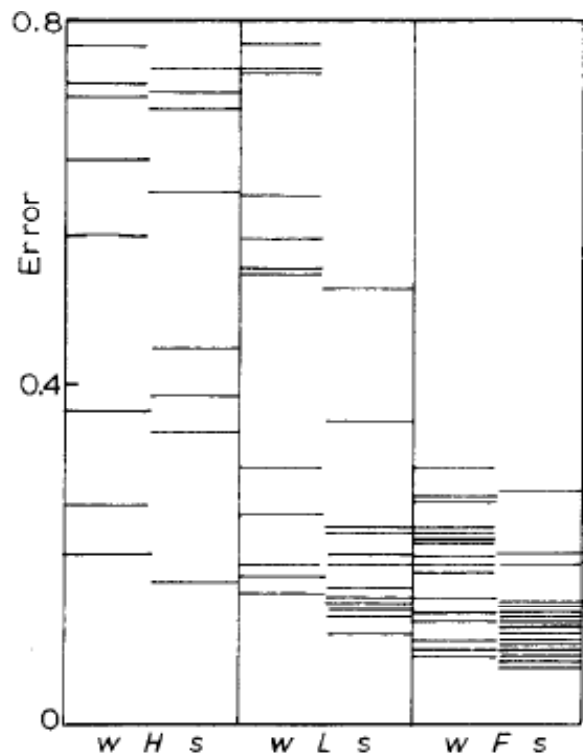
26

Fig. 16. Experimental results part I
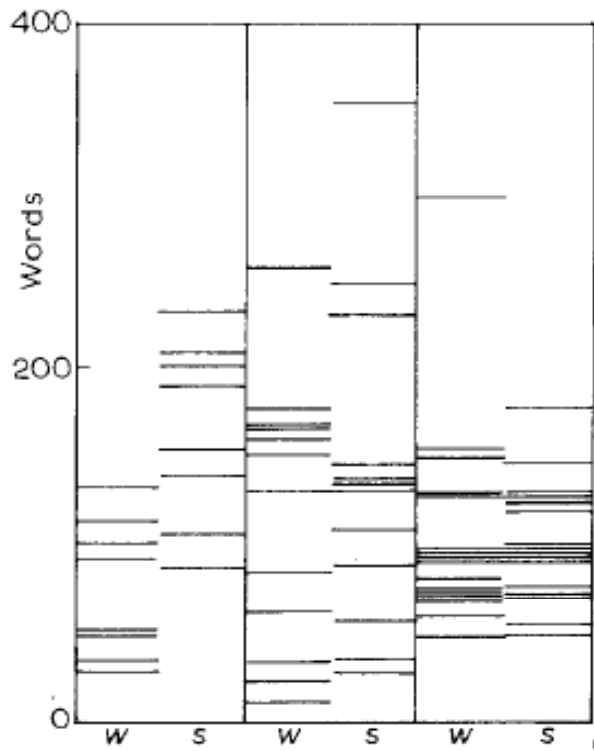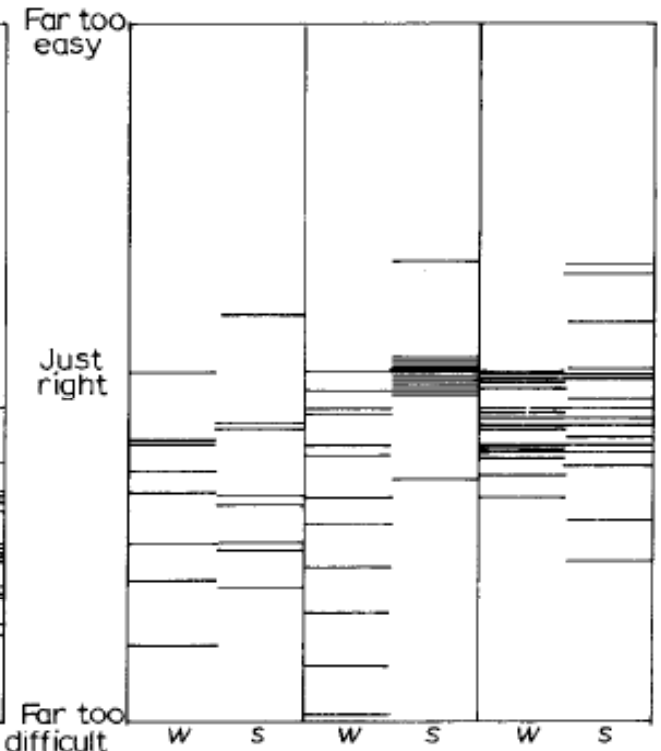
(a) Performance fifth test, $\delta=0.7$, V

