Aviation Research MONOGRAPHS

ARL

VOL. 1, NO. 2, AUGUST 1971

ADAPTIVE TRAINING

Edited by: James J. McGrath and Douglas H. Harris

AVIATION RESEARCH LABORATORY · INSTITUTE OF AVIATION
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Aviation Research Monographs, Vol. 1, No. 2, August 1971

Stanley N. Roscoe, *Editor*
William H. Schmidt, *Associate Editor*

Issued aperiodically by the Institute of Aviation,
University of Illinois at Urbana-Champaign

Ralph E. Flexman, *Director*
Jesse W. Stonecipher, *Associate Director*
Stanley N. Roscoe, *Associate Director for Research*

Library of Congress catalog card no.: 77-171977

Published by the Engineering Publications Office,
College of Engineering,
University of Illinois at Urbana-Champaign

Aviation Research

MONOGRAPHS

AVIATION RESEARCH LABORATORY
INSTITUTE OF AVIATION
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Foreword

This monograph is the second in a new series published by the Aviation Research Laboratory of the Institute of Aviation. These monographs report theoretical analyses, experimental findings, and events of scientific importance leading toward the laboratory's goal of solving real problems in aviation. This issue is a report of a two-day conference on adaptive training conducted by the Institute in April 1970 at the University of Illinois. The conference had two main objectives: One was to further the understanding of adaptive training among those who might wish to study or apply adaptive methods. The other, more immediate, objective was to gain insights helpful to the Institute in designing adaptive training programs for the Synthetic Flight Training System developed by the Link Division of Singer-General Precision, Inc., under contract to the Naval Training Device Center, for the purpose of training Army helicopter pilots.

The participants included scientists who had made significant research contributions to the development of adaptive methods and training specialists who were engaged in solving current problems of pilot training. The services of Anacapa Sciences, Inc., were enlisted to assist in conducting the conference and to edit and report the proceedings. The success of the conference can be attributed to the professional knowledge and insights supplied by these participants. The Institute gratefully acknowledges their efforts.

The Link Foundation has sponsored the publication of this report, as it has many previous Institute of Aviation publications. This public service is, once again, gratefully acknowledged.

RALPH E. FLEXMAN, *Director*
INSTITUTE OF AVIATION

Participants

GENERAL CHAIRMAN

Ralph E. Flexman, *Institute of Aviation*

SESSION CHAIRMEN

Charles R. Kelley, *Dunlap & Associates, Inc.*
Paul W. Caro, *Human Resources Research Organization*
W. Guy Matheny, *Life Sciences, Inc.*
Ralph E. Flexman, *Institute of Aviation*

DISCUSSANTS

Julius H. Gandelman, *Institute of Aviation*
Sanford N. Green, *Reflectone, Inc.*
Eugene R. Hall, *Human Resources Research Organization*
Scott G. Hasler, *Bunker-Ramo Corporation*
Patricia A. Knoop, *USAF Human Resources Laboratory*
John K. Lauber, *U.S. Naval Training Device Center*
John L. Mansfield, *American Airlines Flight Academy*
James J. McGrath, *Anacapa Sciences, Inc.*
Melvin D. Montemerlo, *Singer-General Precision, Inc.*
Don A. Norman, *Life Sciences, Inc.*
H. Kingsley Povenmire, *Institute of Aviation*
Stanley N. Roscoe, *Institute of Aviation*
John B. Sinacori, *Northrop Corporation*
William A. Stewart, *The Rand Corporation*
Richard L. Weekes, *Institute of Aviation*
Milton E. Wood, *USAF Human Resources Laboratory*

Session 1. Fundamental Problems

DR. CHARLES R. KELLEY, *Chairman*

". . . PARTNERS IN IGNORANCE. . . ."

Dr. Kelley

DR. KELLEY: This workshop is an informal, mutual brain-picking arrangement. Insights on the subject of adaptive training will come by asking questions, by exposing our ignorance, and by offering ideas no matter how obvious or obscure they seem to be. We are all expert enough on the subject of adaptive training to realize that we don't know much about it. The technology of adaptive training hasn't really been developed yet. We are all terribly ignorant. I say this to this group in all honesty, because I'm confident that those of you who have worked on problems of adaptive training will have to agree with me. I have been working on research and development of adaptive systems for about eight years, and I still feel lost when people call me in as a consultant. We don't have the data nor the technology to answer the kinds of questions that are being asked. So, if you, too, have felt in the dark about this field, join the crowd. We're all partners in ignorance, but let's attempt to share what little knowledge and experience we do have.

My procedure is going to be to present informally some of my ideas about the fundamental concepts of adaptive training. Most of these concepts have been presented in an article published in *Human Factors* a few months ago (Kelley, 1969); so they are probably familiar to most of you. Still, a brief review might help us to get started in our discussion of adaptive training.

". . . WE USUALLY MEAN BY 'ADAPTIVE TRAINING'. . . ."

Let's begin with a definition of adaptive training. I have written my definition on the blackboard: *Adaptive training is training in which the problem, the stimulus, or the task is (automatically) varied as a function of how well the trainee performs.*

Note the word *automatically*, in parentheses. I take the position that virtually all individual training in significant tasks is adaptive; it has to be. A beginner is not taught to play the piano by starting with a Liszt rhapsody but by playing scales. Whenever training is a significant problem, whenever the task is hard to learn, the instructor necessarily must suit the difficulty level of the training task to the skill of the student. But, if we were to define adaptive training only in this way, we would subsume almost the whole field of training. So I believe that we usually mean by adaptive training some form of training in which the change in the difficulty level is carried out automatically in response to the student's demonstrated skill. This change might be mechanical, machine controlled, or it might be computer controlled. However it is done, there must be an automatic process that determines the relationship between how well a trainee performs and how tough the task or problem becomes.

This definition implies three fundamental elements of any form of adaptive training system. The first element that is implicit in this definition is some kind of *measurement of performance*. Performance must be assessed in some way. There must be as a minimum a binary assessment, for example, whether the student is performing satisfactorily or unsatisfactorily. Without performance measurement, considered broadly, you have no information to serve as a basis for adaptive changes in the training task. The second element that is implicit in the definition is some method for changing the difficulty level of the training task. This is what I have called the *adaptive variable*. In practice, we may have several adaptive variables and several performance measurements. The third element is the *adaptive logic* to connect the performance measurement with the adaptive variable. These three elements constitute, as I see it, the skeleton of any adaptive training system. I suggest, then, that in dealing with fundamental problems of adaptive training, we structure our thinking around these three concepts: performance measurement(s), adaptive variable(s), and adaptive logic. I believe that we will find that concepts, such as training effectiveness or validity, must necessarily be taken into account in making a decision about these three elements. There is no other way to decide what to measure, what to adapt, or how to interrelate the two.

". . . ONE HUNDRED DIFFERENT KINDS OF ADAPTIVE LOGIC. . . ."

Exhibit 1 illustrates the basic difference between fixed and adaptive training. The adaptive training system "massages" measurements of performance through some form of adaptive logic and then changes the problem, the stimulus, or the training task as a consequence.

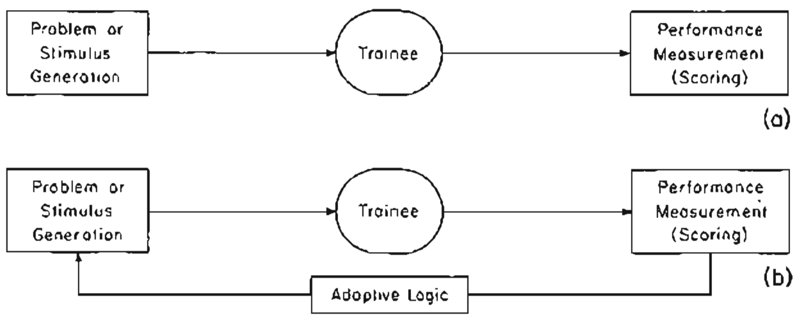


Exhibit 1. Fixed versus adaptive training. (a) Fixed (preprogrammed) training. (b) Adaptive training.

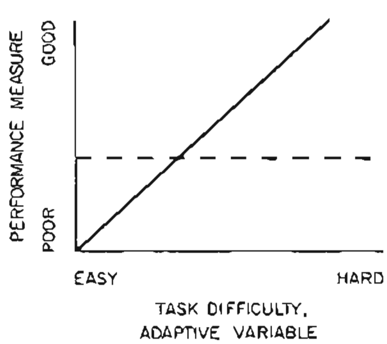


Exhibit 2. Examples of adaptive logic.

Look now at Exhibit 2. The ordinate indicates performance measurement with intervals from *poor* to *good*. The abscissa indicates task difficulty from *easy* to *hard*, which represents the working of the adaptive variable. These are the first two elements implicit in the definition. The third element—the adaptive logic—is the relationship that holds between them. We might set up a relationship like the solid line, which would mean that when a person performs poorly, the task is made easy; when a person performs well, the task is made hard. But, this solid line represents just one of one hundred different kinds of adaptive logic we might use. I concern myself with a class of adaptive systems in which the adaptive logic looks like the broken line. That is to say, I set up a system in which the adaptive variable changes in such a way that the performance measurement does not vary. Task difficulty changes to keep performance at some preset level. This is another possible adaptive logic. There are certain advantages to doing it in this way, particularly in simple systems where you only have one performance measurement; but consider how many other different lines you could have relating the ordinate and abscissa of this graph.

Exhibit 3 is a diagram of the kind of system that Hudson (1962) and Birmingham, Chernikoff, and Ziegler (1962) utilized in their early adaptive techniques applied to tracking tasks. They measured tracking error (averaged error squared over time), and they used this to adjust the amount of display augmentation (quickening) in the display. The constant, b, is subtracted out in such a way that, when performance reaches a certain level, all augmentation is removed because the multiplier controlling the amount of display augmentation receives a zero signal in one channel. This is to say, there will be no display augmentation coming into the display, and the person will be tracking the system as it actually is. Until he can perform that well, however, he has a degree of augmentation, the amount of which is controlled by his time averaged error. When he starts out learning to track, he gets help from the display augmentation, which simplifies his task. When he becomes skilled, he no longer needs or gets that help.

The equation, $y = aX - b$, describes the adaptive technique used by Hudson and by Birmingham. The equation, $dy/dt = aX - b$, de-

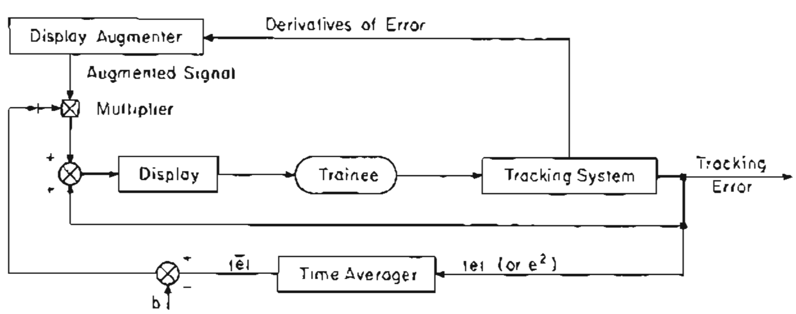


Exhibit 3. Early adaptive trainer.

scribes the technique that I prefer. In both cases, y represents the adaptive variable, X is the performance measurement, and a and b are constants. The real difference between the two is that in the second equation the rate of change of the adaptive variable, rather than the level of that variable, is a function of the performance measurement. This keeps the performance measurement fluctuating about the same value as a person proceeds in training. The system becomes stable (the adaptive variable stops changing) at the point where aX is equal to b. The rate of change can be regulated by regulating a; the difficulty, by regulating b.

Although the symbols are different, the equation,

$$C = K\int_0^t (e_L - e)\,dt + C_i,$$

is essentially equivalent to the integral of the equation, $dy/dt = aX - b$, (there is a difference in the constants). Here, the adaptive variable is equal to a constant, K, multiplying the time integral of the difference beween an error term and an error threshold (which is also constant, e_L) plus an initial condition, C_i, which forms the constant of integration.

$$\dot{C} = K(e_L - e),$$

or

$$C = K\int_0^t (e_L - e)\,dt + C$$

where:

C = the adaptive score which indicates trainee performance and varies the difficulty of the task. A high C score represents good performance and high task difficulty.

e_L = preset criterion of performance or threshold of error.

e = system error, however defined, expressed as a positive quantity, for example, absolute tracking error, error squared, vector error of a multi-axis task, or a weighted error combining several performance variables.

". . . THE EFFECTS OF THESE EQUATIONS. . . ."

These graphs (Exhibit 4) illustrate the effects of these equations. They represent two kinds of adaptive systems, b and c, with a fixed-difficulty system, a, shown at the top for comparison. The ordinate is measured performance, and the abscissa is the period of training or number of trials from the beginning to the end of the training sequence. Two theoretical performance measurements are shown, one being error, which gets better over trials by reducing to some small value (broken line), the other being speed or accuracy, which gets better by increasing (solid line). The relation between the two kinds of performance may, of course, be more complex than the one shown here.

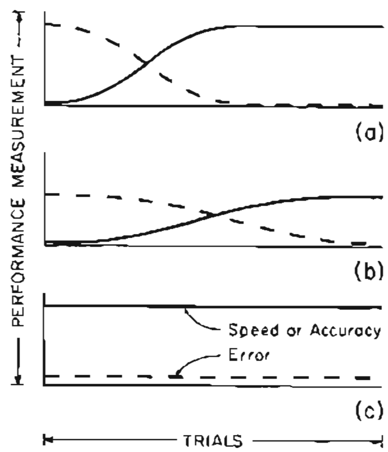


Exhibit 4. Performance change with learning. (a) Fixed difficulty system. (b) Adaptive system 1. (c) Adaptive system 2.

