Open access • Journal Article • DOI:10.1049/PIEE.1968.0209

Training the human adaptive controller — Source link

Brian R. Gaines

Institutions: University of Essex

Published on: 01 Aug 1968

Topics: Training system, Stability (learning theory), Control theory and Learning environment

PROCEEDINGS

THE INSTITUTION OF ELECTRICAL ENGINEERS

Volume 115 **Control & Science**

Training the human adaptive controller

B. R. Gaines, M.A.

Synopsis

The training of human operators for skilled tasks may be regarded as the synthesis of a specific controller from a general-purpose adaptive device, by influencing its adaption through selection and variation of the learning environment. Selection of environments to maximise the rate of learning is itself a control problem, and an automatic feedback training system is proposed which feeds back information about the operator's performance to control the parameters of his environment. The stability and performance of the trainer have been investigated both theoretically and experimentally, and its utility has been tested in a fairly realistic training situation using as trainees both human operators and computer-simulated learning machines.

1 Introduction

The designer of a system containing human controllers has a synthesis problem which, although similar in some respects to the classical problems of automatic-controller synthesis, differs from them in that the human controller already exists as a complete entity and is highly adaptive. Fabrication of an arbitrary controller is thus not possible, and even parameter setting in any conventional sense is difficult. Techniques for the 'synthesis' of the human part of the control system have therefore developed differently to those of classical synthesis, and are based largely on the adaptive capabilities of the human operator.

If advanced adaptive controllers, such as learning machines, became available as general-purpose adaptive devices, these techniques may eventually be applicable to automatic-controller synthesis.[1] For the present, however, it is of interest, in terms of both human engineering and control theory, to treat utilisation of the human operator in control systems as a synthesis problem in itself.

Although the human controller is highly adaptive, its repertoire of possible control policies is not exhaustive, and some families of control policy are far more readily implemented than others. Whether genetically or by experience in its natural environment, the human operator comes to a novel control problem with many limitations on its inputs and outputs and preconceptions about the nature of its environment. To the extent that appropriate coding of information from, and actions upon, its environment can satisfy these preconceptions and call for performance within these limitations, the synthesis problem will be simplified if not solved. Ergonomics is concerned with establishing these limitations and preconceptions, and much is known about good equipment practice in integrating the human operator into a control system. The classical work of Birmingham and Taylor on visual/manual tracking, suggesting that the easiest policy for the operator to implement is a single gain, is an example of this approach to synthetising systems containing the human operator.[2]

The human controller does have the equivalent of direct parameter adjustment through the use of language. However, the manipulation of the parameters of a control policy by verbal direction is very imprecise and not well understood. Luria has studied the control of motor actions by verbal directives in both young children and brain-damaged patients, and demonstrated that the coupling between verbalisation and performance of a skilled task can be very strong.[3] One of the major problems in establishing human control policies through the use of language is that the engineer's conception of what variables define a control function may bear little or no relationship to those variables in the human controller which may be affected linguistically. The experienced operator may make some effort to obey instructions to increase his gain, respond early to fast-changing inputs (increased lead), or take no notice of high-frequency ripple (dominant lag), but there is great interaction between the effects generated by such instructions, and the most sensible prescription may only worsen performance.

Experiments with the equivalent of verbal communication with learning machines make it obvious that language is far more imprecise than is realised. An instruction to increase gain, for example, may mean: 'I assume you are a 3-term controller with proportional, derivative and integral weighting—increase the magnitude of the proportional weighting whilst keeping the others constant', or 'I assume you are a 2-term controller with an overall weighting plus a relative weighting to an integral term—increase the overall weighting keeping the relative weight of the integral term constant', or one of many other possibilities. The human controller is essentially discontinuous and nonlinear and must take its choice as to what is meant and what to do when told to increase gain; even if there is some change in control policy which corresponds to increase of gain in some sense, the side effects may far outweigh the desired one.

Language, for all its defects, is an important tool in the synthesis of human control policies, especially in stating performance criteria, goals and subgoals, which may form the basis for learning. The general effect of communication with the control may be described as 'priming' it with information of some form to prepare it for its control task.

Coding and priming have strong similarities to the techniques of synthesis for classical automatic controllers, but there is a third major technique which has no equivalent, and that is training. The human operator is an adaptive controller and, put in a novel environment and given a performance criterion, it will tend to optimise its control policy. Hence the designer may take the attitude to the human

Paper 5608 C, first received 27th February and in revised form 14th May 1968

Mr. Gaines is with the Department of Electrical Engineering Science, University of Essex, Colchester, Essex, England

elements in the control system that, provided what is required of them is possible and they know what they are required to achieve, no specific synthesis action on his part is necessary because they will adapt themselves to become satisfactory.

