USAF EVALUATION OF AN AUTOMATED ADAPTIVE FLIGHT TRAINING SYSTEM

JAMES E. BROWN, WAYNE L. WAAG, and EDWARD E. EDDOWES
Flying Training Division
Air Force Human Resources Laboratory

INTRODUCTION

In August 1973, the Tactical Air Command (TAC) began acceptance of an Automated Flight Training System (AFTS) built by Logicon, Inc. The device, installed as a parasitic system on one of the existing F-4E simulators at Luke AFB, AZ was designed to provide automated adaptive training for ground-controlled approaches. In December 1973, TAC requested that AFHRL conduct an operational evaluation of the AFTS in the F-4 combat crew training program. Through mutual agreement of both TAC and AFHRL, the evaluation was initiated in May 1974 and concluded in November 1974. The major objectives of the evaluation were: (1) to evaluate the training effectiveness of the Automated Flight Training System (AFTS) in the F-4 Training Program; (2) to identify desired hardware and software modifications for operational devices; and (3) to identify effective methods of operational training use. Since one of the major characteristics of the AFTS was its use of adaptive training, a brief description of the concept and related research literature will be presented.

The term "Adaptive Training" typically is used to represent a training situation... "in which the problem, the stimulus, or the task is (automatically) varied as a function of how well the trainee performs," (Kelley 1971). It can be seen from this definition that adaptive training requires: (1) "A continuous or repetitive measurement of trainee performance," (2) "One or more task variables that can be adjusted to change task difficulty," and (3) "a means for automatically adapting task difficulty as a function of the performance measurement such that the task becomes more difficult as the trainee becomes more skilled," (Kelley and Wargo, 1968).

In most instances the use of the term "Adaptive Training" refers to a training situation in which a trainee works with a device to help him acquire a skill. The properties of the device are such that the trainee receives a series of practice exercises, the difficulty of which is automatically adjusted according to how well the trainee performs. This trainee-device interaction is similar to a non-mechanized learning situation in which ..."The skilled instructor varies the difficulty of the tasks he gives to a student as a function of how well that student has been performing..." (Kelley 1969).

Adaptive training concepts have been extended to aviation and have resulted in the development of several automated systems. The operation of a typical adaptive training system may be illustrated by the student pilot attempting to acquire skill in straight and level flight. The task is to maintain a constant altitude, heading, and airspeed. The difficulty of the task is varied by changes in turbulence, wind velocity direction, and aircraft weight. The student's score on each trial is based on his deviations about the ideal. As a result of his score, he advances to a higher or lower level of difficulty. Once he has completed the highest level, he is considered proficient. Such an automated program is more correctly termed adaptive scheduling rather than adaptive training since the variation in task difficulty occurs between trials rather than within each trial.

To date, studies of adaptive training in aviation have focused on demonstrations of feasibility. Charles and Johnson (1971) developed an automated GCA training program for the Navy. This computerized system was the forerunner on the F-4 Automated Flight Training System (AFTS) at Luke AFB, AZ, which was evaluated in the present effort. The program was developed for the Training Device Computer (TRADEC) System at the Naval Training Equipment Center, Orlando, FL. A Ground Controlled Approach (GCA) flight segment was selected as the initial training task. Procedures for automated data collection recording and student record keeping were programmed into the system. All operations of the system were performed automatically, including on-line structuring of the training course as a function of student performance. Twelve operational F-4 pilots were utilized for the demonstration. Pilot opinion indicated that the system did reflect operational GAC requirements and would be beneficial in operational training systems. It should again be emphasized that the study was primarily a feasibility demonstration of the capability to automate GCA training and was not a comparative evaluation.

Charles et al (1973) later applied the adaptive training techniques to the acquisition of basic instrument flight skills.