With a fixed-difficulty system (Exhibit 4a), the student starts out with the same task he will have to perform throughout the trials. That is, the task he does at the start is just as hard as the task he does at the end of the training. In this case, it is too hard for him at the beginning, and he can't cope. We found this most dramatically in giving trainees, who had never done any tracking until then, a three-axis acceleration tracking task with no rate information displayed. It's hard to do a one-axis acceleration tracking task without a rate display. When the subjects attempted a three-axis tracking task with no rate displays, they came in day after day, and one or the other of the three axes stayed off the scale of the display, outside of range. They really were terribly frustrated and went through several days of this before they began to make any progress. So, to generalize, the beginning period in which there is no evidence of progress might be several days long in very difficult tasks. Then their performance starts improving. We see a period of rapid progress here where the curve is steepest. This is the point where training appears to be really taking effect, as shown by dramatic evidence of progress from time period to time period or from trial to trial.

Therefore, you might say the difficulty of the task is most appropriate at the point where these curves are steepest, as far as we can tell from measured performance. Surely some kind of learning is taking place at the flat portions of the curves, but it is hard to say how much. I feel certain from my own experience that a lot of training time is wasted when problems or tasks are too tough for trainees. This is partly caused by motivational factors; however, it is partly caused by the fact that the desired skills are not being efficiently developed. The student doesn't get effective practice if the task is much too difficult. Nevertheless, with a fixed-difficulty system, the trainee usually masters the task after enough time. Whatever the problems, he may eventually learn to cope with them. But at some point, the problems then become too easy for him. He has the task down pat and makes no further progress that we can measure. He may be reinforcing habits that he has formed, the skill may be becoming consolidated, but in terms of measurable performance he has stopped making progress around the point where the curve levels off. So, at the start, the system is too hard for the student; at the end, it becomes too easy for him.

When we make a system adaptive, we eliminate part of this problem because the system starts out easy for the student and automatically gets harder as he goes along. However, if you use the type of system represented by the first equation, there is a confounding effect (Exhibit 4b). The confounding takes place in this way: measured performance improves in time, but simultaneously the system is getting harder. Thus, the performance measurement early in the trials doesn't mean the same thing as it does later because he is dealing with a different task.

Since I am much concerned with problems of methodology, I did not want to have two things changing at the same time. That is why I developed adaptive systems which work in terms of the second equation, in order to change the task in such a way as to keep the subject's performance at a constant level (Exhibit 4c). By *constant*, I mean the same average performance level over a defined period even though performance will vary from moment to moment or from trial to trial, fluctuating around the desired level. At the end of training, the trainee is getting the same score (for example, the same error, the same accuracy, the same speed score) that he was getting on the first day of training. There is no difference in performance scores, so performance measurements no longer indicate how well the trainee is performing. All of the variance is cast into the adaptive variable. We no longer ask, "How well does a person perform?" in terms of error, speed, or accuracy. Instead, we ask, "How difficult a system can he cope with?" at the desired level of performance.

". . . LINEARITY FOR A VERY LONG TIME. . . ."

Exhibit 5 shows changes in the adaptive variable as a function of training with the three systems. In a fixed-difficulty system (Exhibit 5a) there is no adaptation, so the adaptive variable does not vary, by definition. In the first type of adaptive system (Exhibit 5b) the curve shows gradual progress in most tracking tasks, though it might have a different shape in other kinds of tasks. The task is getting harder with time, and the adaptive variable is going up. At the same time, the performance measurement (Exhibit 4b) is showing improvement. So, we have to keep both curves in mind to have a true idea as to how well the person is performing. With this kind of adaptive system, it is necessary to consider both the level of the adaptive variable and the level of the performance measurement because the two factors co-vary. Using the third system (the type of adaptive system indicated by Exhibit 5c) in tracking studies, we have observed virtually a linear increase in certain adaptive variables over time or over number of trials. In particular, we have found this linear change when we use the amplitude of the forcing function as the adaptive variable in training tracking performance. This would correspond to gust amplitude such as Guy Matheny and his colleagues have used in a few studies. We found that this linearity is maintained for a very long time, for example, for 300 five-minute tracking trials spaced out over several weeks. Only towards the end of those trials was there a tailing off from linearity at the top of the curve, which was the result of the natural ceiling that occurs eventually when amplitude of forcing function is the adaptive variable. That ceiling is reached when the amplitude of the forcing function exceeds the amplitude of the controller's response that can counteract the forcing function. In other words, the system's maximum controlling output will produce

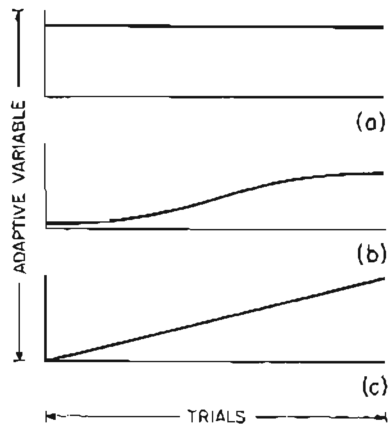


Exhibit 5. System change with learning. (a) Fixed-difficulty system. (b) Adaptive system 1. (c) Adaptive system 2.

a natural ceiling for the forcing function amplitude when it is the adaptive variable.

The primary point, however, is that the performance score stayed the same in this system, and the changes in the adaptive variable, which serve as the index of skill, remained linear with training over an extended training period.

". . . DIFFERENT WAYS OF MEASURING HOW WELL SOMEONE LEARNS. . . ."

Exhibit 6 shows, by actual data, the linearity that we sometimes get in an adaptive tracking system compared with the learning curve for a fixed-difficulty tracking task. The task was a two-axis acceleration control task with no rate displays, two separate meters plus a third meter indicating a score which corresponded to the amplitude of the forcing function on adaptive trials, and an integrated tracking error on fixed trials. On every other trial, subjects performed the fixed-difficulty task. They were given an adaptive trial, a fixed trial, an adaptive trial, a fixed trial, and so on for 20 trials a day for nine days. Exhibit 6 has two ordinate designations, one for the adaptive task and one for the fixed-difficulty task. The initial point and the final point were purposely made to coincide by adjusting the scale of the ordinate. The broken lines represent the standard deviation of the 10 scores that go into each data point. The fixed-task learning curve represents improvement (over pretrial) in RMS error. The adaptive training curve is the amplitude of the forcing function (the adaptive variable) for a constant RMS error. So, at the beginning of these trials, when subjects couldn't really track at all, amplitude of the forcing function cranked down to zero. Quite soon, they could control a very slight forcing function and eventually controlled a large forcing function. Notice the relatively linear nature of the improvement as measured by the adaptive learning curve.

As Stan Roscoe pointed out the other day, there is an artificial restriction of range at the lower end because there is a limit to how much error could be measured. When people are off the scale all the time during early trials, the variance is artificially small because they are exceeding the error that our equipment would measure. However, the relation of mean to standard deviation in these curves

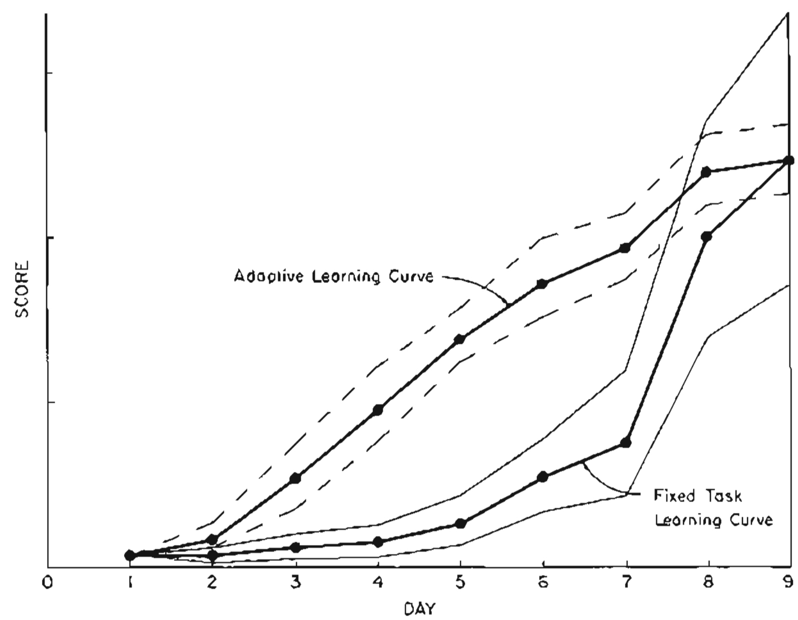


Exhibit 6. Tracking improvement as measured by fixed versus adaptive techniques.

is always much in favor of the adaptive curve. The coefficient of variation is so much smaller for the adaptive trials that, if we just take the trials from Day 4 on when restriction of range is no longer a significant factor, it would be necessary to take from five to 20 times as many trials with the fixed task as it would with the adaptive task to reach the same reliability of measurement.

These data say nothing about the effectiveness of adaptive training compared with fixed training. Since subjects were trained alternately on the fixed and adaptive trials, these curves are simply two different ways of measuring how well someone learns to track. Subjects were just as skilled when performing on the fixed task as when performing on the adaptive task. In the early trials, the fixed task was too difficult. The adaptive task was quite a lot easier than the fixed task, because the forcing function was much less. At about Day 6 or 7, the difficulty of the adaptive task became virtually the same as the difficulty of the fixed task. It was about this time that performance improved most decisively on the fixed task. On Day 8 the subjects already were beginning to level off on the fixed task; which is to say, the fixed task was beginning to get too easy for them.

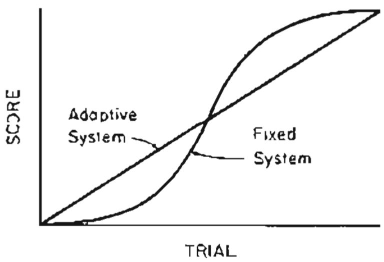


Exhibit 7. Fixed versus adaptive learning curves.

The curve is not carried on until the leveling off is complete, but the fixed-task learning curve would become an S-shaped curve. Exhibit 7 shows what would happen if we had continued the experiment. We ran 180 trials; so it wasn't a short experiment. But, if we had continued it, eventually the fixed system would have leveled off to produce an ogive. If we again matched the first points and the last points on the ordinates, we would have obtained curves similar to those shown here.

Dr. Montemerlo

DR. MONTEMERLO: How do you know you would have obtained such curves?

DR. KELLEY: If the last 40% of Exhibit 7 were omitted and the data rescaled to match end points, it would match Exhibit 6. The curves in Exhibit 6 were truncated and showed only the first part of the full learning curves. Later, these subjects ran another 200 trials on another project. So, there is empirical evidence, though not under truly comparable conditions. But the curve must level off at some point because, for almost all psychomotor skills, the subjects will almost certainly master a fixed task eventually. The adaptive curve doesn't have to level off that way because the system can go on making the problem more difficult indefinitely if the right adaptive variable is used. The subjects would not necessarily reach a natural leveling off point such as they would in the fixed-difficulty system. That one can go on learning much longer when problems are made more difficult than when they are not is really all that this amounts to. Does that answer your question?

DR. MONTEMERLO: Yes. When the two curves crossed, why didn't the adaptive training curve accelerate? Wouldn't the student start learning faster than he did before?

DR. KELLEY: I don't see that he would, at least not necessarily. I think learning was pretty constant throughout training, as the adaptive curve implies, but this is my interpretation. The fixed curve implies integral learning, and I *interpret* this as an artifact of scoring, but my data don't *prove* this the case.

The learning curve in adaptive training also depends on the adaptive variables used. And, there is absolutely no reason why the relationship between trials and adaptive variables should necessarily be linear. It can be anything. It was linear in certain of the tracking systems that I worked with. That was not a coincidence, because I worked very hard to find variables that would show this kind of relationship with performance. Linear functions are very convenient for research purposes.

Unless you have some good reason to believe that more (or less) learning takes place on Day N + 1 than on Day N, I prefer the hypothesis that the same amount of learning takes place. If that is true, then one should have a linear learning curve if performance is measured on an interval scale. Many performance measurements are not interval measurements, of course; and much nonsense about learning rates has resulted from treating them as if they were.

MR. WOOD: In my limited research, I have found a first-order change of difficulty across time. As you have already mentioned, an asymptote can occur in an adaptive system in the form of an artificial ceiling which is a function of the average level of the forcing function and the maximum available control input.

Mr. Wood

DR. KELLEY: With certain adaptive variables. I don't think performance always or necessarily reaches an asymptote.

MR. WOOD: I have found very little variance between subjects as they establish a first-order difficulty curve across time. There appear to be specific difficulty levels for various points in time for a given task and a given adaptive model. In my experience a real learning curve with apparent asymptote does exist in terms of difficulty changes across time.

". . . WHAT CAN BE CHANGED . . . ?"