In practice, the designer may have to carry his synthesis further than this because of two defects inherent in all adaptive controllers—even though a satisfactory control policy is possible for them, they may never attain it and, even if they do, it may take an inordinate time. Automatic adaptive controllers which do not converge, or take far too long in so doing, are rejected at the experimental design stage and never appear in the literature. Human adaptive controllers may also be rejected by means of selection procedures, but for many tasks the designer cannot afford to be selective and for others even the best human controllers may be in difficulty.

Some of the problems caused by overlong transients in adaption, such as jeopardising the controlled system, may be overcome by allowing adaption in a simulated system. This reduces the problems to one of economics, since realistic simulators tend to be expensive and the cost per learning hour high. Since all that is required is that the operator should attain a satisfactory control policy as rapidly as possible (and that it should be stable, i.e. not deteriorate with time or adverse influences), there is no requirement that learning should take place in an environment identical to that to be finally controlled. It may well be that a difficult control task is best learnt in a series of learning sessions starting with a simplified environment and progressively increasing the difficulty of control. Such a process of decreasing the learning transient by varying the environment in which learning takes place is called 'training'.

This paper describes an approach to the training of the human operator for a control skill in which the induction of a satisfactory and stable level of performance in the shortest time is itself treated as a control problem suitable for solution by an automatic controller, i.e. an automated feedback trainer. The structure of a particular form of trainer is described, and its performance, particularly stability, is examined both theoretically and experimentally. Finally, some experiments with human operators are used to demonstrate that the phenomena of adaption defined in theory do occur in practice, and that the feedback trainer is a viable means of decreasing the learning transient.

2 A rationale of training

Consider a control situation, or environment, into which the human controller is to be 'connected' and is then required to perform satisfactorily. Initially its control policy may be such that its performance is not satisfactory but it is expected to become so by adaption through experience.[4,5] However, even if the controller has the capacity for adaption to the environment, i.e. a control policy or a sequence of control policies which are satisfactory and remain so irrespective of further adaption, there is no guarantee that it will succeed, if at all, in a reasonable time. If adaption takes too long or does not occur on interaction with the required environment, then training may be used to shorten the learning phase.

For training to be possible, the required environment must be embedded in a family of environments which have certain topological constraints on them. Training will then consist of connecting the controller to a sequence of varying environments directed in some way by their natural topology. Consider, for example, a 2-parameter family of control situations, such as regulation of a plant whose transfer function contains a pair of complex poles of variable undamped natural frequency and damping ratio. Any member of this family may be represented as a point in a plane with these co-ordinates, and any control policy for the plant splits the plane into two regions, one of stability and the other of instability (Fig. 1). The required environment, or control situation for which training is to be given, may be represented as a point in this plane P_{FIN}, which is outside the region of stability S_O of the controller's initial control policy.

It is a plausible assumption that adaption (the extension of the region S_O to encompass P_{FIN}) will be slow if the controller is placed immediately in the control situation corresponding to P_{FIN}. For, consider the trajectory induced by an unstable control policy in the state-space of the plant. By definition, this trajectory is mainly outside the desired

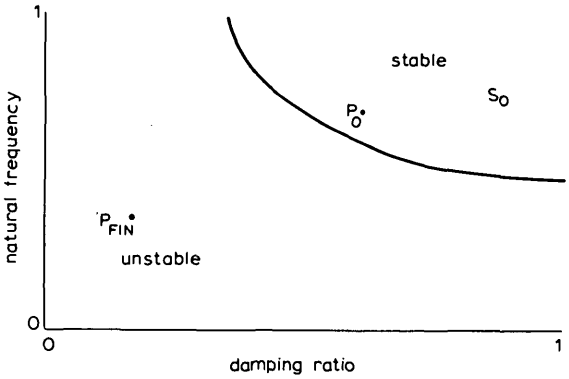

Fig. 1
Stability of control system as a function of plant parameters

region of the state-space, and the experience of the controller is generated by the behaviour of the plant in regions of the state-space which are practically irrelevant to those experienced with a stable control policy. Open-loop adaption of the controller may fail completely because it is based on premises such as linearity which do not hold in these regions, and closed-loop adaption may be impossible because the performance measure shows little or anomalous variation under conditions of instability. Even if learning does occur, the controller is experiencing conditions quite unlike those induced by its ultimate control policy. Hence the transient phase of learning will be not only long but also largely wasted.