Again, the study was performed using the TRADEC system at the Naval Training Equipment Center. Basic instrument flight maneuvers for the F-4 aircraft, straight and level, climbs and descents, level turns, and climbing/descending turns were automated. Variables such as maneuver difficulty, aircraft weight, center of gravity, and atmospheric turbulence were used to control task difficulty. Four trainees representing a wide variety of aviation skills were given training using the automated instrument flight training program. None of the trainees were operationally qualified F-4 pilots. Training was conducted in one hour sessions with each student completing as many runs as possible. Each successive flight began where the preceding one had terminated. Progress and updating were automatically maintained by the computer program. The authors concluded that an automated syllabus for training instrument flight maneuvers could be implemented and that a student performance score reflecting operational standards could be developed.

Charles et al (1972) also conducted a feasibility demonstration on the application of automated-adaptive training techniques for air-to-air intercept training in the TRADEC flight simulator configured as an F-4. The training task included three phases: (1) a climb task under GCI/CIC control; (2) an attack phase under RIO control; and (3) a descent phase also under GCI/CIC control. Missile intercepts including head-on, forward-quarter, and beam runs were incorporated into a training syllabus. Atmospheric turbulence, aircraft configuration, and bank angle were employed as adaptive variables. Performance was measured objectively for each phase, and the syllabus was restructured based on student performance. Since the study was designed primarily to demonstrate the feasibility of automated air-to-air training, only three subjects were used. The results established the technical feasibility of the training.

As this brief review indicates, published reports to date have only documented the technical feasibility of adaptive training programs. No studies have been completed which compare these adaptive training programs with conventional training. The Navy had planned an evaluation of their version of the Automated Flight Training System, but at the time of this writing, the results were not available (Puig et al, 1974). Consequently, the present evaluation is one of the first studies to compare an operational adaptive training program with conventional training techniques. It should be emphasized that the present research was not an evaluation of adaptive training per se, but rather an evaluation of an entire system in which adaptive training represented only one of its many features.

METHOD

Subjects. The sample group for this investigation consisted of 24 replacement pilots assigned to advanced training in the F-4 aircraft. All were recent undergraduate pilot training graduates and had no advanced aircraft experience.

Task Instrumentation. The Automated Flight Training System (AFTS) developed by Logicon Inc., was used in conjunction with an F-4E Weapons System Trainer Set (WSTS) produced by Singer-Link. The basic simulator, WSTS-15, has a limited 2-degree of freedom (pitch and roll) motion base and is used for basic instruments and weapons system training. The device has no visual capability. The AFTS is a parasitic system which obtains information from WSTS-15 and exercises some degree of control over the host device. In its present configuration, AFTS is designed for presenting automated GCA and TACAN training.

The major hardware components of AFTS included:

(1) the monitor buffer, designed for obtaining flight parameter information directly from the host system;

(2) Y-switch, designed for controlling environmental conditions, Data General, aircraft configuration and emergency conditions;

(3) Data General Nova 800 computer system with associated storage devices;

(4) two Metrolab voice generation systems, designed for providing GCA commentary and feedback; and

(5) two Tektronix terminals, one used for control of the AFTS and the other for replays.

AFTS was designed to provide a series of 76 GCA training exercises, ordered according to task difficulty. Four adaptive variables were defined:

(1) wind velocity and direction;

(2) turbulence;

(3) aircraft weight

(4) emergencies

Conditions for each of these variables in addition to their ordering according to difficulty were defined through consultation with experienced flight instructors. Objective performance measures were derived which reflect deviation from the ideal at the "gate" as well as throughout the glidepath. A student's "total" score on one trial as well as his performance on the previous trial would determine the difficulty level on the succeeding trial. (For a complete description of AFTS, see Logicon, 1974.)

Design. An inherent problem in evaluating adaptive training procedures is to insure that the control group receives equivalent training. In the event the control group is given a fixed sequence of trials or if the instructor decides what levels he is to receive, it is unlikely that the resulting schedule will be the same as that received by the adaptively scheduled group. Although such control procedures would better reflect alternative real-world training strategies, they nonetheless would be subject to the criticism on non-equivalent training sequences. To overcome this potential objection, a yoked experimental procedure was followed. Each subject in the adaptive group was randomly paired with a subject in the control group. Subjects in the adaptive groups were adaptively scheduled according to the AFTS system operation. Each subject in the control group received the same sequence of difficulty levels as his counterpart in the adaptive group. In this manner, the training for the two groups were equivalenced in terms of the difficulty levels and the sequence presented.