(b) Performance change under stress

(c) Words on questionnaires

(d) Estimate of difficulty

Fig. 17. Experimental results part II

4.3 Effect of instructions

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

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

4.4 The effect of instruction-induced "stress"

Fig. 17(b) shows, for each operator, the error on the third test minus that on the second, and hence a positive "error difference" corresponds to an improvement of performance. Since the third test is at the same level of difficulty as the second test *(H:δ=0.5),* and follows it after an internal with no practice at the tracking task, any error difference must be due to the effect of events in the intervening interval. During this interval the operator filled in the second questionnaire, and was then informed that his proficiency was to be tested. This information was expected to be stress-inducing, and hence, possibly, to cause a deterioration in the operator's performance. Alone, however, the internal of other activity might be expected to lead to an improvement in performance,

From Fig. 17(b), it may be seen that the effect of the instructions varies widely over the three groups: out of the sixteen operators trained at a high level of difficulty, twelve show a deterioration in performance: the group trained at a low level of difficulty split equally into twelve who get worse and twelve who improve; out of the thirty-two operators trained under feedback conditions, only four show any deterioration, and the general performance improvement is very marked. Only the improved performance of the *F* group over the *H* and *L* groups is significant at the 1% level.

The basis for this improvement is not easy to analyse since the effect of the information that the operator is to be tested is not necessarily one of "stress" and it is in any event too superficial a conclusion to state that the performance of the operators trained under feedback conditions improved with "stress," whilst that of operators trained under open-loop conditions deteriorated or remained unchanged. For example, the mean level of performance of each of the groups

differs widely, and the nonlinearity of the performance scale magnifies changes at the mid-level of performance and minimizes the apparent extent of those at very high, or very low, levels. However, taking account of this effect only increases the contrast between the three groups, since the deterioration of the *H* group would be more pronounced, as would the improvement of the *F* group.

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

4.5 Responses to the questionnaires

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

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

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

The estimates of task difficulty show interesting differences between the groups, apparent in Fig. 17(d) - as expected, the *Hw, Hs* and *Lw* groups, all of whom performed badly, find the task too difficult, but there is a remarkable consensus of opinion in the *Ls* group, emphasized by the availability on this particular scale of a centre point marked "just right." The total number of words written on the questionnaires also brings out an interesting difference between the groups,

in that the Hs group wrote over twice as many words as the *Hw* group. It may be noted from Fig. 17(c) that the *Hs* group has no individual writing less than about ninety words, which is very much higher than the minima of the other groups. This seems to reflect the unique status of the *Hs* group, who were told how to do the task and then found they could not in practice - a situation apparently creating much verbal behaviour.

4.6 DIFFERENCES BETWEEN THE EXPERIMENTAL GROUPS

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

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

The group, *F*. trained under feedback conditions in which 6 was adjusted to maintain their mean error constant, again split according to the instructions given, but not in nearly so dramatic a manner as the *L* group. Both *Fw* and *as* groups learn to a high level of performance, and are significantly better than the *Hw, Hs* and *Fw* groups on all test. The *Fs* group is significantly better than the *Ls* group on the fourth test *(L: δ= 0.25),* which is particularly interesting since this is the level at which the *Ls* group trained. There is not significant difference between the *Fw* and *Ls* groups on any of the tests, and indeed the *Ls* group is slightly better in three out of the five. However, under instruction-induced stress, both the *F* groups show significantly better results than the *Ls* group, and, of course, all other groups.

4.7 Comparison with learning machines

In analysing the corresponding results using the adaptive controllers of Section 3.3 as learning machines emulating the human operator, it is important to be precise about the basis for comparison of the human and machine results. It has been noted in Section 1 that the machines are intended not as detailed models of human tracking behaviour but rather as overall models of the learning behaviour - that is, of performance differences corresponding to different training regimes.

It is also important to distinguish between the range of different learning machines available and the range of different human beings used in the experiments. A number of human subjects have to be used in a given experiment because learning is irreversible and it is necessary to examine

the effects of different training regimes on different, but "equivalent" (in some sense), ensembles of individuals. On the other hand, the learning machines may be reset to a given initial state, and the exact differential effects of training regimes on *identical* machines may be measured. Thus one machine is equivalent to an ensemble of human subjects, and the purpose of examining the learning behaviour of different machines is to match different *sterotypes* of human behaviour rather than to match individual behaviour (which can only be inferred, not measured).