The difficulties in learning created by the instability of naïve control policies are more general in the *dual-control* problem, in which a controller has to exert some control over its environment, at the same time as it learns about it in order to improve its control policy.[6] Any given policy will generate some subenvironment, in that it restricts the states and state transitions of the environment to some subset of the total possible behaviour. The subenvironment generated by the initial policy of a naïve controller may not even intersect that generated by an optimum or satisfactory policy, and learning in the initial subenvironment may then be irrelevant or even deleterious to performance in the desired subenvironment. Alternatively, and especially if the controller deliberately adapts a 'search' policy, the initial subenvironment may be so extensive that learning in it takes an unacceptably long time.

2.1 A training strategy

The difficulties created by the dual-control problem could be overcome if it were possible to maintain a subenvironment similar to the final one throughout learning, no matter what control policy was being implemented. The question of 'similarity' between subenvironments is a difficult one in general and restricts the application of the proposed training strategy in environments where no continuous deformation of structure is possible. However, in most control environments there is a natural topology which enables similarity to be defined. For example, in the second-order system considered previously, small variations in natural frequency or damping ratio cause small changes in the 'feel' of the plant.

Consider the family of environments shown in Fig. 1, with the controller now given a situation P_O within its initial stability boundary S_O. The interaction is stable and unusual behaviour in the plant is not excited, so that adaption takes place and the stability boundary moves to encompass more of the space. It is reasonable to assume that it will not move out indefinitely, since the controller is not gaining experience of conditions far removed from P_O, so that eventually it reaches the position shown in Fig. 2, encompassing the region S_1. The point P_I is now within the stability boundary and the controller may be put into a situation corresponding to this value of the plant parameters. It will once more have a stable interaction and the boundary will move out to encompass the region S_2. The series of training environments, P_O, P_1, P_2 and so on, may be continued until the point P_{FIN}

is within the stability boundary and the learning is deemed satisfactory for this situation. It may be possible to increase the speed of learning for P_{FIN} by overshooting the training sequence to P_K, for instance.

Since, in practice, the stability boundary is defined by some prescribed level of an error functional being exceeded, it is reasonable to suppose that measurement of the controller's performance might be made the basis for adjusting the environment so as to maintain it within, but near, the controller's stability boundary. Thus, after being connected to some environment and the mean, or mean-square, error has been measured over some period, the controller may be given another environment which is more, or less, difficult to control according to whether the error functional is less or greater than some prescribed level. If the parameter of difficulty is continuous, such as damping ratio or disturbance amplitude, then this adjustment might be made continuously —the difficulty is reduced if the error is higher than a certain level, and increased if less. In more general terms, the strategy is one of making the rate of change of difficulty positive when the controller has brought the system within the desired subenvironment, and negative otherwise.

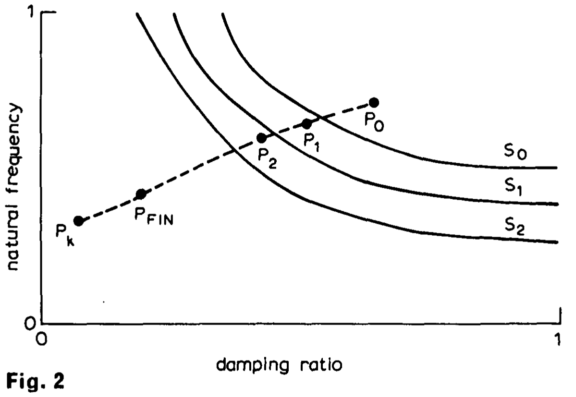

Fig. 2
Generation of a series of parameters for training

Fig. 3 shows a block diagram of this type of training system, i.e. an automated feedback trainer, for a continuous control task. In the main control loop the controller has some form of input to the environment and receives an indication of its error. In the auxiliary training loop, a summing integrator subtracts a constant from the modulus of the error and integrates the result. Its output is fed back to the parameters of the environment, in order to increase the difficulty of the control task if the error is less than the constant, and decrease it otherwise.