Procedure. The present simulator syllabus for F-4 training consists of 22 sorties. GCA training for the present study was accomplished during the first 15 sorties. For sorties 1 and 2, each student received one GCA. For the remaining sorties, each student was given three, thereby providing a total of 41 trials. Due to equipment problems, the GCA's for one sortie were cancelled thus reducing the total to 38. Of this number, six mandatory GCA's were given at fixed places in the sorties in order to meet syllabus requirements. The adaptive group received instruction from the machine while the control group used flight instructors for the first eight sorties and training-devicement for the remaining sorties.

Two criterion rides were given, one following the end of GCA training at Sortie 15 and the other approximately four months later following Sortie 22. Between these two criterion rides, students received no GCA training or practice in the simulator. Each criterion ride consisted of six GCA's. These were to be flown using highly qualified GCA controllers from the Base Communications Squadron while the remaining three were to utilize the AFTS system. Three levels of difficulty were chosen: (1) Level 1 -- no emergency, no turbulence, light weight, and a strong headwind; (2) Level 30 -- no emergency, maximum turbulence, maximum weight, and a strong tailwind; and (3) Level 49 -- single engine failure at seven miles from touchdown, maximum turbulence, maximum weight and a strong tailwind. Students flew these three GCA's under one of the controller conditions, rested for a few minutes, and then flew the same three under the alternative controller condition. Order of presentation was counterbalanced for both criterion rides.

Upon activation of AFTS system, the student is directed to "Contact Phoenix approach controller", descend to 3000 feet and turn to heading 210 degrees. Once these conditions are met and the aircraft weight is adjusted properly, the aircraft is repositioned to approximately 12 miles from touchdown. At 11 miles, the student is instructed to "contact final approach controller, 301.5, Local 17." When accomplished, AFTS replies, "Sim 9-er, how do you hear me?" Once the student responds, AFTS replies, "Roger, do not acknowledge further transmissions." Up to this point, the procedure for both the adaptive and control group was the same. For the adaptive group the machine began giving course heading and glide slope information. For the control group, the machine "voice" was locked out, and the instructor presented glide slope and course information to the student. Such a procedure enabled the use of objective scoring for both groups.

Performance Measures. Time-on tolerance measures are generated for glide slope angle, course heading, and angle of attack. For each parameter, five categories are defined and percentages computed -- well above/left, slightly above/left, on slightly, below/right, and well below/right. These measures are accumulated from the time of glide slope intersection until the student goes out of limits or he reaches 3/4 mile from touchdown. Scores used in the subsequent analysis for each of these parameters were simply the sum of the percentage of time "on target" and half the time slightly above/left and slightly below/right. The path score is a weighted average of these scores for the three parameters. A gate score is also computed which gives information about the parameters at the "gate." Included in this score is information on heading rate and altitude rate as well as deviation about the ideal glide slope angle, course heading, and angle of attack. The path score and gate score are combined to produce the "total" score. It is this score which is used in the adaptive logic to determine the student's advancement.

The AFTS software was modified for the criterion ride in order to yield three additional scores: root-mean-square (RMS) glide slope angle error, RMS course angle error, and RMS angle of attack error. For these scores, measurement was initiated whenever the student reached 7.5 miles from touchdown. It is at this point that the student should begin his descent if his altitude is correctly maintained. Since the AFTS software does not begin scoring until glide path intersection, it is possible for the student to remain low and scoring not be initiated until he is near the "gate." In such a case his score might be higher than would be expected.

RESULTS

Evaluation of Training Effectiveness

An examination of differences between performance data for the adaptive (AFTS-trained) and control (instructor-trained) groups provides information regarding the training effectiveness of the AFTS system. If the AFTS was to be considered an effective training system, then students in the adaptive group should have performed better or at least as well as those students in the control group. Otherwise, the operational utility of such a system would be seriously questioned. Data pertaining to this question could be divided according to source--that obtained during the training period and that obtained during the two criterion sorties.