This argument may be clarified by considering some of the possibilities that might have arisen in testing the variety of different learning machines generated by varying the parameters k and β (Section 3.3)

(i) No machines learn—bad choice of learning machines—chose another or drop experiment.

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

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

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

The majority of learning machines generated by variation of k and , β were unable to learn the task under any conditions (and might be thought of as emulating sub-human, or pathological behaviour). Four machines which did learn, at least under some conditions, were selected as covering the rang of basic behaviour stereotypes —A *(k=4, ,β=0.625)*, B *(k=4, ,β=0.6)*, C *(k=5, β=0.625)* and D *(k=5, β=0.6)*. Fig. 14 shows the variation of difficulty with time for these machines undergoing feedback training (D has a virtually identical trajectory to A and is not shown). These machines were also open-loop trained at the L-level and H-level, together with some intermediate levels.

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

all conditions which corresponds to a few of the human operators. As discussed in Section 4. 1, machine *C* showed unstable learning which corresponded to a similar phenomenon in a few human subjects.

Thus, machine *D* was the most appropriate to form the basis of an "ensemble of identical machines" to evaluate the differential effects of training regimes, and it was used in experiment to determine the critical level of 6 at which learning just failed to take place. This was found to be at δ=0.3, and led to the selection of *δ=0.25* as the low level of difficulty for open-loop training in the experiments with human subjects. This influence of the experiments with learning machines on the design of the experiment with humans illustrates the practical value of having the machines as standard "dummy" subjects.

There is no obvious mechanism for causing the instructions given to human operators to affect the learning machines. It would be adequate for the purposes of this discussion to assume that the instructions would cause the machines to start with a useful level of control and to check whether learning takes place under the different training regimes with this initial control policy when it does not starting with a tabula rasa. However, some experiments were carried out with machine *D* using a simple mechanism for "verbal" communication with the ATLE learning machines.

The ATLE controller described in Section 3.3 was given the capability of accepting statements like, "When the position of the spot on the oscilloscope is *x* and *v*, then a sensible sign of control signal is *c*," and using them to adjust its control policy accordingly. The controller "imagines" the input-pattern it would receive resulting from *x* and *v*, considers that it has emitted the output c, and rewards itself for so doing.

This simple structure is readily extended to take account of non-quantitative specifications, "When the spot is on the far left moving fast to the right . . .," and other qualifiers, "It is *very* sensible . . ." The overall effect of a message is to modify the ATLE controller's policy, or, initially, to *prime* it with a control policy. In the context of the experiments with human operators, it was of interest to discover whether this priming through instructions would enable a controller previously unable to learn a suitable policy to establish an initial sub-environment in which it could do so.

The weight changes of the ATLE are equivalent to adding in the stimulus vector if it should cause a positive output, and subtracting it if the output should be negative. Hence, given the instruction, "If the spot is on the left, press the right-hand button," the ATLE would generate the stimulus vector, (-1,-1,-1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1), where the first seven components are positional information, the next seven components represent a lack of velocity information, and the last component is always set. The output required is positive, and hence this stimulus vector becomes the weight vector. The corresponding control policy involves only positional feedback and is sensible but not very effective.

A variety of sets of initial instructions were tried in this way, some of which gave rise to extremely powerful initial control policies and enabled machine *D* to learn the control task under open-loop training at *δ=0.5*. Thus, it is possible to model with the learning machines the effect of instructions in overcoming problems of learning perceptual-motor skills. One outcome of these experiments seemed to be that good sets of "instructions" could not be chosen in advance. In practice it was found that the instructions could be generated most effectively by giving one, examing its effect on performance, and then giving another—that is, "telling" was ineffective

(Lewis and Cook, 1969), but instructions based on feedback as to their effect could be used to control learning behaviour.

In summary, the experiments with the learning controllers were of benefit in the design of the experimental system for human operators. They also gave rise to an adequate set of patterns of learning behaviour to account for the human operators who could learn the task under F and L conditions not being able to do so under H conditions. The overall results suggest that, since similar patterns were shown by humans and machines, the results obtained derived from the epistemological problems posed by the tracking task not from any particular human peculiarities in learning it. The "fatigue" or "boredom" of machine C, and the effect of "instructions" on the learning of machine D, are of less weight, but illustrate possible extensions of the studies with learning systems to the modelling of other aspects of learning behaviour.