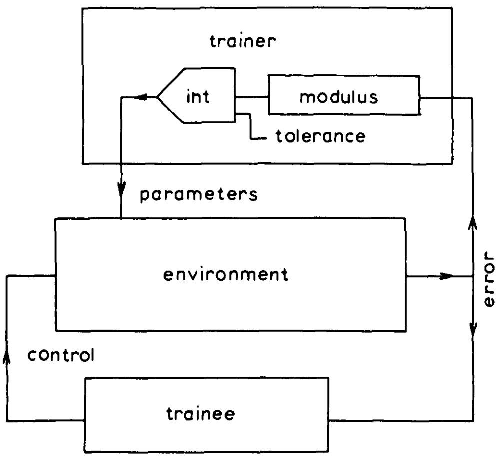

Fig. 3
Feedback training system

2.2 Stability of the feedback trainer

Although clearly the only stable value of the integrator's output in Fig. 3 corresponds to an environment of such difficulty that the mean error of the controller is equal to the prescribed constant, it is by no means obvious that the outer loop is stable and that the output is finally near this value. Since the overall system is nonlinear and the characteristics of both controller and environment are assumed to be imperfectly known, no exact stability analysis is possible. With reasonable assumptions about these, however, it is possible to linearise the outer loop, analyse its stability, and check it empirically.

If the environment is linear (three integrators in cascade, for example) and the controller is nonlinear (say, a simple relay controller), then a limit cycle will be established which determines the mean error modulus. When the parameters of the environment are varied, for a given control policy the size of the limit cycle will also vary, and may be supposed to increase with the difficulty of control. This variation will not be instantaneous, and the envelope of the error modulus will change with certain time constant. If the controller is linear, however, there will be no limit cycle set up, and if there are no disturbances in the system, the error modulus will, exponentially, rise if the control loop is unstable and fall if stable.

In both cases the behaviour of the error modulus can be approximated by the relationship

$$a(M - M_0) + b^2 sM = D - A \qquad (1)$$

where M is the smoothed error modulus, M_0 is a constant to account for M being greater than zero, D is the difficulty of the task, and A is the ability of the operator. The constant a will be large and b small for switching mode controllers, while a will be a function of the disturbance and b large for linear controllers.

The training feedback is of the form

$$sD = -c^2(M - M_0) \qquad (2)$$

so that, combining the two equations,

$$D + \frac{a}{c^2} sD + \frac{b^2}{c^2} s^2 D = A \qquad (3)$$

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

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

In the experimental situation, b was small in its effect compared with a, and eqn. 3 then reduces to

$$D + \frac{a}{c^2} sD = A \qquad (4)$$

so that again D follows A, but this time through a simple exponential lag of time constant, a/c^2.

These results indicate that a feedback trainer of the form in Fig. 3 should be stable and follow changes in the operator's ability without introducing transient artifacts of its own. The analysis is obviously only qualitative, however, and requires experimental confirmation. In the following Sections, a feedback trainer for a particular environment is described, together with some experiments on its behaviour coupled to adaptive/nonadaptive automatic controllers and human operators.

3 Experiments with an automated feedback trainer

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[7] or submarine, was selected as being both realistic and of fundamental interest in manual and automatic control.

Preliminary experiments and comparison with aircraft dynamics indicated that a second-order stable transfer function, with undamped natural frequency in the range between

0 and 8 Hz, and a damping ratio in the range between 0 and 1, was most suitable. However, the human operator is capable of compensating such a system fairly easily, and the dynamics were increased to third order by addition of a rate control. The overall transfer function was thus of the form

$$F(s) = 1/s(a^2s^2 + 2Kabs + b^2) \quad . \quad . \quad . \quad . \quad . \quad (5)$$

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

The operator was provided with an input to the transfer function above by means of a manual control, and a second, repetitive, 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. A block diagram of the complete training system is shown in Fig. 4.

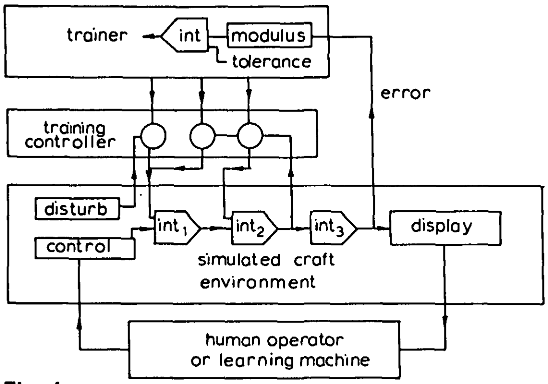

Fig. 4
A feedback trainer for a high-order tracking skill

3.1 Behaviour with fixed controllers

The difficulty in controlling the environment of Fig. 4 increases as the damping ratio decreases; the undamped natural frequency decreases; and the disturbance amplitude increases, i.e. within the range of values used in the experiments. Fig. 5 shows the variation in undamped natural frequency with time when this is varied as the parameter of difficulty D (eqn. 1), the damping ratio and disturbance having fixed values.