Training Data

The original experimental design called for one GCA to be administered on the first two sorties and three GCA's on the remaining 13. Due to system failure, GCA's for one sortie were cancelled, thereby reducing the total number to 38. Of these, six were emergencies required by the training syllabus on specified sorties. Therefore, only 32 GCA's were administered as originally intended for each student. Furthermore, one student in the control group suffered a fractured collarbone and was placed on a medical-hold status, thereby eliminating him from the study.

Since both the adaptive and control groups received the same levels of difficulty, inferences were based on performance data rather than measures derived from difficulty levels. Nonetheless, data reflecting changes in difficulty level as a function of training trials are characterized by increasing means and variability as a function of training trials.

An examination of individual learning curves reveals certain trends. It appears that the major difference among individuals is not the slope of the learning curve, but rather the number of trials before consistent advancement in difficulty level occurs. In other words, once the student has mastered the concept of the GCA, and what the requirements are, then he will consistently advance. Otherwise, he remains at the lower levels of difficulty. For such students who do not master the GCA concept quickly, it is unlikely that the introduction of emergency conditions will be of any value. The individual learning curves for several students indicate this to be the case.

The most important question to be addressed by the training data concerned potential differences in performance for the adaptive and control groups. Data for the 32 GCA training trials were combined to produce 11 training blocks. The first block consisted of the first two GCA's given on the first two sorties. The remaining ten blocks consisted of consecutive sets of three training trials. For each performance measure, the block score was simply the mean of all trial scores within that block. Seven measures of performance were used -- path completion (scored 0 or 1), glide slope score, course angle score, angle of attack score, path score, gate score, and total score. Mean scores for each measure pooled across all training blocks are presented in Table 1.

Table 1

Mean Scores for Training Sortie Data

MEASURE	ADAPTIVE	CONTROL
Path Completion	.613	.606
Glide Slope Score	58.011	59.877
Course Angle Score	64.007	66.359
Angle of Attack Score	30.279	28.277
Path Score	57.806	58.481
Gate Score	25.445	26.695
Total Score	165.158	162.663

Each dependent measure was analyzed by an analysis of variance having one between factor (adaptive vs control) and one within factor (training block). These results are presented in Table 2. As indicated, only the training block factors produced a significant effect for the dependent measures. Neither the group factor nor its interaction

Table 2

Summary of Analyses of Variance for Training Data

MEASURE	A	B	AB
Path Completion	.0058	8.1400**	.3290
Glide Slope Score	.7018	3.8310**	.7127
Course Angle Score	.8727	2.1881*	.7268
Angle of Attack Score	.1915	2.9454**	1.0656
Path Score	.1426	8.2004**	.8340
Gate Score	.1071	5.3284**	.3622
Total Score	.0322	10.7876**	.3117

*p <.05 A Group Factor
**p <.01 B Trial Block Factor

with the training block factor was significant. In other words, the data revealed no difference in performance between the adaptive and control groups. A priori t-tests were computed to determine whether any group differences existed during the first training block. Again no differences were found suggesting the initial ability levels for the two groups to be equivalent.

Since no group differences were indicated, the data were pooled and found to be consistent for all dependent measures. There is an initial increase in performance through the first training blocks followed by a decrease and another increase. It seems likely that the decrease noted in blocks 5 through 9 reflect the introduction of difficulty levels requiring emergencies. In any case, it is apparent that measures of performance do change as a function of training trials. The adaptive logic does not alter difficulty level so as to maintain the performance data constant.

To summarize, the data reflect no differences between the adaptive and control groups during the training period. An examination of the descriptive statistics reported elsewhere reveal the results to be nearly identical. It seems safe to conclude that the data indicates both groups received equivalent training. No differences could be detected.