4.8 Implications of experimental results

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

The utility of feedback training is best examined by considering separately the groups under w and s conditions of instruction. With the non-informative instructions, w, the interaction between learning how to control the system and learning how the system operates, the dual control problem (Section 2.3), is predominant and the sub-environment phenomenon may be expected to strongly influence learning. This is strongly borne out by the experimental results in that the Fw group, under feedback training, shows overwhelmingly better performance at all test levels of difficulty, than either of the Hw and Lw groups under open-loop training. Thus, in a situation where the task is complex and poorly defined, and where interactions between performance and gaining knowledge may be expected, the experimental results clearly demonstrate the predicted advantages of feedback training.

With the informative instructions, s, the operator had the possibility of overcoming the sub-environment phenomenon by setting up a control policy "verbally," and initially taking a cognitive approach to the perceptual-motor tracking skill. This will only be possible if the level of performance required of him is not too high. Comparison of the results for the *as, Ls* and *Fs* groups gain shows a significant advantage to feed-back training, but now the L group is more similar to the F group than to the H group. This interaction between the effects of verbal instruction and the level of difficulty in training is, perhaps, the most important outcome of the present series of experiments. It not only provides experimental evidence of the meaningfulness and applicability of the approach to problems in learning advanced in Section 2, but is also

relevant to the practical instruction situation in flying and driving, where verbal instruction and variation in task difficulty are closely combined.

The demonstration of an interaction between verbal instruction and the various modes of training, and its relationship to the sub-environment phenomenon, is, in particular, a vindication of the approach taken to the study of learning and training by Pask (1960; 1961; 1964; 1965a; 1956; 1971), who has emphasized the importance of language in learning, and the linguistic nature of all processes in the learning hierarchy. The theoretical developments in Section 2 indicate that feedback training will be most effective when there are complex interactions between the "sub-skills" required for the learning of a particular task, and it is these interactions which are most amenable to description through language—thus, it is no coincidence that bqth feedback training and verbal instructions exert a profound influence on learning in the experimental situation chosen.

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

5. Summary

The theoretical discussion in Section 2 of this paper is based on the premise 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 primitive, 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 automation.

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 automation. 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, by 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-place, 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 basis 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 proves to be 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 non-adaptive human operators and automatic controllers. The feedback trainer used in this way forms the basis for a very sensitive *test* of control ability.

The operator in experiments on various modes of training on this control system were divided into three main groups. Two of these groups trained at fixed levels of difficulty, one fairly low, the other high, and the third group underwent feedback training. Each of these groups was further sub-divided into those who were informed about the nature of the task, and those who were not.

It was found that those trained under conditions of high difficulty showed little learning (although this level of difficulty was exceeded in their training by over sixty per cent of the feedback training group). The group who were trained under conditions of low difficulty split into those with informative instructions, who learnt, and those with non-informative instructions, who did not. The best performances, and the closest spread of learning, was shown by the feedback training group, who showed some effect of instructions, but otherwise learnt uniformly to a high standard.

Adaptive controllers used as trainees in the same experiments gave rise to the same phenomena. Those which were able to learn to a high level under feedback training did not learn under conditions of high difficulty, but were able to learn at low difficulty. The effect to instructions, in setting up an initial control policy, was to extend the range of conditions under which learning took place.

5.1 Conclusions

(i) It is possible to establish a theoretical basis for the study of learning and the investigation of training procedures which is independent of the learning device and the nature of the task to be learned. The foundations for this theory lie in automatic control and general systems theory, neither of which is complete nor sufficiently far advanced to solve many problems of great practical interest. Indeed the human operator is the only physically realized adaptive controller of any complexity available as yet, and, hence, is of great interest to the control engineer. Advances in the understanding of human learning may, therefore, be expected to be closely allied to developments in control engineering and general systems theory.

(ii) At present it is possible to set up models of learning behaviour in the human operator by appropriate adaptive controllers subject to the same physical (physiological) constraints as the human. As models of learning, these have the same status as linear describing functions for modelling control policies - they are well-known and give a good account of certain overall phenomena, without accounting for the fine structure of behaviour or relating to known physiological mechanisms. As more powerful learning artifacts are developed it may be expected that they will have increasing use as models of human learning in a wider variety of situations.