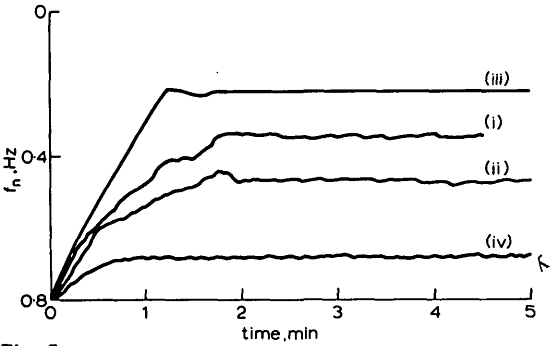

Fig. 5
Dynamic behaviour of trainer with fixed controllers

The trajectories (i) and (ii) were generated with human operators, while (iii) and (iv) were generated by simple relay controllers, for different slopes of switching line in the position/velocity phase plane. In all cases the response is of the form predicted from eqn. 4—initially the difficulty is so low that the error generated is small, and the difficulty is forced to rise asymptotically to a level at which the mean error is at a prescribed level. In practice, the initial rise is rate-limited because M is essentially greater than zero, and an exponential climb is seen only at the turnover. The slope at turnover varies with the asymptotic value of D, since the constant a tends to increase with A.

Even with human operators, there is no noticeable effect of learning during the few minutes of the experiment, and the asymptotic value of natural frequency may be taken as a measure of the quality of the control policy implemented; it defines a point on the stability boundary for the controller in the natural-frequency/damping-ratio plane. If this value is measured for a number of values of the damping ratio, the complete stability boundary may be mapped out. Fig. 6 shows such boundaries measured for the three human operators (i), (ii) and (iii) and two relay controllers (iv) and (v). These have the same form as those obtained in the study of human reactions to aircraft dynamics, the linearity or nonlinearity of the control policy adopted, and the subjective feel of the simulated vehicle.[8]

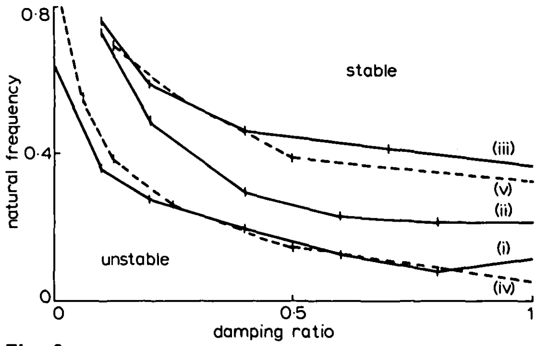

Fig. 6
Stability boundaries of fixed controllers measured with trainer

When the controller attached to the feedback trainer is adaptive and improves its control strategy with experience, the asymptotic value of difficulty, if it exists, will not be reached as rapidly as those of Fig. 5. Instead, the difficulty will rise rapidly until the mean-error modulus is at the prescribed level, and will thereafter vary slowly to follow any changes in the controller's ability. Most importantly, it will maintain the desired subenvironment while all the time converging towards the final environment in which the controller is required to perform satisfactorily. While it is plausible that this is a useful training technique, its advantages can only be demonstrated by experiment. The following Section briefly reviews experiments on feedback training and outlines the results of some studies with the system of Fig. 4.

3.2 Previous work on feedback training

Documented research on automated feedback trainers for tracking skills has been very slight, comprising one major theoretical study, one major experimental study and a number of minor apparatus studies. The earliest mention of the possibility appears to be of Stockbridge and Siddall in 1956,[9] who suggested the use of a guided-weapons-tracking trainer in which 'the difficulty of the task is proportional to the success of the subject'. In a number of papers Pask[10–14] has made available a very deep and comprehensive discussion of automated training, and has placed it in the general context of interactions between self-organising systems. The only major experiment reported to date is that of Hudson,[15] who trained over 70 subjects for ten hours each on a third-order two-dimensional tracking task.

A short exposition by Senders[16] of the principles of 'adaptive teaching machines' is a good example of the many studies of feedback training which have been reported informally. Hudson's devices centre has proposed feedback-training systems; Ziegler, Birmingham and Chernikoff[17] have described a 'teaching machine for the selection and training of operators of higher-order vehicles' which removes 'quickening' as the operator's mean-error modulus decreases; Chernikoff's report[18] on this machine and the ensuing discussion are particularly interesting. Bowen, Hale and Kelly[19] have described a 'general vehicular research tool'; Briggs[20] has described experiments on scheduling augmented feedback according to the operator's performance which might be automated in feedback training; Kelley[21] has utilised an adaptive system similar in principle to that described here as a 'secondary loading task'.