Criterion Sortie Data

For each criterion sortie, half of the GCA's were controlled by the machine while the other half were administered by highly qualified GCA controllers. Performance under the experienced GCA controllers was assumed to represent the major criterion for evaluating the training effectiveness of the AFTS system. In other words, were there performance differences between the adaptive and control groups using experienced controllers? The primary concern was whether the adaptive group, trained on the machine, could effectively transfer to the GCA task using actual controllers. Aside from this question, the data permitted an evaluation of the automated machine controller using the experienced GCA controller as the standard.

To answer these questions, an analysis of variance was performed for each dependent measure. The factorial design consisted of one between subjects factor (group) and three within subjects factors - criterion sortie (1st vs 2nd), type of controller (machine vs human operator), and level of difficulty (Level 1 vs Level 30 vs Level 49). A summary of the resulting F-values for all main effects and interactions are presented for each dependent measure in Table 3.

As the data indicates, the only significant main effects were those factors reflecting type of controller and level of difficulty. Professional GCA controllers produced significantly better scores for all measures except GCA completion, RMS angle of attack and the gate score. The group means are presented in Table 4. For the levels of difficulty factor, only the RMS angle of attack measure was not significant. Group means for this factor are presented in Table 5. The data indicate that measured performance for Levels 1 and 30 to be roughly the same, while significantly degraded for Level 49, the single-engine emergency. The exception is the path score measure wherein Levels 1 and 49 are equal while Level 30 yielded the better performance. However, it must be recalled that Levels 30 and 49 were flown under conditions of maximum turbulence. The measurement algorithm adds 15 points to the path score to compensate for such turbulence. Consequently, subtracting this

Table 3

Summary of Analyses of Variance for Criterion Ride Data

SOURCE	RMS GLIDE SLOPE	RMS COURSE ANGLE	RMS ANGLE OF ATTACK	PATH SCORE	GATE SCORE	TOTAL SCORE
A	.026	.300	.137	.053	1.308	.688
B	.639	.168	3.820	.981	.613	1.046
C	10.325**	24.707**	.804	40.800**	2.199	5.554*
D	26.284**	30.517**	1.078	21.322*	12.880**	11.406**
AB	.068	.000	.169	.060	1.320	.423
AC	.788	.155	1.314	1.458	.914	1.178
AD	.062	.171	.503	.175	1.028	.293
BC	.010	.111	.002	.209	.314	.025
BD	.466	1.504	.458	1.415	.336	.019
CD	1.258	1.077	.924	.286	.143	.038
ABC	.978	2.810	.008	2.736	.060	.281
ABD	1.414	.324	1.539	.129	.192	.024
ACD	.907	2.971	1.401	.877	.288	1.548
BCD	.122	2.249	2.431	.378	4.485*	4.560*
ABCD	.519	.082	.896	.356	4.192*	3.272*

A - Group
B - Criterion Sortie
C - Type of Controller
D - Level of Difficulty

* p .05
**p .01

Table 4

Mean Scores as a Function of Type of Controller

Measure	GCA Operator	Machine
Path Completion	.833	.812+
RMS Glide Slope	.241	.294*
RMS Course Angle	.320	.454*
RMS Angle of Attack	1.885	1.820
Path Score	80.141	71.387*
Gate Score	46.627	42.629
Total Score	220.126	207.472*

*Significant Differences
+No Significant Test

Table 5

Mean Scores as a Function of Levels of Difficulty

MEASURE	LEVEL 1	LEVEL 30	LEVEL 49
Path Completion	.935	.870	.663[t]
RMS Glide Slope	.214	.208	.380*
RMS Course Angle	.284	.323	.553*
RMS Angle of Attack	1.737	1.834	1.991
Path Score	71.078	85.386	70.838*
Gate Score	52.882	48.439	32.650*
Total Score	221.448	226.794	193.157*

[t]No Significance Test
*Significant Differences

Table 6

Mean Group Scores GCA's Administered by Professional Controllers

MEASURE	ADAPTIVE	CONTROL
Path Completion	.861	.803
RMS Glide Slope	.236	.247
RMS Course Angle	.339	.299
RMS Angle of Attack	1.895	1.874
Path Score	76.976	75.981
Gate Score	48.251	44.974
Total Score	221.412	218.723

amount from the mean reveals Levels 1 and 30 to be roughly equivalent, with performance on Level 49 significantly degraded. Since the path score is part of the total score, the same logic applies. In summary, the criterion data suggest performance on Levels 1 and 30 to be roughly equivalent. Performance is significantly degraded only on Level 49, the single engine emergency.