(iii) The experiments demonstrate the utility of feedback training in the teaching of a novel perceptual-motor skill, and the theory suggests that it will be generally applicable. Further experiments are required to elucidate the applicability of feedback training in other situations, particularly those of immediate practical interest. The influence of direct communication on the learning of a perceptual-motor skill was also apparent, and future experiments should make great use of, and exert greater control over, this variable. The present system is being extended by incorporation of a fully-branching audio/visual teaching machine, which will not only be capable of administering the questionnaires, but also of giving verbal instruction to the trainee at the discretion of the automatic trainer. Advice might then be given when, for example, the trainer finds that the level of difficulty has not risen to a prescribed value after a given time.

The integration of feedback training procedures, based on the variation of difficulty, with direct instruction through natural language, has a tremendous potential, not only in teaching perceptual-motor skills, but also in training the operators and service-engineers of specialist equipment, such as communication, navigation, and computer systems. The linking of a feedback training simulator to an audio/visual teaching machine is within the scope of modern computer-based

technology, and may prove to have better cost/effectiveness than conventional applications of teaching-machines in education.

Acknowledgements

References

Andreae, J. H., Cashin, P. (1969) "A Learning Machine with Monologue," *Int. J. Man-Mach. Studies 1.* (2) pp. I—20.

Bekey, G. A. (1965) "Description of the Human Operator in a Control System," in *Modern Control Systems Theory,* New York: McGraw Hill.

Bigland, B., Lippold, O. C. J. (1954) "The Relation Between Force, Velocity and Integrated Electrical Activity in Human Muscles," J. *Physiol.* 123(1) 214—224.

Blakelock, J. H. (1965) *Automatic Control of Aircraft and Missiles,* New York: Wiley.

Feldbaum, A. A. (1963) "Dual Control Theory Problems," *Proc. 2nd Int. Congr. IFAC,* Basle.

Gaines, B. R. (1965) "Automated Feeback Trainers for Perceptual-Motor Skills," Techn. Rep. I, Ministry of Defence Contract, Cambridge, UK.

Gaines, B. R. (1966) "Teaching Machines for Perceptual-Motor Skills," Proc. Programmed Learning Conf., Loughborough, in *Aspects of Educational Technology,* London: Methuen 1967.

Gaines, B. R. (1967) "Automated Feedback Trainers for Perceptual-Motor Skills," Final Report Ministry of Defence Contract.

Gaines, B. R. (1968a) "Training the Human Adaptive Controller," *Proc. IEE.* 115 (8) pp. 1183-9.

Gaines, B. R. (1968b) "Adaptive Control Theory," in *Encyclopaedia of Information, Linguistics and Control,* Oxford: Pergamon.

Gaines, B. R. (1969) "Linear and Nonlinear Models of the Human Controller," *Int. J. Man Mach. Studies* 1 (4) 333-360.

Gaines, B. R. (1972) "Axioms for Adaptive Behaviour," *Int. J. Man Mach. Studies* 4(2) 169-199.

Gaines, B. R., Andreae, J. H. (1966) "A Learning Machine in the Context of the General Control Problem," *Proe. 3rd Int. IFAC Congress,* London.

Gibbs, C. B. (1951) "Transfer of Training and Skill Assumptions in Tracking," *Quart. J. Exptl. Psychos.* 3 99—*110*.

Hall, 1. A. M. (1963) "Study of the Human Pilot as a Servo Element," *J. Roy. Aero. Soc.* 67351-360.

Hudson, E. M. (1964) "Adaptive Training and Non-Verbal Behaviour," NATRADEVCEN TR 1395—1.

Jex, H. R., McDonnell, J. D., Phatak, A. V. (1966) "A Critical Tracking Task for Manual Control Research," *IEEE Trans.* HFE-7 138—145.

Kelley, C. R. (1962) "Self-Adjusting Vehicle Simulators," IRE Int. Congr. Human Factors in Electronics, Long Beach, Calif.

Kelley, C. R., Kelley, E. J. (1970) "A Manual for Adaptive Techniques," Nonr-4986 (00) NR 196—050, Dunlap and Associates.

Lange, G. W. (1965) "Syntheses of a Model of the Human Operator Engaged in a Tracking Task," PhD Thesis, University of London.

Lewis, B. N., Cook, J. A. (1969) "Toward a Theory of Telling," *Int. J. Man-Machine Studies,* 1 (2). 129—176.

Licklider, J. C. R. (1960) "Quasi-Linear Models in the Study of Manual Tracking," in *Developments in Mathematical Psychology,* Glencoe: Free Press.

Lowes, A. L., Ellis, N. C., Norman, D. A., Matheny, W. G. (1968) "Improving Piloting Skills in Turbulent Air Using a Self-Adaptive Technique for a Digital Operational Flight Trainer," NAVTRADEVCEN 67-C-0034-2, Fort Worth, Texas.