Since the level of difficulty of the control task is varied according to the operator's ability, a feedback trainer may be used to test this ability and measure it in terms of highest tolerable task difficulty. Jex, McDonnell and Phatak[22] have carried out a comprehensive programme of research for NASA on the use of control systems with varying dynamics to measure some parameters of the human-operator describing function. This is the only published work which discusses the viability of different feedback loops, from error functionals to task dynamics.

The obvious empirical test of the dynamics and stability of a feedback training system is to use as experimental subjects simulated operators, i.e. automatic controllers, of different but non-time-varying 'ability'. This does not appear to have been done in any previous work on feedback training, although Taylor and Birmingham[23] have plotted the 'learning' curve for a variable-gain element on various control systems in order to deride too close an analysis of the shape of learning curves.

Hudson's feedback training technique was not successful in operation because it related the mean error to the absolute dynamics of the controlled element. He suggests that a loop be mechanised instead to 'keep the error level at some desired point', and the training loop shown in Fig. 3, described by eqn. 2, may provide a means of doing this. Although his automatic adjustment of controlled-element dynamics was not effective, he was able to maintain the operator's performance at a fixed level by manual adjustment of the plant dynamics. Comparison of operators trained under this manually effected feedback training régime at two levels of performance, with operators who train on the required system throughout (a third-order transfer function), shows quite clearly that there is an optimum level of difficulty for the operator, defined in terms of his performance, which induces the most rapid learning. Training at levels either below or above this gives a slower rate of learning. A feature of particular interest in Hudson's experiments is his use of a variety of plant parameters for maintaining the difficulty constant for the subject, but these variations seemed to have little effect on the overall result.

3.3 Experimental evaluation of feedback training[24]

Since the literature did not provide clear evidence of the viability and utility of automated feedback training, formal experiments were carried out with the system of Fig. 4 to evaluate feedback training in a fairly realistic situation. While the results of Figs. 5 and 6 were obtained using a conventional joystick control, this was replaced for training purposes by a pair of pushbuttons which gave opposite impulses to the system but reversed their sense at each push. These controls not only eliminated operator fatigue, so that long training runs could be used, but added complexity to the control task so that extensive learning was required of all operators.

The major experimental variable was the difficulty, and variation in difficulty, of the environment. Given the requirement to train operators to control some novel environment which is variable in difficulty, there are three main strategies of interest:

(*a*) *Fixed training*, in which the operator is allowed to control the desired environment immediately

(*b*) *Open-loop training*, in which the operator is given some sequence of environments, graded in difficulty, to control, but without reference to his varying ability

(*c*) *Feedback training*, in which the difficulty of the environment is continuously varied according to the operator's ability

The form of open-loop training investigated was one in which the operator was trained at one level of difficulty and tested at another. By testing at a number of different levels, it was possible to use the same results as an evaluation of fixed training.

Another variable investigated was the effect of verbal instruction on learning and its interaction with the training procedure. Half the operators were given informative, or *strong*, instructions which helped them understand the nature of the controlled system, while the others received only the performance criterion, or *weak* instructions.

In order to extend the generality of the results and demonstrate their independence of peculiarly human factors in learning, adaptive controllers were also used in the same series of experiments. These consisted simply of a perceptron-like[25] adaptive-threshold-logic unit receiving a binary pattern generated by the position and velocity of the error, and giving a binary output corresponding to pushing one or other of the buttons; an error functional was utilised to give reward/punishment-performance feedback.

Fig. 7 shows the variation in difficulty with time for three human operators undergoing feedback training. A level of

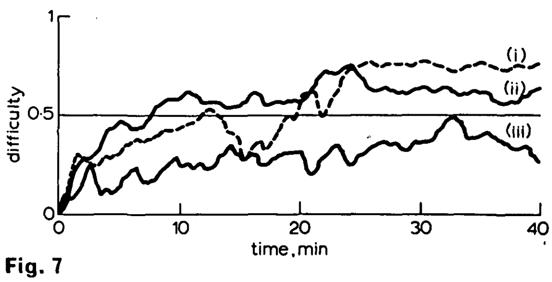

Fig. 7
Dynamic behaviour of trainer with human operators

difficulty, $D = 0{\cdot}5$, was taken to be the required level, and it can be seen that operators (i) and (ii) exceed this fairly quickly, while (iii) never attains it. The results with adaptive controllers shown in Fig. 8 are similar, but (iii) (broken line)