DR. KELLEY: Here is a diagram (Exhibit 8) of an adaptive vehicle trainer that includes the displays, controls, and problem generator of a typical flight simulator. The trainer is made adaptive by superimposing an adaptive logic that connects the performance mea-

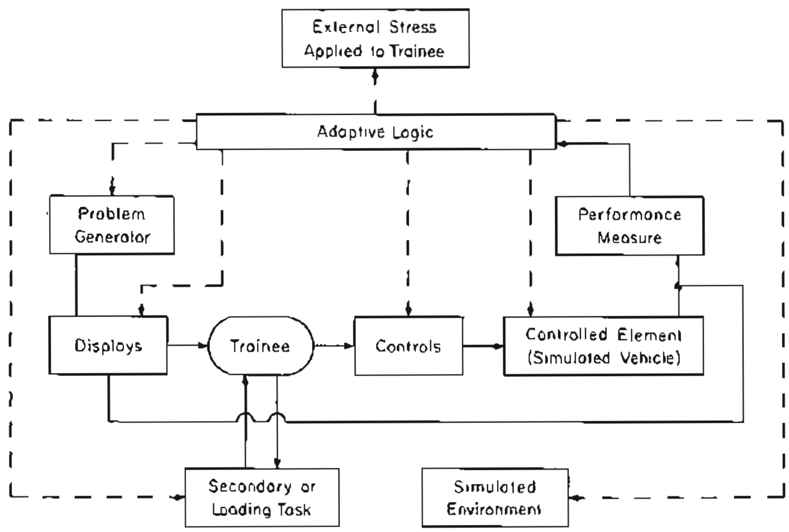


Exhibit 8. The adaptive vehicle trainer.

surement equipment with any part of the system in a way that makes the task automatically harder or easier as a function of that performance.

What can be changed? One thing is the stress that is applied to the trainee. Anything that stresses the trainee in any way that makes the task tougher for him will work; if it makes the task tougher for him, it can be an adaptive variable. The simulated environment can be changed: the oxygen pressure can be lowered in a flight trainer; vibration can be introduced; other factors in the environment, CO^2 level, temperature, g forces (in a centrifuge), and many other such things, can be varied. As long as they affect the difficulty of the task, they could work as adaptive variables. The most common way of adapting, however, is to change the problem or the task itself in some way to make it more or less difficult. This is, by all odds, the way that it is usually done, but it is by no means the only way it can be done. Characteristics of the display can be changed. A display signal can be cut out that was employed to help the novice along, such as the use of quickening or prediction. Display hysteresis, or lag, can be varied as a function of how well the person is performing. The controls can be diddled with backlash or damping or some other factor. One of the most useful methods, for a lot of purposes, is to use a secondary or loading task to make the main task harder or easier. In an operational flight trainer, varying the information or communication load has especially high potential for adaptive training. Are there any questions about this basic diagram of an adaptive vehicle trainer?

MR. FLEXMAN: When you remove a display, it is an all-or-nothing change. Doesn't it do violence to your adaptive logic where you call for a gradual change in difficulty?

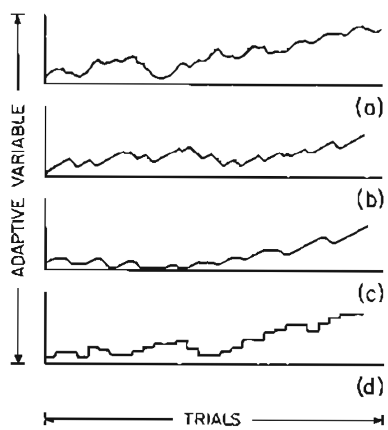

Mr. Flexman

DR. KELLEY: Well, in tracking logic, we often gradually change elements; but in other adaptive systems, we might change them through discrete steps of one sort or another. When using secondary communication-load tasks, for example, either we are feeding the trainee messages, or we are not. That is usually the easiest way to do it. Or, we may have to change task problems by discrete jumps. We might go from one kind of maneuver which we decided that the novice trainee has mastered to a different type of maneuver that is more difficult and that involves a discrete change. I think our adaptive variable, broadly considered, might have both features. Some of the time, it will change in discrete jumps; some of the time, it will change in a continuous gradation. Tracking is the easiest kind of system to implement because you can change in continuous, often linear, fashion. Adaptive training devices for practical tasks cannot often be made as easily as the adaptive tracking systems that a lot of my early work featured.

". . . STRATEGIES OF ADAPTIVE CHANGE. . . ."

Exhibit 9 shows that an adaptive variable can be manipulated in many different ways. It can vary continuously, as in the first curve; it can vary at one of two rates, as in the second and third curves; it can vary in discrete jumps, as in the fourth curve.

In the second curve, the system is either getting harder at one preset rate or it is getting easier at another (greater) preset rate. It is always doing one or the other. There is no point at which it is not either gradually getting more difficult or gradually getting easier. Thus, this curve is made up of ramps with steeper down-slopes than up-slopes. Often this kind of system is used when responses are graded right or wrong and you want the student to get most of them right. If he is to get 90% of them right, the slope down must be nine



(a) $\frac{dy}{dt} = K(X - X_T)$

(b) $\frac{dy}{dt} = K_1$, when $X > X_T$
$= K_2$, when $X < X_T$

(c) $\frac{dy}{dt} = K_1$, when $X > X_1$
$= 0$, when $X_1 > X > X_2$
$= K_2$, when $X < X_2$

(d) $y_{i+1} = y_i + KX$,
when $X > X_1$
$= y_i$,
when $X_1 > X > X_2$
$= y_i - KX$,
when $X < X_2$

where:

i = trial number
y = adaptive variable
X = score (performance measurement, a higher score being better)
X_T = desired score (a constant)
K, K_1, K_2 = constant coefficients

Exhibit 9. Adaptive learning curves.

times as steep as the slope up, a fact that becomes obvious when such a system is implemented. Only at the 90% threshold is the system stable.

Many times it is undesirable for the system to be changing at all times. The first two curves represent systems that are never the same from moment to moment or trial to trial. They are always changing. To keep that from happening, a deadband can be inserted, as in the other two curves. The deadband provides that, whenever the student is performing within specified limits, the system will stay fixed; it won't change. Only when he exceeds these limits will the task be made easier or more difficult. The flat areas in Exhibit 9c represent those areas where he is performing within these limits and the system does not change. That will add stability to a lot of systems that wander too much. It is by no means a panacea. Stability is added by going from adaptation to fixed performance.

In implementing an adaptive training system with a computer, changes are often made in discrete jumps, and there may or may not be a deadband. In the fourth curve it is assumed that there is a point where the system does not change. When changes are made, they can be made by fixed amounts or by different amounts, depending on how the task is implemented. What I mean is, the increment of change can be made proportional to how good the student is performing, or the increments can be the same size. The change can be at fixed time intervals or can be keyed by the performance measurement.

These points, in the long run, do not matter too much. Any one of these strategies of adaptive change can function virtually as well as any others. The nature of the training problem and the nature of the training system implementation will more strongly determine how it should be done than what might theoretically be the *most desirable* way of doing it.

MR. FLEXMAN: Have you experimentally varied the length of the plateau or deadband time? For example, it could be shorter at the easy end and longer as the task gets more difficult.

DR. KELLEY: No, I have not. In the task illustrated, the period in the deadband is entirely a matter of how well the trainee is performing.

". . . ONE DIFFERENCE BETWEEN A GOOD AND A BAD ADAPTIVE SYSTEM. . . ."

There is one important point that is not recognized by everyone. In the kind of adaptive system that we've been talking about so far, when there is a tight performance standard, the student has an easy task to practice on because he cannot meet the tight performance standard except when the task remains easy. Conversely, when he has a very loose performance standard (if you allow him a lot of error), he has a tougher task to do and is forced to practice on harder problems. Now, think for a moment of the training implications of this statement and of reasons why it is important. It is undesirable for trainees to practice only on easy tasks and never on difficult tasks. At the same time, you may not want them to perform sloppily and get in the habit of doing so. The relation here is defined by the adaptive logic, and it is very crucial. One difference between a good and a bad adaptive system can be the difference between the performance standards used.

When we get to more sophisticated implementation of adaptive systems, I would not be surprised if we wanted to vary the performance standards in such a way that a person got a range of difficulty levels from the start in training. We might find that one shortcoming of the simple kind of adaptive tracking system is that it does not give the novice a range of problem difficulty to deal with. Of course, this limitation is not necessarily built into all adaptive systems; it is just the way that a lot of us have done it in the past. If we vary the performance standard, we concomitantly vary the difficulty of the problems that the student must deal with. So, the easiest way to get adaptive variation in difficulty is simply to adjust the performance standard—that is, during the task, to change the performance limits that call for a change in task difficulty. For example, these adjustments can be programmed to go from easy to hard in each training session. If the adaptive system is very responsive to the performance threshold, and most are, this is very easy to do.

". . . THE DEEPER QUESTIONS. . . ."

I think this just touches on the kind of issue we must discuss during this conference. We must deal with some of the deeper questions of training validity. What makes an adaptive system a good or a bad adaptive system? I know that we are all beyond the point where we think that just because a system is adaptive it is going to be good in terms of training effectiveness. Of course it isn't! There is no inherent reason why an adaptive trainer is going to be an effective trainer. There are too many variables in the training situation, aside from the question of whether the training system is adaptive, that bear on training validity and effectiveness. I think that the adaptive technique is an extremely powerful one, but I hate to see it relied on to the neglect of the other factors which we know are critical in the training situation.

Mr. Green

Mr. Green: I have a comment pertaining to your system diagram in Exhibit 8. You showed the performance measurement block as the only input to the adaptive logic block. I felt that there might be an input to the adaptive logic block from some point after the trainee, specifically from the controls block. There are situations where, even without performance measurement, you would like to have an input into the adaptive logic block after the controls or controlled element block, then further down the line you would have significant performance measurements which would also be injected.

Dr. Kelley: I suppose so, but I would tend to think of such inputs as performance measurements in themselves.

Mr. Green: The equations that you showed before may be implemented in a performance measurement block but on an automated preprogrammed basis. In the procedure that I am talking about, those equations need not be implemented at that point in time. In other words, the problem can be run on a time-sequence basis where performance may have significant milestones. This is especially desirable when the training task is not very complex, for example, when you are teaching the trainee to recognize signal patterns. You might just score the sequence on a time basis, and then later in the training program, you might have the trainee respond with a key word or something of that nature. You would then be in a position to make performance measurements.

DR. KELLEY: But, it is not adaptive until you get through that sequence. As long as you give time-sequencing, you are not adapting, and this is a diagram of an adaptive system.

MR. GREEN: The only place where it would be appropriate would be if the student had to make a recognition milestone input to the system. Then it would not be purely time sequencing. However, in that case, you are monitoring and not necessarily making a performance measurement.

There is also the possible system where it is desirable to maintain a constant in the performance measurement block. Here simultaneous inputs to the adaptive logic block would be required from other source blocks with a controlled feedback loop to the performance measurement block in order to maintain a constant of performance measurement or a narrowly constricted performance measurement band.

DR. KELLEY: Performance measurement should be interpreted very loosely, because it may be just the touch of a switch that is being measured. I agree with you that there should be the possibility of feed-in from different points to the adaptive logic. But, I think that time-sequencing per se is not part of the adaptive feature, even though it may be used in a lot of training devices.

". . . A CHANCY BUSINESS. . . ."

DR. MCGRATH: I don't share your enthusiasm for the use of secondary or loading tasks as adaptive variables. Secondary tasks can sometimes be used successfully to improve the signal-detection performance of a sonar or radar operator. The secondary task increases the operator's activity level and helps him maintain vigilance during an otherwise monotonous watch. But, you can never be certain of

Dr. McGrath

how performance will be affected. Sometimes the use of a secondary task will enhance performance on the primary task; sometimes it will degrade performance. In either case, it has important effects upon the operator's approach to the primary task. These effects are bound to have implications for training, but they depend upon the relative difficulty levels of the primary and the secondary tasks, the payoffs or importance that the operator perceives in the two different tasks, the relative work loads involved in performing the primary task versus the secondary task, and a whole host of variables that make it very difficult to predict how the secondary task will influence primary-task performance. Manipulating primary-task difficulty by manipulating secondary-task loading is a chancy business. Consequently, although I have never performed any adaptive training studies using secondary tasks, I know from using secondary tasks on other kinds of problems that they should be used only with the greatest discretion.

Dr. Kelley: Surely, a secondary task has a different significance in a vigilance situation than it has in a tracking situation. I've never implemented an adaptive training system for sonar operators, but in using adaptive techniques to study their performance, it never occurred to me to use secondary tasks as an adaptive variable. The only thing we thought about was how we could systematically vary the difficulty of the stimulus material itself. Even there it gets hairy, because these signals vary on several dimensions of difficulty, and it is not an easy task to scale difficulty.

Dr. McGrath: That's true. Moreover, the transition from the completely performable task to the impossible task occurs very rapidly when the critical signal parameters are varied. Therefore, in most signal detection tasks, difficulty is hard to scale because performance is sensitive to minute changes in task variables.

Dr. Matheny: That can also happen in a closed-loop tracking task. For example, introducing a transmission lag into the system will cause performance to deteriorate in a hurry. But the important point is that the difficulty introduced into a task by a given variable may vary as a function of the task into which it is introduced in some interactive sense. You have to be very careful how you do it.

We cannot say that two units of Variable A give one unit of task difficulty of the same type given by two units of Variable B. In other words, the change in task difficulty is unique to the loading variable and unique to the primary task. I think this is relevant to what you're talking about, Jim, and it is a problem that has to be attacked or talked about at this conference.

Dr. Matheny

DR. KELLEY: The general way you stated the problem is also important, Guy. We can't say that difficulty caused by one variable is equivalent to that caused by another. The change in difficulty caused by one adaptive variable is not equivalent, as far as training is concerned, to the same change in difficulty brought about by different adaptive variables. And I have never been able to get around the fact that, in an adaptive training system, both the performance measurement and the adaptive variable affect what you teach the trainee to do; training is affected both by what you measure and by how you make the task tougher or easier. I cannot see how you can disentangle these two fundamental aspects of the task. I think these remarks are relevant to the fact that we talk about task difficulty as if it were the significant variable, but there are lots of different ways to make the task harder, and they are by no means equivalent, and they do not have the same training implications. Do you have a comment, Paul?

". . . ADDING A NEW DIMENSION. . . ."