Luria, A. (1961) *The Role of Speech in the Regulation of Normal and Abnormal Behaviour,* Oxford: Pergamon Press.

MacDonald Ross, M. (1969) "Programmed Learning —A Decade of Development," *Int. J. Man-Machine Studies,* 1 (1) 73—100.

McRuer, D. T., Graham, D., Krendel, E. S., Reisener, W. (1965) "Human Pilot Dynamics in Compensatory Systems," AFFDL-TR-65-15.

Mesarovic, M. D. (1960) *The Control of Multivariable Systems,* Cambridge: M.I.T. Press.

Newell, A., Shaw, J. C., Simon, H. A. (1959) "Report on the General Problem Solving Problem," *Proc. Int. Conf. Inf. Processing,* Paris.

Pask, G. (1960) "The Teaching Machine as a Control Mechanism," *Trans. Soc. Instr. Tech.* 12 72 82.

Pask, G. (1961)An *Approach to Cybernetics,* London: Hutchison.

Pask, G. (1964) "Adaptive Teaching Machines," in *Teaching Machines,* Oxford: Pergamon.

Pask, G. (1965) "Teaching as a Control Process," *Control.*

Pask, G. (1965) "Comments on the Cybernetics of Ethical, Sociological and Psychological Systems," *in Progress in Biocyberneties* 3, Amsterdam: Elsevier.

Pask, G. (1971) "A Cybernetic Experimental Method and Its Underlying Philosophy," *Int. J. Man Maeh. Studies* 3 (4) 279—337.

Preyss, A. E., Meiry, J. L. (1968) "Stochastic Modelling of Human Learning Behaviour," *IEEE Trans*. MMS-9 36—46.

Weir, D. H., Phatak, A. V. (1967) "Model of Human Operator Dynamic Response to Step Transitions in Controlled Element Dynamics," NASA CR-671.

Weizenbaum, J. (1966) "ELIZA—A Computer Program for the Study of Natural Language Communication Between Man and Machine," Comm. ACM 9 36—45.

Widrow, B., Smith, F. W. (1964) "Pattern-Recognizing Control Systems," in *Compute; and Information Sciences*, Spartan.

Widrow, B. (1966) "Bootstrap Learning in Threshold Logic Systems," *Proc. 3rd Int. Congr. IFAC,* London.

Winograd, T. (1971) "Procedures as a Representation for Data in a Computer Program for Understanding Natural Language" MAC TR-84, M.I.T., Project MAC.

Young, L. R., Green, D. M., Elkind, J. I., Kelly, J. A. (1964) "Adaptive Dynamic Response Characteristics of the Human Operator in Simple Manual Control," *IEEE Trans*. HFE-5 6—13.

Young, L. R., Stark, L. (1965) "Biological Control Systems—A Critical Review and Evaluation," NASA CR-190.

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.

Appendix

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.

[Alternative i—Weak—Non-informative]

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.

[Alternative ii—Strong—Informative]

The red push-buttons on the arms of your chair are to be used as controls. Depressing either 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.

TABLE I: Weak and strong forms of instruction

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

TABLE II: Number of operators in experimental groups

Activity	Duration	Description
Read Instructions	Variable (7 — 18 min.)	Instructions reproduced in Table I — weak and strong forms.
Train₁	20 min.	Track under one of three conditions, H, L, F; the feedback group started with $\delta=0.00$.
Test₁ $\delta=0.5(H)$	5 min.	Continue tracking without interruption and without knowledge of change, but with $\delta=0.5$, and performance measured.
Questionnaire₁	Variable (18 — 53 min.)	Fill in Questionnaire shown in Figs. 11 and 12; this involves reading instructions again.
Train₂	20 min.	Track under same conditions as Train₁; the feedback group started with δ at the same level as it was at end of Train₁.
Test₂ $\delta=0.5(H)$	5 min.	As for Test₁.
Questionnaire₂	Variable (12 — 31 min.)	Fill in Questionnaire; informed that there will be three tests, each of five minutes.
Test₃ $\delta=0.5(H)$	5 min.	Measure performance over last 4 minutes of test; this is at same level of difficulty as previous, unannounced test.
Test₄ $\delta=0.25(L)$	5 min.	Same as Test₃. but at level of difficulty as that at which L group trained.
Test₅ $\delta=0.70(V)$	5 min.	Same as Test₃ but at higher level of difficulty than that met by any operators, except a few of F group, during training.

TABLE III: Experimental procedure