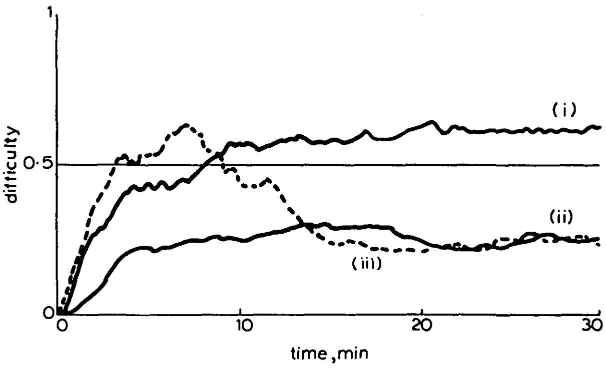

Fig. 8
Dynamic behaviour of trainer with learning machines

is of particular interest since it shows instability of adaption. This was also noted in a few human operators trained under nonfeedback conditions.

The details of experimental procedure and results are available elsewhere[24] and will not be described in this paper. However, a summary is given in the following Section.

3.3.1 Experimental procedure

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

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

The operators were 72 RAF pilots at an advanced stage of selection and training. They were chosen as a homogeneous group suitable for comparative studies of training techniques.

The main results of the experiments are as follows:

3.3.2 Effect of training conditions

The operators trained at a high level of difficulty show little or no learning and perform badly on all tests. The strong instructions have a significant effect in improving learning, but do not enable the operator to overcome basic difficulties. The high level of difficulty is not in itself unattainable, however, since over 65% of the feedback group reached it, or a much higher level, during their training, and hence could perform well at the high level.

The operators trained at the low level of difficulty split clearly according to the instructions given—those with weak instructions do not show appreciably better performance than those trained at a high level, while those with strong instructions show a spread in performance from very good to very poor throughout the tests.

The operators trained under feedback conditions all learn to a high standard. Those with weak instructions do not differ significantly in their performance from the group which trained at a low level with strong instructions. The feedback group with strong instructions is significantly better than all other groups.

3.3.3 Effect of instructions and verbalisation

The overall effect of informative/noninformative instructions is that strong instructions give significantly improved performance in all groups. Informative instructions are shown by the results to be capable of compensating for poor training conditions, provided they are not too poor.

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

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

3.3.4 Transfer and feedback

In terms of 'transfer of training'[26] the results demonstrate clearly that the question as to whether easy/difficult or difficult/easy transfer is best is not meaningful. They indicate, however, that transfer from a difficult level to an easy one is better than learning solely on the easy one, provided that the difficult level is within the operator's ability to perform reasonably well.

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

3.3.5 Results with learning machines

The results with computer-simulated learning machines parallel those with human operators completely. 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. More specifically, there were individual machines which learnt under feedback conditions to a high stable level but were completely unable to learn at a high level of difficulty. Among these machines were some that learnt at a low level of difficulty and others that did not. In this last event, suitable instructions—whose final effect was to give the machine a reasonable initial policy—enabled learning to take place. This demonstrates not only the utility of these devices as dummy operators but also the generality of the learning and training phenomena investigated.

4 Summary and conclusions

The training of human operators for skilled tasks has been treated as a technique for synthetising a specific controller from a general-purpose adaptive device by influencing its adaption through selection and variation of the learning environment. Selection of environments to maximise the rate of learning is itself a control problem in the space of states of the adaptive device, and, by consideration of the circumstances which would make learning generally difficult, it is possible to propose a system for solving this problem automatically.

The stability and performance of one such system, an automated feedback trainer based on performance feedback, has been investigated both theoretically and experimentally and shown to be satisfactory. Further experiments on the utility of this system, comparing feedback training with fixed training in a single environment and open-loop training in a sequence of environments unrelated to performance, demonstrate the superiority of feedback training in at least one fairly realistic training situation. That the same results are obtained using simple learning machines as trainees demonstrates that the phenomena investigated are fundamental to the learning situation, and that the training technique is, as required, insensitive to the type of trainee. The strong interaction between the effects of the form of verbal instructions and the technique of training indicates that linguistic variables cannot be neglected in the study of training. It would be advantageous if the actual instructions given could also be adaptively controlled by the feedback training system.

The results have implications both for the structure of future automatic control systems and for human teaching and training by machine. Priming, coding and training as techniques for making the best use of a general-purpose adaptive device may well become standard synthesis techniques for automatic control systems. However, at present, automatic learning systems are so few that there is little scope for experiment with these techniques and little impetus to theoretical studies of them. The human adaptive controller is the one outstanding example of a learning system which is viable, exists and has important applications. Thus the study of human learning phenomena offers a proving ground for techniques ultimately to be applied to automatic controllers.