DR. CARO: It was more appropriate earlier in the discussion concerning the use of secondary tasks. I was going to comment on the problem of distinguishing between whether you are changing task difficulty or changing the task. Rather than just making the original task more difficult, you might be giving the trainee something different to do.

Dr. Caro

DR. KELLEY: Would you elaborate on that?

DR. CARO: Well, I suspect that in some cases where you are introducing secondary loading tasks, such as by adding a communications or malfunction problem in a flight simulator, the task becomes different. I don't know whether it is perceived by the trainee as different or whether the difference can actually be measured, but it is not just making the original tracking task more difficult. You are adding a new dimension to the original task, not just requiring harder work. I think this is something beyond just changing difficulty.

Dr. Kelley: In one sense, there is no such dimension as just difficulty, because whenever you change an adaptive variable, you create a different task. Maybe, by throwing in a secondary task, you might be training some kind of time-sharing behavior rather than, say, the skill that is involved in the primary task per se.

Dr. Caro: Yes, there is the time-sharing element of it. There is also the consideration that once you change the task by superimposing a different type task, such as communication that must be time-shared, then you may be starting a new task again at a low difficulty level and need to adapt that task up to some higher level before you go on to another task. So, you are going in discrete steps from learning one task to learning another rather than extending the learning of the basic task that you are trying to train.

Dr. Kelley: That ties in with the idea that most complex performance involves a number of different kinds of tasks, all of which must be used to achieve training by an adaptive device. You might need as many adaptive variables as you have different kinds of activities that you are teaching a person, and each one of those might vary along a continuum of difficulty.

Dr. Caro: It gets messy very rapidly.

Dr. Kelley: Yes.

". . . THE FUNDAMENTAL DILEMMA. . . ."

Dr. McGrath: Is this the fundamental dilemma of adaptive training? By using adaptive techniques, are you teaching the student merely to adapt to the particular variable that is being manipulated? Does adaptive training narrow the scope of training?

Mr. Green: I don't think it really narrows training. You are changing the task from several dimensions to a single dimension, or you are integrating several functions to a single function. For otherwise, your performance measure becomes just as complex as your task and subtasks.

Dr. Kelley: In many cases, performance measurement is a function of many different behaviors that are required to perform the task. For example, we can adapt tracking tasks easily by varying something simple like the amplitude of a forcing function. When you go from the simple tracking task to a real-world problem, such as operational flight training, you find out that tracking is only a small part of what is involved in flying an aircraft safely and skillfully. All of the other performances that go into flying an aircraft are part of the training process, and they have to fit into your concept of performance measurement. For example, instead of simply measuring RMS tracking error and saying, "This is performance measurement," you reach a point where you say, "Well, if the student is individually performing Tasks A, B, C, D, and E correctly (which might include tracking, communications, and display-monitoring tasks), then he is performing the total task appropriately, and he is ready for the next more difficult task." The next task might involve adding another variable or changing the required maneuver. There would be many variables adapting at the same time; I don't know how it would be handled practically, but conceptually there is no reason not to do it. In developing adaptive training systems for practical use, we have to look at the parts of performance that count. We cannot just take the parts that are convenient to measure and say, "That is our performance measurement." Our performance measurement has to represent some reasonable approximation to what we are trying to teach.

DR. MCGRATH: That is true, but the narrowing of the scope of training in an adaptive system would not be a consequence of using a simple performance measurement; rather, it would be a consequence of using a simple adaptive variable.

DR. MONTEMERLO: You brought up this dilemma that adaptive training possibly teaches the student merely to adapt along one variable. This is one reason why the use of secondary loading tasks is not really realistic. I am not talking about laboratory experiments but about pilot training or other real-life tasks. The object is not to increase difficulty just for the sake of increasing difficulty and making the task adaptive but to make the task more realistic as training goes on. The training process must approach the real-world skill you want the trainee to learn. I mean we are not worried about getting a fine scale of difficulty that would produce a nice graph; we are worried about teaching the criterion skill. So, I think the use of secondary loading tasks is more relevant for use in the basic research laboratory than for practical use in the field.

". . . WHAT ARE THE GERMANE CRITERIA. . . ?"

DR. McGRATH: I would like to comment on that. All of us can see the obvious methodological advantages of adaptive techniques in laboratory experimentation. So, we can evaluate these techniques as research tools. But, can we evaluate the effectiveness of adaptive techniques as practical methods for training operationally relevant skills? It does not suffice to say that they produce more sensitive or linear learning curves. That is fine for research applications, but that is not necessarily a criterion for a practical training application. So, what are the germane criteria? If adaptive training is good, where can we see its benefits? Will the trainee ultimately reach a higher performance level? Will he reach a given level faster? Will transfer of training to the operational task be greater? Will the student retain his trained skills longer? Will it cost less to train him? Will there be an increase in the student's motivation or his acceptance of the training situation so that the attrition rate will be reduced? Will adaptive methods enable us to train students with lower-level aptitudes? These are the practical criteria for operational training, even though none of them applies to the utility of adaptive techniques as research tools. I think one of the valuable outputs of this conference might be a specification of the appropriate criteria for evaluating adaptive training techniques.

MR. FLEXMAN: Obviously your list of criteria could equally apply to any training technique.

DR. McGRATH: Yes, of course.

MR. FLEXMAN: One of the reasons why we are vigorously pursuing the development of adaptive training methods is to achieve better control of the training situation. Adaptive training will undoubtedly use automatic control of the task difficulty. When we apply adaptive logic, we have precise control of the learning experience for the first time. In the existing situation, good learning principles and methods are normally applied as a function of the motivation and background of the individual instructor, so learning environments vary considerably because of differences among instructors. The instructors almost always are the biggest source of variance in the efficiency or the adequacy of training programs. When and if we develop a system or a logic that is appropriate for a particular learning task, we can structure the learning situation in such a manner that all students will receive equal opportunity for learning. Adaptive training is a step towards the more formal structuring of the learning environment.

Miss Knoop

MISS KNOOP: Sooner or later, these discussions always get around to performance measurement as being one of the most significant problems we have to face before we can ever implement adaptive training. Performance measurement is also basic to Jim McGrath's questions about the value of adaptive training. To answer such questions, we must be able to measure performance objectively and validly. In addition, I wonder if anyone has given any thought to the relationship that holds between the type of adaptive logic that is used and the type of performance measurement that we must employ. For years, people have worked in the area of performance measurement to different ends. We can measure performance for a number of reasons, and adaptive training is just one more reason. I think that perhaps the objectives we seek in adaptive training may be laying the groundwork for the new and broader objectives we should be seeking in the general area of performance measurement.

DR. KELLEY: I believe that one of the things that adaptive training does for us is to focus our attention at the very outset on the problem of performance measurement, because you cannot make the system adaptive until you can measure performance. Looking at existing training systems, almost all of which endeavor to assess performance in some way, you see that the trainees normally get some general score to indicate that their training is completed and, presumably, that they have learned whatever it was that they were being trained to learn. You very quickly realize how seldom good performance measurement is done in a lot of these systems. The problem is absolutely fundamental. There is no way of getting a decent adaptive training system without a decent performance measurement.

". . . THE QUESTION HAS NO ANSWER. . . ."

The criteria that Jim McGrath outlined are not criteria for evaluating an adaptive system. They are, in fact, criteria for telling whether Training System A is better than Training System B. And one or both training systems may be adaptive. The question of whether adaptive training is effective, compared with fixed training, has no answer. Whenever you test this question in the laboratory, you are always comparing a particular adaptive training system with a particular fixed-difficulty training system. One or the other may be better, but this says nothing about adaptive training as such.

My predilection for adaptive training is not based on laboratory data or anything of that sort. It is based on logical considerations and the fact that I observe skilled instructors changing the difficulty level of the material they present the trainee as a function of how well that trainee performs. That is why I say that adaptive training is effective and not because I have graphs showing progress on an adaptive versus a fixed system. You simply cannot assess adaptive training per se on the basis of experiments which purport to compare adaptive versus fixed training. What you can do is ask, "Is this particular adaptive system, in this or that situation, better than some particular fixed system for the training of a certain skill?" Which is to say that you need criteria of effectiveness which go beyond the particular systems you are dealing with. If the criteria did not go beyond the specific systems tested, if they were not more general, you would have no way of comparing two different training methods in terms of effectiveness.

Mr. Wood: One of the most fascinating things about the technology of adaptive training is the fact that it provides an opportunity to study the training implications of changing task load during the learning process. It is very appealing to have a training technique which allows even nominal control and study of task load. For years we have conducted part-task and whole-task research with little recent success. I see adaptive training as perhaps providing a new approach to many research issues.

Dr. McGrath: Nevertheless, we should be able to specify what practical benefits can be expected from adaptive training. In other words, when and why should you develop an adaptive system? When you have low-skilled people who must learn complex tasks, is that when you use it? Or, when you want to reduce training time, is that when you use it?

Dr. Kelley: When you want to reduce the number of instructors, would that be a good time to use it?

Dr. MGrath: That is the point. We should be able to answer such questions. What are the training problems that adaptive techniques are best suited to solve?

Mr. Wood: Adaptive training provides a technique for individualized instruction of perceptual-motor skills. This is something we haven't had before.

". . . DOING AWAY WITH THE INSTRUCTOR. . . ."

DR. KELLEY: I think Jim McGrath's point is well taken. When do you make the system adaptive, and why? To me, it is self-evident that you can get rid of instructors when you automate their adaptive processes. So, that is one time to go in the direction of automatic adaptive systems.

DR. MCGRATH: When you want to reduce instructor variance?

DR. KELLEY: When you want to get rid of instructors, period. They are expensive as well as variable. With adaptive training you can automate a major function that in the past has been performed by human instructors.

MR. STEWART: It does not necessarily follow that an automated training system has to be adaptive in the sense of providing a continuously varying task. The minute you put a man in front of a programmed teaching machine you have already met the criterion of eliminating or minimizing the need for an instructor. But doing away with the instructor does not hang on the system's being adaptive as you have defined it.


Mr. Stewart

DR. KELLEY: Now, wait a minute. Let's be sure we are clear on this. Does it, or doesn't it? You eliminate the instructor, and now you have a student facing a console and doing some kind of work. You can give him a task that is unvarying, or one that is preprogrammed for a fixed sequence, or one that is adaptive, in which case the sequence is dependent upon how well he performs: there are the three options. Let's just consider the second two, which are the important ones for significant training problems. Either you provide a preprogrammed sequence of tasks —everybody goes through it, including the potentially slowest learner, and everybody gets the same problems according to a fixed schedule—or you provide the students with different problems appropriate to their individual levels of skill. A corollary is that all students go through a training program which is of the same duration, you see, or they go through a training program which brings them to a particular level of skill in whatever time is required. It seems to me that either the difficulty levels of the task are preprogrammed or the task is adaptive. I know no other alternative.

Dr. McGrath: But either method can reduce the need for instructors. I believe Bill Stewart's point was that, if the chief objective of adaptive training is to eliminate instructors, this simply calls for some programmed learning procedure. To reduce the role of human instructors, you need automation, not necessarily adaptation. So, surely adaptive techniques seek more important goals, such as the individualized tailoring of learning tasks that you have stressed.

Dr. Kelley: Yes, but if you simply automate the system without making it adaptive, then you have a fixed program. You can eliminate instructors by providing an automated, fixed-program sequence; there is no doubt about that.

Dr. McGrath: So, why make it adaptive?

Dr. Caro: We are making it adaptive because we want the decision function to be wholly rational. The average instructor is incapable of reacting rationally in all situations, but a properly programmed computer could do so. The adaptive training model provides a possible way to make at least one instructor function wholly rational. Adaptive training is really instructor simulation.

Dr. Kelley: I thought I had answered that question by implying that it is not appropriate to give everybody the same sequence of problems in teaching a complex and difficult skill. If you reject that point, then we would disagree; if you accept it, you accept the need for some form of adaptation, be it via an instructor or an automated adaptive system.

Dr. McGrath: Chuck, if I am an iconoclast, it is because I want this group of experts to state explicitly why they are developing adaptive systems and to specify what problems they expect adaptive techniques to solve.

Dr. Montemerlo: We train people to do many different things; some would call for adaptive training, and some would not. What times would adaptive training be more useful than some other type of training? Is it only in perceptual-motor skills? Is it only in one class of perceptual-motor skills? Is that what you are asking?

Dr. McGrath: Partly. Suppose someone asked, "When should one use overlearning as a training technique?" Well, we could answer that. Overlearning should be used when long-term retention of skills is required in the absence of interim practice. There may be other reasons, but that is the main reason for using that technique. Can we provide a comparable reason for using adaptive techniques? What specific problems will they solve that other methods will not?

MISS KNOOP: Well, I have a guess or a trial balloon. I have always felt that adaptive training will help us to reach a point where we can achieve overlearning for individual students sooner than we can with ordinary techniques of training.

". . . A CARROT IN FRONT OF A DONKEY. . . ."

MR. STEWART: Here is another problem: if I were to perform an adaptive task, I would have no knowledge of whether my performance was good, bad, or indifferent. It masks motivation.

DR. KELLEY: Let me disagree with you very decidedly on that point. The performance measurement may no longer be an index of how well you are doing because it is constant, but there is something else that can be used, the index of how difficult the system is, the adaptive variable. Keep that available to the trainee; it is a measurement of how he is doing. It provides exactly the same motivation as the performance measurement score he used to have before. The only time you run into the problem that you mentioned is when people make the mistake of leaving the adaptive score or difficulty index off the display console. Then the student has no way of knowing how well he is doing. I believe that one of the real advantages of adaptive training is that it has such a good effect on motivation because the student has a very accurate gauge of the difficulty of the task he is performing. He can see how well he is doing with real precision. He can see progress from trial to trial and day to day, for example, that he cannot see with the best scores that we have in fixed tasks. Whatever you did before with a performance measurement, you can now do with an adaptive score.