As previously indicated, the critical comparisons were between the two groups for the GCA's conducted by the actual controllers. Table 6 presents the mean values for each of the dependent measures. A priori t-tests were computed for each measure. No statistical differences were obtained. In other words, the adaptive group trained by the machine controller performed as well with the actual GCA controllers as did those trained by human operators.

The only significant interactions involving a group effort were for the gate score and total score measures. An analysis of the fourth order interaction for the gate score revealed that the adaptively trained group performed significantly better on Level 30 during the first criterion sortie in which the machine served as the controller. The total score produced similar findings. T-tests between the groups for each difficulty level revealed no differences whenever the human GCA controllers were providing commentary.

Considering the data collected during the two criterion ride sorties, it is apparent that no reliable differences in performance could be detected between the adaptive and control groups. However, the trends of the criterion test data reflecting somewhat better performance of the adaptively trained students and the high total scores for both groups lead to the conclusion that the AFTS is an effective system for training GCA's. The machine appears to train as well as the instructor and apparently does not train any adverse GCA responses as measured during this study. Although the data have clearly established AFTS to be an effective training device, the cost-effectiveness of the system remains a question beyond the scope of the present evaluation.

DISCUSSION

The results of the evaluation indicated the AFTS to be an effective system for training GCA's. A major concern was to determine whether the AFTS provided any negative training. Throughout the evaluation, no information was gathered suggesting this to be the case. Aside from the major conclusion regarding the training effectiveness of the system, the data seemed to warrant implications in several areas.

Modification and Use of the System

The evaluation surfaced a number of areas in which modifications to the system should be considered.

Difficulty Levels

The requirement for 76 difficulty levels is questionable. The individual student learning data suggest that fewer difficulty steps will be adequate for training. Indeed, the use of difficulty levels raises several questions which should be examined.

How many difficulty levels are required in the AFTS? Although students in the study received more GCA training than the normal F-4 syllabus required, none of the students reached the top difficulty level (e.g., Student #2 reached Difficulty Level 68 in 32 trials). Thus, normal students probably will never be exposed to all of the training conditions available in the system. Without bypassing the adaptive scheduling feature of the AFTS, lower achieving students will not be exposed to GCA emergencies presently required in the F-4 syllabus. Thus, the number of difficulty levels should be reduced to the number of steps that can be accomplished realistically in a training program. This modification will result in a reduction in software program size and will set up attainable flying training goals. The exact number of difficulty levels will depend upon the GCA training objectives identified by the Tactical Air Command.

Which difficulty levels should be retained in the AFTS? The answer to this question lies most appropriately in the domain of instructional systems development. Detailed specification of the AFTS training objectives will not only reduce the number of difficulty levels but will result in identification of the type of GCA training to be accomplished. The data from the evaluation indicated that factors such as wind direction and speed, aircraft weight, and atmospheric turbulence do not significantly affect pilot performance in GCA's after the basic GCA task has been mastered in the AFTS. Emergency conditions which require either an aircraft configuration change from other than normal or which increase the task workload of the pilot are factors that change the real difficulty of the GCA task. Application of instructional systems development principles will assist in identification of the desired difficulty levels. Additionally, it is recommended that a student data bank system to collect data on pilot performance for specific GCA difficulty levels will be of significant benefit in determining which difficulty levels should be retained in the AFTS.

Adaptive Scheduling Algorithm

The performance formula by which the AFTS increases or decreases the difficulty level of GCA training is called the adaptive algorithm. Within the formula, Path Score and Gate Score receive equal weighting in the scoring algorithm. Similarly, glide slope, course, and assigned angle of attack receive equal weighting in the computation of Gate Score and Path Score. Yet, the adequacy of these formulas should be empirically verified to determine the contribution of each performance parameter to pilot performance.