On the other hand, machine-aided instruction and adaptive training both offer the possibility of improvements and economies in the educational and training processes; yet there is little foundation either in theory or in a common terminology for the study of these processes. While the learning systems involved are multidimensional, nonlinear and of discrete action, topics about which control theory has as yet little to say, the present study shows that a highly simplified analysis can be reasonably accurate and very useful. Since these topics are also the subject of much current research, probably analysis in far greater depth will become equally simple in time. For the present, clearly, automatic control and human studies have much to gain in mutual development from a common viewpoint and terminology.

5 Acknowledgments

The experiments on human operator training were carried out under contract to the Ministry of Defence at the Department of Experimental Psychology, University of Cambridge.

6 References

1 GAINES, B. R., and QUARMBY, D. J.: 'The convergence of adaptive pattern classifiers in control systems', Department of Electrical Engineering Science, University of Essex, RM 68/2

2 BIRMINGHAM, H. P., and TAYLOR, F. V.: 'A design philosophy for man-machine control systems', *Proc. Inst. Radio Engrs*, 1954, **42**, pp. 1748–1758

3 LURIA, A.: 'The role of speech in the regulation of normal and abnormal behaviour' (Pergamon, 1961)

4 ZADEH, L. A.: 'On the definition of adaptivity', *Proc. Inst. Elect. Electronics Engrs*, 1963, **51**, pp. 469–470

5 GAINES, B. R.: 'Adaptive control theory' *in* 'Encyclopaedia of information, linguistics and control' (Pergamon, 1967)

6 FELDBAUM, A. A.: 'Dual control theory problems', Proceedings of the 2nd International IFAC Congress, 1963

7 BLAKELOCK, J. H.: 'Automatic control of aircraft and missiles' (Wiley, 1965)

8 HALL, I. A. M.: 'Study of the human pilot as a servo element', *J. Roy. Aeronaut. Soc.*, 1963, **67**, pp. 351–360
9 STOCKBRIDGE, H. C. W., and SIDDALL, G. J.: Memorandum to the Commonwealth Advisory Commission on Defence Science, 1956
10 PASK, G.: 'The teaching machine as a control mechanism', *Trans. Soc. Instrum. Technol.*, 1960, **12**, pp. 72–82
11 PASK, G.: 'An approach to cybernetics' (Hutchison, 1961)
12 PASK, G.: 'Adaptive teaching systems' *in* 'Teaching machines' (Pergamon, 1964)
13 PASK, G.: 'Comments on the cybernetics of ethical, sociological and psychological systems' *in* 'Progress in biocybernetics—Vol. 3, (Elsevier, 1965)
14 PASK, G.: 'Teaching as a control-engineering process', *Control*, 1965, **9**, pp. 6–11, 153–158, 211–216
15 HUDSON, E. M.: 'Adaptive training and nonverbal behaviour', US Naval Training Device Centre AD610572, 1964
16 SENDERS, J. W.: 'Adaptive teaching machines' *in* 'Programmed learning and computer-based instruction' (Wiley, 1961)
17 ZIEGLER, P. N., BIRMINGHAM, H. P., and CHERNIKOFF, R.: 'An equalization teaching machine', USNRL Report 5855, 1962
18 CHERNIKOFF, R.: 'Human equalization: Concept, machine and measurement', Conference on 'Complex vehicular controls', Farnborough, 1962
19 BOWEN, H. M., HALE, A., and KELLEY, C. R.: 'Tracking training *v.* field study of the effectiveness of the general vehicular research tool', US Naval Training Device Centre AD401583, 1962
20 BRIGGS, G. E.: 'Research on augmented feedback and the acquisition and transfer of skill', *ibid.*, AD293884, 1962
21 KELLEY, C. R., and PROSIN, D. J.: 'Further research with adaptive tasks'. Final report to ONR (Dunlap and Associates, 1967)
22 JEX, H. R., MCDONNELL, J. D., and PHATAK, A. V.: 'A critical tracking task for manual control research', *IEEE Trans.*, 1966, **HFE-7**, pp. 138–165
23 TAYLOR, F. V., and BIRMINGHAM, H. P.: 'That confounded system performance measure—a demonstration', *Psychol. Rev.* (USA), 1959, **66**, pp. 178–182
24 GAINES, B. R.: 'Automated feedback trainers for perceptual-motor skills', Final report to M.o.D., 1967
25 ROSENBLATT, F.: 'Principles of neurodynamics' (Spartan Books, 1962)
26 HOLDING, D. H.: 'Transfer between difficult and easy tasks', *Brit. J Psychol.*, 1962, **53**, pp. 397–407