DR. MONTEMERLO: But if you keep the student's score constant, it is like putting a carrot in front of a donkey. No matter how fast the donkey runs, he can't get to the carrot. Have you experienced having any of your students getting frustrated because they could not do any better or get a better score?

DR. CARO: They are always getting a better score, in a sense.

DR. KELLEY: But the *score* is the status of the adaptive variable and is not kept constant; pursuing a goal in terms of adaptive variables is not pursuing an unreachable carrot. It is true that the student cannot experience complete mastery of the task. The only time this has been a problem was in working with astronauts who found an adaptive system inherently frustrating because they could not attain the kind of mastery over it that they could over some fixed set of maneuvers. However, they could improve their performance regularly, and they could also see a gauge that showed how their skills improved. I really think that the reason the astronauts got frustrated was that a couple of firemen trained on the adaptive task were so much better than they were at controlling a simulated spacecraft.

Mr. Flexman: That would do it.

Mr. Wood: At least to some degree, an adaptive task is frustrating to a student. The student never gets closure on the task, because as he gets better the task gets more difficult. Traditionally, he expects to do better as he learns. Of course, in the adaptive situation, as Chuck Kelley has suggested, there are many forms of feedback that can be provided. I think it is appropriate to ask what effect these additional forms of feedback have on training because they are somewhat artificial. They are not directly related to the criterion task because the operational criterion task is rarely adaptive, it is more *fixed* in nature. When the operational transfer is made, artificial forms of feedback are removed with possible effects on performance level. Possibly one of the advantages of some form of fixed practice is the fact that the person can use the task-inherent feedback for motivational and learning purposes.

". . . QUESTIONS AND . . . ANSWERS FOR THE RECORD. . . ."

Dr. Kelley: Let's take these questions and write out our individual answers for the record. When are adaptive training methods called for? What problems do adaptive training systems solve? What times would you use an adaptive trainer instead of a fixed trainer? What are the reasons for using adaptive training methods?

Mr. Flexman: Adaptive training methods are called for: (1) when the task to be learned is of sufficient difficulty to require a significant amount of time to achieve mastery, (2) when task improvement is a linear, or progressive, phenomenon, (3) when the training situation lends itself to automated controls, (4) when the subject population can be controlled to the extent necessary, and (5) when the training task bears sufficient resemblance to the final behavioral model to have a reasonable transfer value.


Mr. Norman

MR. NORMAN: Adaptive training systems are appropriate primarily for the "big Skinner box" approach to promoting precise control over learning of the task, assuming that terminal expected performance can be defined and that important pitfalls during acquisition have been identified. This should make training more efficient by reducing the amount of time spent in nonessential-to-final-task situations. There is some evidence that adaptive training can be used to attain the criterion skill faster, with better transfer performance and retention, than preprogrammed training.

MR. STEWART: Adaptive training systems, Kelley style, are probably most useful, I think, in situations requiring considerable overlearning and high retention over time. They are obviously quite useful in equipment design, as in Chuck Kelley's work on head-up display design for the F-111 aircraft. They are most appropriate in systems where variables can be closely, perhaps infinitely, controlled, but application to part-task/whole-task grouped learning patterns may not be amenable to the close control over levels of difficulty that seems essential to the technique. The important problem is the relevance of the adaptive variable used as a forcing function (turbulence, for example) to the task for which the student is being trained.

DR. KELLEY: Adaptive training methods are called for when training is already adaptive but instructor controlled and when it is desirable, for reasons of cost, standardization, or (when it can be demonstrated) training effectiveness, to mechanize the instructor's adaptive function. These methods are also appropriate when it can be shown that training time is significantly reduced or that a higher level of skill is reached in the same time by replacing a fixed schedule with an adaptive training schedule. The criterion for using adaptive training should be cost-effectiveness, broadly interpreted and applied.

DR. LAUBER: Adaptive training techniques are justifiable whenever the goal is to automate the training situation, i.e., to replace the

Dr. Lauber

instructor with a machine. It has been pointed out that a good instructor utilizes adaptive techniques in that he tailors the training situation to the individual trainee. The primary problem, it seems to me, is to formalize the decision logic used by a good instructor and then to implement this logic in the automated device. This approach, hopefully, will allow at least a standardization, and perhaps an optimization, of the training situation.

Dr. Caro: Adaptive training requires specification of a decision model which relates observed or measured performance to task difficulty or complexity. In the real-world model, i.e., a human instructor working with a single trainee, decisions are often made on the basis of the instructor's subjective evaluation of trainee performance, and any change in the tasks that the instructor makes in response to that evaluation is an expression of his gut feel. His decision model is not wholly rational. By removing the instructor, a wholly rational decision model can be substituted for him. Intuitively, I must assume that adaptive training then has at least the potential of improving upon the real-world model and should be employed whenever an appropriate decision model can be developed.

Mr. Mansfield

Mr. Mansfield: The main reason for choosing adaptive versus fixed training methods is to arrive at a desired result in a complex training task at less cost by reducing the number of instructors and by reducing the trainees' total time in the training program, thereby increasing the trainees' (and instructors') time that can be made available for performing the operational task as skilled personnel. I emphasize *complex* as in present pilot training programs where performance evaluation and the increment and scheduling of additional training tasks or difficulty levels are dependent in large part on instructor judgment. I also emphasize that this statement is contingent on many unanswered questions concerning how effectively and to what extent adaptive training can actually replace fixed training.

Mr. Hall: The adaptive technique allows for the formulation and application of decision rules for instructor use in changing the difficulty of a current problem or in presenting new materials to students. The question of transfer of training in some ways becomes secondary; so long as the basic task is mastered and the terminal performance goal is achieved, the method can be considered effective. Adaptive training is called for when the computer can effectively substitute for the instructor and when the task is of such a high order of difficulty that it cannot be mastered unless it is broken down into component parts. What is needed for implementation is an analysis of the task or tasks and then the development of a strategy for teaching them. Adaptive training techniques as defined by Chuck Kelley probably will not often be required for training practical operational tasks.


Mr. Hall

MISS KNOOP: The development of adaptive training methods is linked to the goal of automating many instructor functions (in flight simulator training, for example). One reason for doing this is *not* to replace the instructor but to relieve him of time-consuming routine duties, so that he can use his teaching skills more effectively. It is impossible or impractical to automate some instructor functions. So let's not talk of getting rid of him.

If we desire to automate a large portion of the tasks normally performed by the instructor, we must include his capability for observing and evaluating performance and making a judgment about what to teach next or what to have the trainee practice next in order to accomplish the training objective most efficiently. Interpreted, this means we need a capability to automatically evaluate or measure performance and to adaptively train. Using computers, we should be able not only to automate this instructor function but also to accomplish it more reliably, more objectively, and to a greater extent or depth than could any human.

DR. MCGRATH: The availability of a performance measure, an adaptive variable, and an adaptive logic is a condition for the use of adaptive training, but of course it is not a reason for its use. As a hard-line empiricist, I'm frankly unimpressed by the circular reasoning that adaptive methods are good because good teachers use adaptive methods. So why use them? Well, automated adaptive training may be used for the purposes of reducing the need for human instructors, standardizing training procedures, making the decision function more objective and rational, or achieving greater precision of control in presenting tasks and instructional materials. These are important goals, but they can be met by other methods of automated, structured, or programmed training that are not necessarily adaptive. The purposes that can be served uniquely by adaptive training must be those which are served by maintaining a continuously closed loop between the student's performance output and the task-stimuli input or which can be accounted to Chuck Kelley's creative idea of a stabilized performance criterion. My impression is that nobody truly knows what clear benefits are offered by adaptive training and only by adaptive training. Every potential benefit of adaptive training that I can suggest as being unique to this technique is properly a hypothesis that requires empirical testing.

Mr. Green: First, I do not believe a generalized concept can be evolved where the answers to these questions are cut and dried. I feel that the technique of adaptive training must be implemented on the basis of the training problem objective, that is, the specific type of training to be accomplished and the relative degree of complexity of control of the training problem desired. Second, I feel the focal point is the requirement for monitoring and evaluation. Where preprogrammed instruction suffices, adaptive training should be minimized. Where a sterilized objective in training is desired, adaptive training surely may be implemented as a programmed learning system.


Mr. Weekes

Mr. Weekes: What problems do adaptive training methods solve? They provide the assistance required to prevent calamity in the real-world situation. For example, in teaching a boy to ride a unicycle, some arrangement (safety wheels) must be made to prevent instant failure. They assist in training division-of-attention or time-sharing when the real-world job requires the performance of more than one task. For example, to add juggling to the unicycle rider's act, the safety wheels must be reinstalled, although their use had been discontinued when control of the unicycle was the total task. A specific application: it furnishes a way to relieve a flying trainee of as many of the control details inherent in the total task of maneuvering and navigating an aircraft as necessary to enable him to absorb instruction concerning a specific aspect of the total task. Another way of saying it: adaptive methods can be used to lighten the housekeeping duties which might divert a trainee's attention, in order that the trainee may better concentrate on the specific item under instruction.

DR. MONTEMERLO: In the introduction to the December, 1969, issue of *Human Factors*, Jim Regan stated that surprisingly little work has been done on what specific aspects of adaptive training theoretically make it more effective than nonadaptive training. In the same issue, Chuck Kelley stated that the reasons are obvious and need not be researched. Hudson observed that adaptive training is not necessarily efficient in itself, it is good adaptive training that is efficient. Yet, this holds true for any method of training. So, what is it about adaptive training that makes it more efficient, if it is more efficient? And more efficient than what? Surely, testing adaptive training versus fixed-difficulty training is shooting down a straw man, for nonadaptive, rigidly fixed, unvarying training procedures are not typical of modern training programs. I think that adaptive training methods are most probably called for in perceptual-motor tasks that are too difficult for someone to handle at the outset, such as the task of controlling a helicopter. But, generally, the systems approach must be taken to identifying training applications, using behavioral objectives and task analyses. If a need for adaptive training shows up, then use it. But, it is a mistake to work in the other direction, starting with the assumption that adaptive techniques will be used and then trying to fit them into the program.

DR. MATHENY: We don't have enough empirical and definitive data to answer the question, "When is adaptive training called for?" We can list certain requirements to be met before we can consider using it. These, as I see them, were outlined in my *Human Factors* article on the effective time constant as an adaptive variable (Matheny, 1969). These criteria for an adaptive variable are: (1) capable of being described and quantified, (2) related to task difficulty, (3) capable of being varied in a systematic way, (4) if used in a device for training for transfer, does not inhibit or interfere with transfer, (5) acceptable to the trainee, and (6) capable of being adjusted over a range commensurate with the trainee's skill. As a general answer to the question, "What are the reasons for using adaptive training?" I think it is to reduce variability—variability in instructor's guidance to the student and variability in matching the task difficulty to the student's skill when seeking to keep the student challenged and progressing.

Session 2. Flight Training Applications

DR. PAUL W. CARO, *Chairman*

DR. CARO: This afternoon we will consider some of the ways adaptive techniques can be applied to training aircraft pilots. I'd like to begin by asking Milt Wood to describe his recent experiment concerning the use of adaptive methods for training pilots to control pitch and roll.

". . . A QUASI-ADAPTIVE METHOD. . . ."

MR. WOOD: The purpose of this study was to compare two methods of training naive subjects in a two-dimensional, second-order, tracking task that corresponded to the control of pitch and roll.

The first method was an automatic, self-adjusting technique, such as that described this morning by Chuck Kelley. Task difficulty was adaptively varied while the subject's error was maintained at a constant average level of moderate magnitude. This method is appealing because of the many reasons already mentioned, such as the ability to maintain the task load at a level that is neither too difficult nor too easy. If it is too difficult, the student loses control and gets little useful practice. If it is too easy, he sits and worries about why his girl friend was mad at him last night, and he doesn't work very hard.

The second method employed fixed-difficulty levels which increased systematically during practice. The fixed-difficulty levels were determined by the median difficulty levels in each session on the continuously adaptive task. This might be considered a quasi-adaptive method, because the difficulty levels were fixed and the subject's error was allowed to vary, but the difficulty level was increased systematically as training progressed. This method should retain some of the advantages of the fixed-difficulty approach to training. For example, the student can be motivated by a decrease in error as he learns, and he can use the feedback that is inherent in a fixed-difficulty task. Therefore, the results from the continuously adaptive task were used to establish a schedule of difficulty changes for the fixed-difficulty task so that comparisons could be made between the two approaches and the advantages of the two approaches could perhaps be combined.

Fourteen male college students served as subjects. All were in good physical health, possessed 20-20 vision, and had no prior experience in the control of second-order vehicular systems. They were randomly assigned to two groups of seven subjects each.

DR. KELLEY: You did not match their initial skill levels when you assigned them to the groups?

MR. WOOD: No. I certainly wish I had done so, and would do so in any subsequent experiments.

An 8,000-word, general-purpose digital computer supplied the dynamics for the second-order compensatory tracking task which was free to move in the dimensions of pitch and roll. The subject used an off-on type of side-arm control to maintain a constant null position of pitch and roll as displayed by an attitude-director indicator. Task difficulty was defined as the average amplitude of the forcing function which moved pitch and roll away from a null position of zero error. The subject's display and side-arm control were contained in a realistic cockpit assembly which isolated the trainee from distracting environmental factors. Exhibit 10 shows a block diagram of the experimental task.

DR. KELLEY: Was the forcing function a thrust disturbance?

MR. WOOD: It was not a thrust disturbance but simulated an outside force acting upon the vehicle. It was integrated and smoothed; so it was not a discrete disturbance. The display acted as though the vehicle were flying through rough air.