An issue which is separate but related to algorithm design concerned the number of steps that pilots should be incremented or decremented based upon performance. The AFTS presently will increase difficulty levels up to a maximum of three levels. However, the efficiency of this limit is subject to question. In fact, is it necessary for the AFTS to set students back in difficulty levels based upon performance? Several instructors in the evaluation suggested that the system should increment but not decrement students. The individual student learning data collected in the study suggest that early in GCA training, pilots could be retained at a given difficulty level until a high level of GCA performance skill is attained. Succeeding changes in difficulty level then might be increased (in relation to skill level) up to a maximum of four, five, or more steps. The precise number of steps for changes in difficulty level should be re-evaluated so that the AFTS progression formula can be made more efficient.

Maintenance of Flying Skills

A question which is rarely considered by most proponents of adaptive training or adaptive scheduling is what happens when the student reaches the highest difficulty level. The AFTS does not consider this question. Yet designers of automated systems in flying training should consider the issue. If the student reaches the top step in the program, how should his skill be maintained in a given task? If he stays at the most difficult GCA training step, then skills on lower difficulty levels (i.e., various GCA emergencies) may not be retained. Other approaches would suggest that the AFTS program should restart the student at the beginning difficulty levels for emergencies or that a special test program incorporating selected emergency GCA difficulty levels be used. The recommended solution based upon experience from this evaluation is a combination of previously mentioned approaches. When pilots attain the top GCA difficulty level of the AFTS, the software program should automatically change the student to a special skill maintenance program. This program would consist of selected GCA emergency difficulty levels. If the student had difficulty with a particular emergency, the program would branch automatically to the main AFTS training program for remedial training. Upon completion of the remedial training, the students GCA training would be returned to the skill maintenance program for continuation training. Other equally effective skill maintenance programs can be conceived; however, all approaches should be carefully evaluated with respect to instructional objectives.

Adaptive Training

It should be re-emphasized that the present study was not an evaluation of adaptive training, but rather an operational system in which adaptive training was only one of its many characteristics. The experimental group, in addition to the adaptive scheduling based on their own performance, received GCA training with the following characteristics: (1) standardized instruction for all GCA's; (2) knowledge of results from the performance measurement printouts; and (3) feedback using the replay capability. Consequently, it was impossible to assess the contributions of each of these characteristics to the training effectiveness of the system. Nevertheless, there were characteristics of the data which do reflect upon the concept of adaptive training.

One of the major requirements of adaptive training is that variations in the adaptive variable should produce changes in task difficulty. It is assumed that the resulting sequence of tasks is arranged in order of increasing difficulty. As indicated previously, the 76 levels of difficulty in the AFTS were defined with the aid of experienced instructors from the F-4 Instructional System Development Team (ISDT). That these discrete steps actually represent a series of increasingly more difficult tasks -- as measured by actual performance -- remains unverified. The collection of such data using a sample of experienced pilots would be tedious and time-consuming. The data available from students within the study are confounded by the fact that the information was collected during actual training. As indicated in the results, performance data varied as a function of training trials. It is apparent that task difficulty is not varied sufficiently in order to maintain a constant level of performance. Such information suggests that either: (a) the adaptive variables used in this training system do not actually produce difficulty changes; (b) the sequence of 76 tasks does not represent a set of increasingly difficult tasks; or (c) the adaptive scheduling algorithm is inappropriate.

An underlying assumption of adaptive training is that learning represents a continuous process. On each successive trial, skill is incremented by a certain amount. While each continuous increase in skill level may be seen from the group learning curves, it is rarely the case with individual learning curves. As stated earlier, the learning curves generated by the 12 students suggest mastery of the GCA to represent a process of insight. In other words, students will not advance until they have mastered the concept of the GCA. Once mastered, however, students advance at much the same rates. Such data suggest that once the student "learns" to fly the GCA, variations in wind velocity, direction, aircraft weight, and turbulence have little effect on his performance. Only emergencies in which the aircraft configuration is dramatically changed will affect his performance.

It is the authors' opinion that while variations in wind, weight, and turbulence may add realism to the task, these changes based on performance within the adaptive context do little to facilitate learning. It is suggested that a random presentation of GCA's under these conditions may be as effective as the present system utilizing adaptive scheduling. However, such a statement is a matter of conjecture and is certainly in need of empirical validation. Future studies comparing adaptive scheduling with random and/or fixed presentation would test the utility of the adaptive scheduling feature, per se. The present investigation demonstrated the training effectiveness of the entire AFTS system. Future research could determine the effectiveness of various components, of which adaptive scheduling represents one of the most important.

REFERENCES

Charles, John P., and Johnson, Robert M. Automated Training Evaluation (ATE). NAVTRADEVCEN 70-C-0132-1. Naval Training Device Center, Orlando, Florida, March 1971.

Charles, John P., Johnson, Robert M., and Swink, Jay R. Automated Flight Training (AFT) GCI/CIC Air Attack. NAVTRAEQUIPCEN 72-C-0108-1 Naval Training Equipment Center, Orlando, Florida, 1972.

Charles, John P., Johnson, Robert., and Swink, Jay R. Automated Flight Training (AFT) Instrument Flight Maneuvers. NAVTRAEQUIPCEN 71-C-0205-1. Naval Training Equipment Center, Orlando, Florida, February 1973.

Kelley, Charles R. What is adaptive training? Human Factors, 1969, 11, 547-556, pg 547.

Kelley, Charles R. Fundamental problems. In McGrath, James J. and Harris, Douglas H. (Eds.) Adaptive training. Aviation Research Monographs, 1971, Vol 1, No. 2, pg 2.

Kelley, Charles R. and Wargo, Michael J. Adaptive Techniques for Synthetic Flight Training Systems. NAVTRADEVCEN 68-C-0136-1. Naval Training Device Center, Orlando, Florida, October 1968.

Logicon Inc., Automated Adaptive Flight Training System (AFTS) for F-4E Weapon System Training Set, Volume I, System Operation Manual. Logicon Inc., San Diego, California, January 1974.

Puig, J. A., Johnson, R. M., and Charles, J. P. Evaluation of an Automated GCA Flight Training System. Reported in NTEC/Industry Conference Proceedings, 19-21 Nov 1974, Naval Training Equipment Center, Orlando, FL, pp 269-276.

ABOUT THE AUTHORS

MR. JAMES BROWN is a Research Psychologist in the Air Force Human Factors Laboratory (AFHRL), Flight Training Division, Williams Air Force Base, Arizona. He holds a Master's degree in Industrial Psychology. He has over 16 years experience in human factors and flying training research. Prior to joining AFHRL in 1970, he was employed in private industry for 11 years. He has served as principle investigator and program manager for several flight simulation and flight training research programs. Since employed by the Air Force, he has worked on instructional systems development programs, future Undergraduate Pilot Training programs, development and tests of an advanced F4E flight simulator, and performance measurement studies in Undergraduate Pilot Training. He has authored numerous technical reports and professional papers.

DR. WAYNE L. WAAG is a Research Psychologist with the Air Force Human Factors Laboratory (AFHRL), Flight Training Division, Williams Air Force Base, Arizona. He holds a Ph.D. in Engineering Psychology from Texas Tech University. He has more than 4 years of experience in military psychology. Since joining AFHRL in 1973, he has directed the development of an automated quantitative pilot performance measurement system and associated student data system for handling the performance data collected. He has authored more than 30 research reports and presentations.

DR. EDWARD E. EDDOWES is a Supervisory Research Psychologist in the Air Force Human Factors Laboratory, Flight Training Division, Williams Air Force Base, Arizona. He holds a Ph.D. in Experimental Psychology from the University of Florida. He has more than 16 years of experience in human factors, training system development and higher education. Since joining AFHRL in 1972, he has developed a program of research focusing on the measurement of pilot performance, innovations in flying training, pilot information processing behavior, cognitive aspects of aircrew performance and student-pilot/instructor-pilot interactions. He has authored more than 35 professional articles and research reports.