DR. KELLEY: One more question. How often did the thrust come in?

MR. WOOD: One of the two pa-

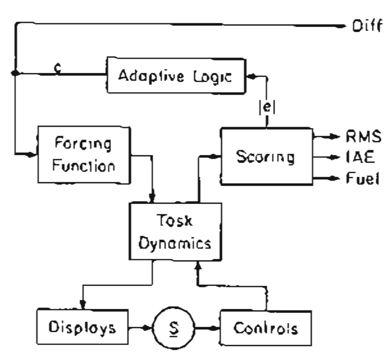


Exhibit 10. Block diagram showing functional components of experimental task.

rameters, either pitch or roll, was selected for perturbation every two seconds. The subject's task was to maintain pitch and roll within 11° of null.

Dr. Matheny: How did you arrive at 11° as the threshold value?

Mr. Wood: I blue-skied it, using the results of Chuck Kelley's research, some preliminary data, and my past experience as a pilot. However, it was quite a reasonable threshold value for this task.

Dr. Caro: Were the subjects instructed to maintain the error at zero, or were they told simply to maintain the error within 11° of zero?

Mr. Wood: They were told to maintain null error.

Mr. Flexman: Did they know the tolerance threshold?

Mr. Wood: They did not know the threshold. Of course, they knew the continuously adaptive task got easier or more difficult. It was quite obvious to them that if their error became high, the task got easier, and vice versa. In fact, it occurs to me that this might be a good way to identify malingerers. Some people don't like to work, and one way to avoid hard work in a continuously adaptive task is simply to let the error go up and the task, therefore, become easier.

". . . A FEW MORE DETAILS ON THE METHOD. . . ."

Dr. Kelley: You had a two-axis tracking task. How did you combine the errors for determining adaptive changes?

Mr. Wood: We picked the axis of larger error as our measure of performance.

Dr. Lauber: Can someone summarize for me the arguments pro and con for summing integrated absolute error in a two-axis tracking task like this one?

Dr. Kelley: There are several ways to score two-axis tracking performance, and the score that Milt Wood described is a very effective one. The score is highly correlated with RMS error but easier to measure. When the larger error in the two axes is used, the student finds that he cannot let one error grow while he is learning to control the other; he realizes that he must control both axes simultaneously to obtain a good score.

Mr. Wood: Another interesting thing happens. When the score depends on the parameter of largest error, the student tends to direct his attention to that parameter. In a way, the task conditions him to attend to parameters which he might otherwise overlook.

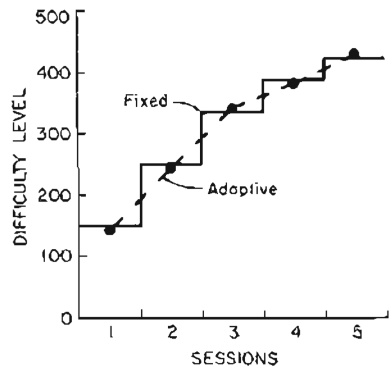


Exhibit 11. Levels of practice as determined by median changes in difficulty for the adaptive group. Difficulty is set for the fixed group.

Let me give a few more details on the experimental method. In order to define the sequence and magnitudes of the increases in difficulty levels under which the fixed-difficulty group would practice, it was necessary first to obtain performance data from the continuously adaptive group. These data were obtained from five daily training sessions, each consisting of eight five-minute trials. The median adaptation level for each session determined the level of difficulty in corresponding fixed-difficulty sessions. Consequently, the median levels of task difficulty were of the same magnitude for both groups, as shown in Exhibit 11. You will notice a very nice exponential progression of the adaptive curve, which is somewhat at odds with the data Chuck Kelley showed this morning. I have obtained this kind of curve with almost all of my subjects, and it has very little variance. I find its stability quite amazing. Anyway, by matching difficulty levels in this fashion, we could evaluate the relative merits of the two training conditions with one of the primary variables—task difficulty—held constant.

Exhibit 12 summarizes the experimental design. In order to provide a basis for the comparison of intra-group performances during training, a series of criterion tests was administered to both groups after specified intervals of practice. A one-minute criterion task of fixed difficulty was given to both groups after every four training trials (that is, after every 20 minutes of practice). A high difficulty level was chosen for the criterion test; it was the level normally attained after four sessions of practice. A two-minute adaptive criterion test was given to both groups after each practice session (that is, after 40 minutes of practice). After five training sessions, all subjects reported for a final transfer session. During this session, each subject was required to perform for five minutes under each of five different fixed levels of difficulty. Except during the transfer trials, all subjects received knowledge of results in the form of

Exhibit 12. Summary of experimental design.

Session	Trials	Adaptive Difficulty	Fixed Difficulty
1	1–4	Continuous △	Fixed at 150
	*	Fixed Criterion Test	
	5–8	Continuous △	Fixed at 150
	*	Fixed Criterion Test	
	*	Adaptive Criterion Test	
2	9–12	Continuous △	Fixed at 250
	*	Fixed Criterion Test	
	13–16	Continuous △	Fixed at 250
	*	Fixed Criterion Test	
	*	Adaptive Criterion Test	
3	17–20	Continuous △	Fixed at 335
	*	Fixed Criterion Test	
	21–24	Continuous △	Fixed at 335
	*	Fixed Criterion Test	
	*	Adaptive Criterion Test	
4	25–28	Continuous △	Fixed at 385
	*	Fixed Criterion Test	
	29–32	Continuous △	Fixed at 385
	*	Fixed Criterion Test	
	*	Adaptive Criterion Test	
5	33–36	Continuous △	Fixed at 420
	*	Fixed Criterion Test	
	37–40	Continuous △	Fixed at 420
	*	Fixed Criterion Test	
	*	Adaptive Criterion Test	
Transfer Session	41–45	Fixed Levels 100–200–300–400–500	

* Same test for both groups.

error or difficulty scores presented verbally at the end of each practice or criterion trial.

". . . SEVERAL INTERESTING CONTRASTS. . . ."

Performance across groups during the training process and during transfer was compared by measures of RMS error and fuel seconds consumed. The results showed several interesting contrasts between the techniques of continuous adjustment and fixed levels of task difficulty during training.

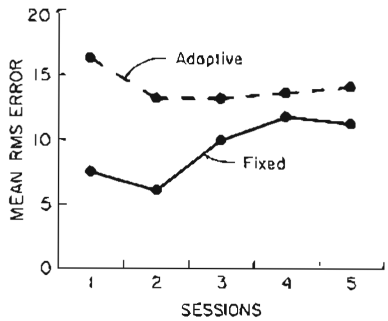


Exhibit 13. Mean session RMS error obtained by both groups during acquisition.

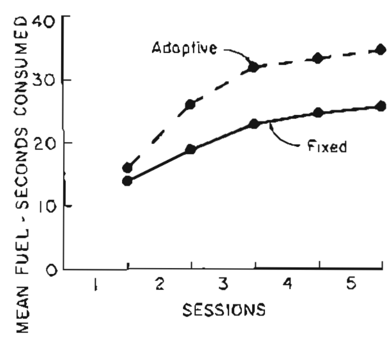


Exhibit 14. Mean fuel-seconds consumed for each minute of flight as averaged over session intervals.

As shown in Exhibits 13 and 14, there was a difference in mean error and fuel consumption between groups across the five practice sessions. The median task difficulty levels were equated between groups; so it was reasonable to expect no significant difference in mean error between groups. This expectation was not borne out. Subjects on the fixed-difficulty task made significantly ($P \leqslant .01$) smaller mean errors during Sessions 1 and 2 and consumed significantly ($P \leqslant .05$) less fuel across all five training sessions than subjects on the continuously adaptive task.

This apparent difference in training effectiveness was further reflected in the results of the fixed-criterion tests, shown in Exhibits 15 and 16. During fixed-criterion performance, subjects trained on the fixed difficulty task made significantly ($P \leqslant .05$) smaller mean errors across all five training sessions and consumed significantly ($P \leqslant .05$) less fuel during the last four practice sessions than subjects trained on the adaptive task.

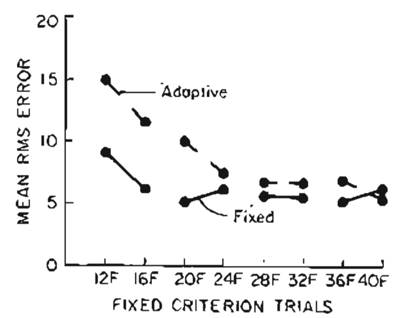


Exhibit 15. Changes in mean RMS error for both groups during fixed-criterion test.

Transfer performance was also differentiated between groups. Although no between-group differences were statistically significant in terms of RMS error, significantly ($P \leqslant .01$) less fuel was used by fixed-difficulty subjects while performing under the middle range of difficulties presented in the transfer trials (Exhibit 17). In general, the overall efficiency of fixed-difficulty practice was greater than that of continuously adaptive practice and was characterized by lower RMS error in addition to less expenditure of fuel during fixed-criterion tests.

I conclude that, within the context of this study, practice under a schedule of increasing levels of fixed difficulty was superior to practice under continuously adaptive levels of difficulty. The superiority of fixed-difficulty practice appears to be a function of several features of the continuously adaptive task which generally reduce training efficiency. One such feature involves a tendency for the subject to reach levels of task difficulty which greatly exceed his momentary level of skill. This situation is brought about through lags in the self-adjust model and the nature of practicing under a random forcing function. When these high difficulty levels are obtained, two events occur which can reduce training effectiveness. First, when task difficulty reaches a level which the subject cannot control at that point in training, his error rapidly increases with a corresponding increase in overall task load. Second, when the adaptive program

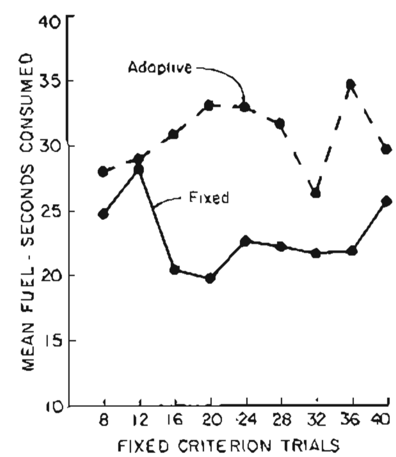


Exhibit 16. Mean fuel-seconds consumed for adaptive and fixed groups during fixed-criterion trials.

senses this out-of-threshold error, task difficulty is rapidly reduced to a level which results in a very low error, and the consequent task load is much less than ideal.

This exaggerated action of the continuously adaptive process produces an inefficiency which is avoided in the fixed-difficulty approach. When increasing levels of fixed difficulty are selected for practice, the result is a gradual learning-curve variation in error for each training session. This technique provides one approach for designing an adaptive system which does not lose efficiency through rapid and extreme adjustments in task difficulty. As Chuck Kelley noted, learning is a gradual process and should be treated in this way by any adaptive logic. That is, the adaptive model should provide a gradual sort of learning experience. In my opinion, this increasing level of fixed difficulty provides just such a gradually changing learning experience and avoids the extreme excursions of continuously adaptive practice.

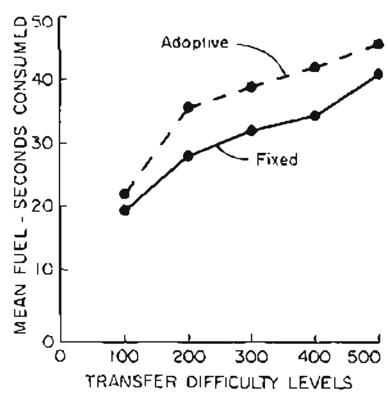


Exhibit 17. Mean fuel-seconds consumed at various levels of difficulty during transfer.

". . . REALLY A SLOWLY CHANGING 'GROUP-ADAPTIVE' TASK. . . ."

DR. KELLEY: Many aspects of this study bother me, but I'll mention only the most important ones. First, I don't think you had a fixed-difficulty task at all. Since the results of performance on the adaptive task were used to adjust the difficulty level on what you call the fixed-difficulty task, you really have a comparison between a slowly changing group-adaptive task—one in which average performance of the group over an extended period is used to control the difficulty level—and a very quickly changing individual-adaptive task, which could go from maximum difficulty to zero in little more than one minute. So, I think you have shown that if the difficulty level of an adaptive task is changed slowly rather than quickly, performance perhaps is better. (I say *perhaps* because I have some questions about your measure of performance.) In other words, if this had been truly a fixed-difficulty task, the difficulty level would have remained the same throughout training.

Mr. Wood: I said the task was quasi-adaptive.

Dr. Kelley: Well, it changed difficulty in discrete steps rather than in continuous variations, but it did change; it changed quite frequently and furthermore changed quite optimally in the sense that you used the measurement of adaptive task performance to decide exactly what changes to make in the fixed-difficulty task.

I also must question your measurement of performance, which really bothered me very much. If you maintain error constant in a continuously adaptive task, then it follows that you cannot use error as an assessment of performance. In your study, you have to use the amplitude of the forcing function as the assessment of the performance of the continuously adaptive group. When you use RMS error as the performance measure in a continuously adaptive task, you have an uninterpretable measure. When the difficulty level changes in the adaptive task, subjects tend to make larger error excursions. Naturally the RMS error will be higher when you permit the difficulty level of the task to fluctuate within the trial, simply because these extreme excursions will be heavily weighted in the performance score. This really doesn't say anything about how well the subjects are performing the task within that trial; it is merely a characteristic consequence of using RMS error as a measure of performance in a continuously adaptive task.

But, I gather from the results of the criterion trials, which are the best data you have, that there was indeed a difference in favor of what I will call the slowly varying adaptive task as opposed to the quickly varying adaptive task.

Dr. Montemerlo: You say there were really two adaptive-training groups in this study? Well, I thought that adaptive training required the task to change, either by becoming easier or by becoming more difficult, as a function of the performance of the individual. Now, one group in this study experienced a change in task difficulty as a function of their individual performances. But the other group did not receive individualized adaptive changes. The task difficulty, although staged upwards, was the same for everybody. So it really wasn't adaptive training that the second group received.

DR. KELLEY: The second group did not receive individualized adaptive training; they received an optimized progression of difficulty levels, a kind of "group-adaptive" training.

MR. WOOD: It was stereotypic adaptive change.

DR. MONTEMERLO: It isn't much different, after all, from what is usually done in the lecture method. The lecturer normally attempts to stage his material to the learning pace of the average student. I thought one of the points of efficiency with adaptive training was that instruction is paced to the individual rather than to the average of a group. Here we have experimental results which showed that pacing instruction to the average of a group was more efficient than adapting to individual levels of skill.

DR. KELLEY: Individual skills were, in fact, measured, and this measurement of adaptive performance determined the stages of the fixed-difficulty task. It certainly was not preprogrammed after the fashion of the lecture method.

DR. MONTEMERLO: It certainly was not adaptive training either.

". . . SENSATIONS OF TASK MASTERY. . . ."

MR. WOOD: Let's call it a multiple-learning-curve approach. An advantage of this approach is that it gives an opportunity for the student to utilize inherent task feedback, which is valuable for both training and motivational purposes. As the student practices under any given level of fixed difficulty and as his skill increases, he experiences a reduction in error and a decrease in subjective task difficulty. This is in direct contrast to the continuously adaptive approach which maintains constant some selected level of error. In this situation, subjective task difficulty remains constant, and the subject does not experience sensations of learning unless additional modes of feedback are introduced. There are many ways to introduce such feedback, such as displaying an index of changing task difficulty, a measure of out-of-tolerance error, or time on target. Even though these forms of augmented feedback may be effective, they are artificial to the basic task and eventually will be withdrawn because the criterion task will most likely behave in a nonadaptive manner. If the subject has been trained in a task which provides some degree of error change during training, however, he should be in a better position to interpret the task-inherent feedback of the criterion task and should not be handicapped by the withdrawal of the augmented feedback information. He will also have had the familiar and expected sensations of task mastery at each difficulty level of the practice task.

Dr. Kelley: The subjects in the continuously adaptive group were not shown their adaptive scores during their performance of the task. Isn't that correct?

Mr. Wood: Right. I provided summary feedback only. At the end of each trial the subjects were given their error and difficulty scores.

Dr. Kelley: But, while they were tracking, they were not shown these scores. I consider it a mistake to omit such a knowledge-of-results display. The subject can get no information from his instruments as to how well he is performing an adaptive task if error is being maintained constant. The presence or absence of a display of the adaptive score during the performance trials makes a big difference. The difference is in the student's motivation, because the display of the adaptive score provides him with immediate knowledge of results. If the student is simply told after a trial is over how well he had performed, the motivational properties of the knowledge of results are negated. He must have such knowledge while he is performing the task.

Mr. Wood: Certainly there was some knowledge of results available to the individuals practicing the continuously adaptive task. I mean, there was some task-inherent feedback. For example, when a subject achieved null error, he knew he was doing a pretty good job because his main objective was to null the errors. Surely, you will admit to that.

Dr. Kelley: But, did you not say that the error score was maintained constant no matter how well they performed?

Mr. Wood: The trial average was constant. During the trial, error increased or decreased as a function of the subject's ability to null the forcing function.

Dr. Kelley: But the average error score had no relation to how well they were truly performing the task, because task difficulty was simultaneously being varied.

Dr. Caro: If nulling the error gave them information about how well they were doing, then, by definition, error could not have been held constant.

Mr. Wood: Error was held constant only in terms of a time average. On a moment-to-moment basis, some degree of feedback was available as the student responded to individual system perturbations.

DR. MATHENY: There is no doubt that an error display can be confusing in a continuously adaptive task, because its behavior depends upon the length of time over which the performance measure is averaged in order to determine a change in the difficulty level. It is entirely possible that the error display in an adaptive task can give the student negative feedback. This can happen.

MR. NORMAN: I might comment on an exception that would prove your rule on knowledge of results. In our study (Lowes, Ellis, Norman, and Matheny, 1968), we gave no knowledge of results. If a student asked how he was doing, he was invariably told that he was doing well.

DR. MATHENY: There is inherent feedback in the system that you described, Milt, and it's quite dramatic, actually.

MR. WOOD: You mean with the fixed-difficulty approach?

DR. MATHENY: No, I mean with your continuously adaptive approach. The student doesn't need a display to tell him how well he is doing on the adaptive task.

DR. KELLEY: I think he does.

DR. MATHENY: That question can be answered empirically.

MR. WOOD: I know that the student's comprehension of task load is quite reliable. I asked subjects if the task seemed any more difficult today than it was the day before. Without exception, they reported the subjective level of difficulty to be quite constant.

DR KELLEY: At any given moment, subjects will judge how well they are doing as a function of how much error is present. Even when the adaptive logic has reduced the objective level of difficulty and made the task simple, if the subjects are doing poorly, they will feel that the task is hard to perform. Now, by the adaptive logic, you keep the subject in a position where he is always doing approximately equally poorly. He can get a feel of how well he is doing only after a very long period of practice, and I don't think it could be a very accurate feel in this two-axis acceleration tracking task.

". . . SOME ADDITIONAL IMPLICATIONS. . . ."

Mr. Wood: Let me conclude with a few more observations. I believe that the multiple-learning-curve approach to training combines the advantages of both fixed-difficulty training and adaptive training. It *adapts* the level of difficulty to the expected level of skill, yet it also provides the motivational and feedback advantages inherent in the fixed-difficulty approach.

When the student's performance score is automatically held constant during practice on an adaptive task, there are some additional implications that are worthy of note. When error is maintained constant during practice, the subject is constrained to a learning situation which may not allow practice in many aspects of the criterion task. In the adaptive task used in this study, for example, the maintenance of 7.5 units of mean integrated absolute error results in an average stimulus deflection of approximately 11°. Because error is merely a statement of stimulus position relative to a null position, a constraint on error provides a similar constraint on the basic stimulus characteristics of the task.

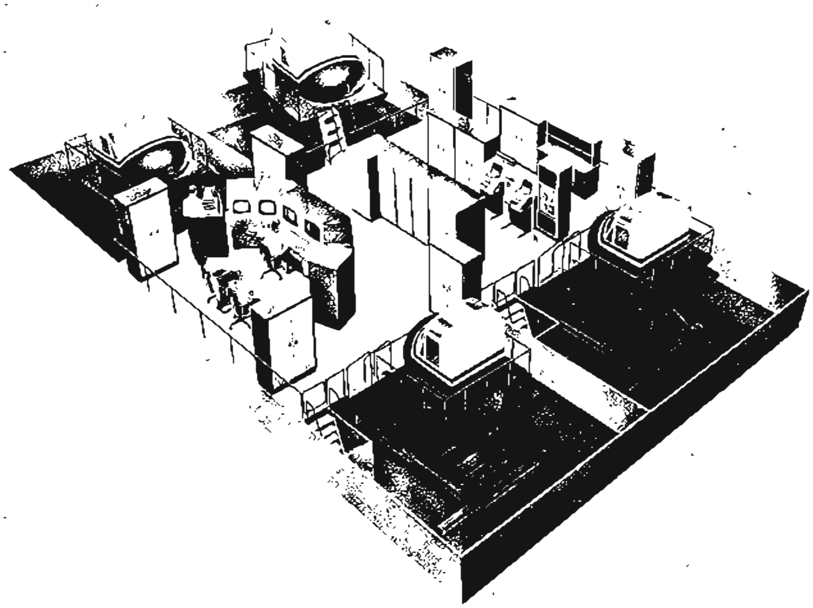
If the student is required to practice under a constrained set of stimulus conditions, it is likely that optimum transfer will occur only when the transfer task is characterized by similar constraints. Ed Hudson (1964) makes this general point with the finding that some medium level of constant error during practice gives the best compromise training for transfer tasks of varying difficulties. Fixed-difficulty practice may provide improved transfer to tasks of varying difficulty because it allows practice under a broader range of error conditions. Error is free to vary, so the student produces a learning-curve effect through practice at each stage of task difficulty. During the early portions of the training session, error is relatively high, and the student has the opportunity to practice under conditions of relatively high perturbation. The task at this point is characterized by large instrument excursions and requires a tracking strategy different from that required at lower error levels. As practice continues and error is reduced (though never to zero), the student has opportunities to practice those aspects of the task characterized by small instrument excursions.

In general, I would like to suggest that a broader range of meaningful tracking strategies can be more efficiently practiced in fixed-difficulty training. This feature, in combination with the more appropriate levels of task difficulty and the better task-inherent feedback, may provide some basis for the possible superiority of the fixed-difficulty method. I think the outcome of this study shows that, under special conditions, a fixed-difficulty mode of practice can provide training equal or superior to that of a continuously adaptive mode. At least, the outcome emphasizes the need to look closer at fixed-difficulty modes of practice which operate in an adaptive manner.

DR. CARO: Thank you, Milt. Your comments were very interesting.

". . . A FEW WORDS ABOUT THE SFTS. . . ."

How many of you are familiar with the SFTS, the Army's Synthetic Flight Training System? Is there anyone here who has never heard of it? Good! Before proceeding further in our discussion of flight training applications, I would like to say a few words about the SFTS, because it represents hardware in which provision has been made for adaptive training. As you know, the SFTS is a multi-cockpit system of training devices which, in addition to simulating helicopters, has several features which relate specifically to adaptive training.



Synthetic Flight Training System.

In 1965, the Army Aviation School asked me to prepare a statement of functional requirements for helicopter training devices that would meet projected training requirements and replace obsolete Army equipment. I determined that one of the training requirements was a reduced reliance upon instructors through the automating of instructor functions. Adaptive training per se was stated as a design requirement because research was beginning to suggest it would be a promising way to make training more efficient. Ed Hudson and Chuck Kelley were saying things about adaptive training that supported my own ideas about automated training and performance measurement, so I asked that adaptive training be included as one of the automated training features of the system.

I would like to emphasize, though, that the SFTS is not primarily an adaptive training system. Adaptive training is one of several features of the SFTS which are intended to reduce reliance on instructors during simulator training by automating some of their instructional functions. It has always been conceived of as a relatively small part of the system, but one that perhaps would grow.

Since 1965, the SFTS has gone through a series of concept formulation studies, and a contract has been awarded to Link to build a four-cockpit subsystem. Ralph Flexman's group here at the Institute of Aviation has been given a subcontract to prepare the SFTS's automated training material, including the material that deals with adaptive training. Dick Weekes and Julius Gandelman have been working on it, so I've asked Julius to give us an introduction to the adaptive training capabilities of the SFTS.


Mr. Gandelman

Mr. Gandelman: I'll start by outlining the structure of the SFTS training program. Each period of automated training runs about 105 minutes and is broken down into elements consisting of well defined phases of learning. For example, the ILS training period contains the following training elements: orientation, interception, localizer tracking, and glideslope tracking. The student starts with a simple part of the maneuver, and by the time the training period is over he is attempting the entire ILS approach.

Each element is broken down into four phases: briefing, demonstration, guided practice, and adaptive practice. In the briefing phase, the computer tells the student what he is expected to do, what the standards are, and so on. The student is then given an automated demonstration of how he is supposed to perform. During the demonstration, he watches all the instruments move and feels all the vehicle motions to get some idea of what the task is like. The next phase is guided practice. The student now actually performs the task, but he is "talked through" the required maneuver. His hands are on the controls, and he's being told what to do. If he exceeds the error tolerances, he hears an audio alert over the headphones. For example, he might hear the word *altitude,* indicating that he is out of tolerance on altitude control. The last phase is adaptive practice. In this phase, the student's performance is scored, and the concepts of adaptive training are implemented in a series of tasks. In the case of flying ILS approaches, the adaptive variable is simulated air turbulence—the amount of rough air—which effectively varies task difficulty.

So, in each training period we have various elements, and each element has briefing, demonstration, guided practice, and adaptive practice phases. When the student moves to a new element, all four phases are repeated. We have now structured two complete training periods in this manner. Basic aircraft control and instrument interpretation comprise the first period. The second period is ILS training.

Now, in these two training periods we use measures that we call performance tolerances. A performance tolerance is a statement of how well the student is expected to perform at the end of training. For example, it might be a statement that he should hold his altitude within plus or minus 100 feet. The error criterion for adaptive changes is based on the accumulated time-out-of-tolerance divided by the elapsed time. At present, we are using 10% time-out-of-tolerance as the error criterion.

DR. KELLEY: Does time-out-of-tolerance refer to one or to some combination of the performance tolerances?

MR. GANDELMAN: As it now stands, it refers to only one parameter. If we were to consider more than one parameter, we would use a weighting function. We would apply the appropriate weights for each parameter and compute a vectored sum of the time-out-of-tolerance for all parameters.

MISS KNOOP: How would you select the weights? Would it be on the basis of opinion?

MR. GANDELMAN: The Army needed a lot of flexibility; in giving them this flexibility, we did not attempt to specify weights, numbers, or parameter values. We tried to make the programming accessible so they could change the weights as their experience with the system grew.

". . . ADAPTIVE VARIABLES FOR THE SFTS. . . ."

We have selected three adaptive variables for the SFTS. The first one is air turbulence; the second is control damping. Turbulence will be used to make the task more difficult. Control damping will be used to make the task easier. Both variables will be used in the aircraft control training period, but control damping will not be used in the ILS training period, because by that time the student doesn't need that particular kind of assistance. However, in the ILS period, we have introduced the third adaptive variable, horizontal wind. It is not a crosswind; it is a headwind or a tailwind that would either increase or decrease the simulated ground speed and thus change the time limits of the task. For example, if the task were to intercept a localizer beam, a headwind would give the student more time to think and to respond. With a tailwind, he's really got to hustle. The same horizontal wind may be used as a crosswind. If the student is given a tracking problem, we can increase its difficulty by adding a crosswind so that his heading angles have to be recalculated and his work load is increased. In effect, we really have four adaptive variables for the SFTS: turbulence, control damping, along-track wind, and crosswind.

Mr. Norman: Control damping. Could you enlarge on that a little? You don't actually mean the control response itself, do you?

Mr. Gandelman: Actually it is a change in the moment of inertia of the airframe response.

Mr. Weekes: Any rotary-wing aircraft is inherently unstable, and by decreasing the unstableness of the simulated helicopter, we make it easier for the student to control.

Dr. Caro: In effect, you're making it a heavier helicopter, which makes it more stable.

Mr. Gandelman: The engineers back at Link would argue that this is not stability augmentation.

Mr. Weekes: But that is what it amounts to. The argument for introducing control damping is the same as the argument for having the instructor assist the student in an actual helicopter. The aircraft is so skittish that if the instructor told the student, "You've got it, boy," his hands would be so full trying to keep the thing right side up that he wouldn't have any time left over. So the instructor has to override the student on the controls and take care of many housekeeping details to allow the student the time he needs to concentrate on the point that's under scrutiny in the lesson.

MR. GANDELMAN: We used that particular concept throughout all the training periods, not just in adaptive training. In introducing multiple-axis control tasks, we knock several axes off in the beginning, so the student need only concentrate on one axis. As he improves, we introduce the other axes, as appropriate.

DR. CARO: Go ahead with your description.

". . . WHAT IS THE ADAPTIVE LOGIC. . . ?"

MR. GANDELMAN: Whether the performance tolerances are too loose or too tight can also be a subject for debate. If the tolerances are too tight, the changes in the adaptive variable can become frequent and quite obvious to the student. If this is the case, I suspect that the student is going to be coping with the adaptive variable rather than with the primary task. I favor inconspicuous adaptation. In other words, the change in the adaptive variable should be so gradual that it really is not obvious to the trainee.

DR. CARO: Would you describe how this change takes place? What is the adaptive logic?

MR. GANDELMAN: The SFTS provides the first real opportunity to have both the application and the research tool combined in one device. As the students start using it, they're going to find that certain values do not work, and they will have to find out what values do work. In effect, they will be meeting an application and, at the same time, doing research.

To see what we do now, let's take the first training period and the element called bank control. On the first bank-control task, the student is told to turn to heading 120° from 30°. The task interval will be a function of the rate of the turn, which we know should be about 3° per second; so a 90° turn would require 30 seconds. That is a very short task interval, and it leads to a very basic question: How much of his ability can you assess in 30 seconds? Well, not very much, in my opinion. Therefore, after performing the first task, the student is given five or six succeeding bank-control tasks. The logic is that, whatever the entry level on the first task, there will be adaptive changes in the difficulty level; and the difficulty level at which he completes the task is the entry level for the succeeding task.

The difficulty scale runs from Levels 1 through 9 in discrete steps. Level 3 is the anchor condition in which there is no rough air and the controls operate normally; the first task is introduced at this nominal entry condition. Levels 4 through 9 are defined by successively increasing amplitudes of turbulence. If the student cannot perform within established error limits at Level 3, control damping is introduced at Level 2 and increased at Level 1 to give the student further assistance.

The entire task interval has been divided into 10-second segments for the purpose of measuring performance. Why 10 seconds? Again, we don't know the optimum length and have arbitrarily selected 10 seconds. Every 10 seconds we measure the accumulated time-out-of-tolerence and the error rate to see whether the student has exceeded the error criterion for adaptive changes. We presently plan to work at a criterion rate of 10% time-out-of-tolerance.

At the end of a task interval of 30 or 40 seconds, you have a chance to sample and evaluate performance only three or four times. The student may have risen from an entry difficulty level of 3 to a difficulty level of only 5, which does not satisfy the exit criterion. The exit criterion is the highest difficulty level possible, in this case, 9. If the student does not meet that criterion and is at Level 5, then we give him a brand new, but similar, task. He enters that task at Level 5 and proceeds for as long as is allowed for that task. We continue giving him additional tasks until he gets to a difficulty level of 9. To exit from adaptive practice with the SFTS, the student must be proficient at the highest difficulty level possible. But, how difficult that is depends upon what have been established as the adaptive functions. How realistic is the turbulence at Level 9, and what training has been achieved at that level? These are some very fundamental questions that have not yet been answered.

When the student is able to achieve one element, he then goes through the same sequence—briefing, demonstration, guided practice, and adaptive practice—on a new element which is yet more difficult. That summarizes the adaptive logic and some of the basic concepts that we've employed in the SFTS.

Dr. Caro: Thank you. Bill Stewart has asked to exercise his right to interrupt at any time and ask a question.

". . . 'FLYING' THE ADAPTIVE VARIABLE. . . ."

Mr. Stewart: It may be useful to have a quidnunc who is willing to step forward where other people are reluctant to go. Apparently there is some concern about whether the student will simply learn to adapt to the turbulence variable, whether he will be "flying" the adaptive variable or "flying" the task. It seems to me that if he's flying the variable, he is going to learn the task. That's really what adaptive training is all about.

MR. GANDELMAN: I doubt that. If he is being trained to fly in very heavy turbulence, how realistic is it for the operational task of straight and level flight in the absence of turbulence? The control dynamics, his response, and everything else are quite different.

MR. MANSFIELD: The assumption is that the student has demonstrated that he can fly satisfactorily in still air before you make the task more difficult by introducing turbulence.

MR. GANDELMAN: That is the assumption, but is the student learning how to fly turbulence or is he merely learning how to fly straight and level in the presence of turbulence?

MR. STEWART: Well, the latter is the case, isn't it? The specific task being taught is to fly straight and level with varying amounts of turbulence. As the turbulence increases, the student has more difficulty flying, but presumably he eventually flies turbulence at the level you define as the exit criterion. It seems to me that if he achieves this criterion, he has learned to fly straight and level. He might even be getting to the atrociously overlearned stage, but he has developed a very fine skill in dealing with this particular beast under unstable conditions. And, that is really the whole purpose of the training.

MR. GANDELMAN: Operationally, flying in very heavy turbulence is a different situation from flying with little or no turbulence. For example, a pilot will tolerate altitude errors of a lot more than plus or minus 100 feet. He is unable to anticipate turbulence or gusts, and his behavior is really not the same as it would be with no turbulence. So, I question what training is achieved by having a student learn to fly straight and level in very heavy turbulence.

DR. CARO: Well, Julius, you say that error tolerance is greater in heavy turbulence. Operationally this may be the case, but is it the case in training?

MR. GANDELMAN: No. The issue of changing tolerance limits for adaptive changes has long been discussed. I believe the issue is basic to the differences among the views of Birmingham, Hudson, and Kelley.

DR. CARO: The same tolerance band is maintained while the task is made more difficult on the assumption that the student should be able to maintain that tolerance band operationally with gust levels up to that which you simulated. Surely you will not include unrealistic turbulence levels. Level 9 won't be any more than can be handled operationally by a competent pilot within the specified tolerance, will it?

Mr. Gandelman: That's right, but the question should be raised. I'm not really sure of the answer. I merely suspect that the student is learning to fly turbulence rather than learning to perform the primary task.

Mr. Wood: I think past research bears you out. Hudson demonstrated that the level of difficulty under which you practice has a lot to do with the kind of criterion task you will perform best. In other words, if you practice under conditions of high error, then you will do a better job on a criterion task with high error. If you practice under conditions of low error, you will do better on criterion tasks with low error.

Mr. Gandelman: Also, we're using a constant tolerance band, and I suspect that in the flight situation we're going to have the pilot flying at the tolerance altitude rather than the assigned altitude.

Mr. Stewart: What you really need is a variable performance criterion as well as an adaptive variable so that you train a man to fly in very smooth weather perfectly or in rough weather less perfectly. But, requiring the student to hold altitude within plus or minus 100 feet at Level 9 turbulence may be unreasonable.

". . . THE STUDENT GOES WHERE YOU WANT HIM TO GO. . . ."

Dr. Kelley: There are several points in this discussion that we ought to partial out. One concerns whether or not the adaptive variable is the appropriate one for this task; which is to say, should turbulence be used as one of the adaptive variables? If turbulence is not the significant adaptive variable, there is no point in teaching the student to fly in different levels of turbulence. So we should separate this question from the question of what technique of adaptation, or what adaptive logic, is used to progress through the different levels of difficulty. Turbulence might not be an appropriate adaptive variable if unrealistic levels of turbulence were used or if turbulence so interacted with the performance tolerances that a change in one required a change in the other.

I believe, as a matter of principle in designing an adaptive task, that it makes sense to find out how the task is performed in the real-world situation, because where you begin in training may not be so important as long as you end at the right place. But, it is very important that the student goes where you want him to go and that he does not go someplace else as a consequence of your adaptive variable. So, until we know a great deal more than we know now, I think we have to take a real-world situation as an implicit criterion and establish a simplification, if simplification is necessary, when we're going to use adaptive training. The student will adapt along a dimension set by some adaptive variable. But, you have to choose and manipulate that variable on the basis of your understanding of the task. You have to choose an adaptive variable that brings you along some dimension of simplicity-difficulty towards the real-world task. Now, this seems to me to be pretty much what they have done in the SFTS.

We could talk about whether or not these are the right adaptive variables (they don't look too bad to me), but I'd be much more concerned about the performance measurement and what goes into it. I would have real qualms about that. On the basis of this presentation, I can't judge whether the measures are likely to be unreliable or insensitive, but they really have to be investigated.

MISS KNOOP: Could I ask what performance variables are being measured?

DR. CARO: They will be measuring parameters of aircraft control, such as altitude control.

MISS KNOOP: Rotor RPM?

DR. CARO: Rotor RPM, heading, altitude, turn speed, bank, and things of that sort.

MISS KNOOP: Well, the reason I brought it up was that Dr. Billings at Ohio State found that rotor RPM in helicopter flying is one of the more significant measures.

MR. WEEKES: That was with an old-fashioned helicopter.

DR. CARO: The design-basis aircraft for the SFTS uses a turbine engine with a governor control that reduces that problem.

MISS KNOOP: That's too bad; it would be a good measure. The other point I want to make is that Chuck Kelley is perfectly right. We should not pretend to know, at this point, whether or not turbulence is the correct adaptive variable. I think we will all agree that we do not know.

". . . THE CRUX OF THE CRITERIA. . . ."

Mr. Green: The contention that Patty brought up this morning, that performance monitoring seems to be the crux of the criteria for adaptive training, holds true for this particular trainer. And I think that's going to be true for every problem we try to solve with adaptive training. I don't think a set of generalized criteria can be made. When you try to implement adaptive training into a simulator, a specific set of design criteria for that particular adaptive feature must be made. In other words, although the generalized scheme of adaptive training is to be implemented in the SFTS, a rotary-wing device, a completely different set of criteria might be required for a fixed-wing device.

I had another question pertaining to the exit criteria. Can the student exit from any point in the element? Although Julius Gandelman pointed out that he had to reach Level 9, I'm curious about why he could not exit, say, at Level 4 before going into the next element.

Dr. Caro: The way it's programmed now, the student must perform one complete task at Level 9 before he exits.

Mr. Gandelman: This is what we're hoping for. We don't know whether it's possible.

Mr. Green: I asked the question because training emphasis should be placed on relevant tasks in respect to each other whether or not they follow the same sequence with the same weighting.

Mr. Gandelman: You see, when the student exits from an element, he goes into an entirely new training situation with a new set of related tasks.

Mr. Green: I was thinking of the task in which he will be trying to intercept the glide path first and then make a certain type of turn, depending on altitude, speed, and how close he is to the actual touchdown point when he intercepts the glide path. These factors are all interrelated.

Mr. Gandelman: We assume the student must be proficient in each task prior to going to a subsequent task.

Mr. Green: But, your definition of proficient means performing at Level 9.

Mr. Gandelman: That is true at this stage of the SFTS development.

". . . GETTING FEEDBACK. . . ."

DR. MCGRATH: You said earlier that the changes in the adaptive variable or changes in the level of task difficulty should not be obvious to the students, and yet those changes might be an important source of knowledge of results to the student.

MR. GANDELMAN: A digital display will tell him what difficulty level he is operating at. The changes should be continuous or smooth enough so that his only indication is a feeling that, "I'm now working harder, but I don't know exactly when I started to work harder."

DR. MCGRATH: Why is it important that the student not be able to discriminate these levels without the digital display?

MR. GANDELMAN: Well, this may be a personal rationalization. Part of the flying task is to anticipate and/or be sensitive to cues—motion cues, instrument cues, and so on. Developing sensitivity to a wrong cue would be detrimental to his flight skill. In other words, if turbulence changes suddenly and obviously, the student knows he now must behave differently. But, the pilot's overall task is to be sensitive to gradual changes as they actually occur.

MR. STEWART: But, isn't that a violation of the feedback principle on which Chuck Kelley criticized Milt Wood's study?

DR. CARO: No. The student is getting feedback via the score display.

MR. STEWART: Then that satisfies Jim McGrath's point.

DR. CARO: I believe so. The feedback score has two digits. The first represents the task number; the second represents the level of difficulty of that task. The feedback score is continuously displayed and updated, although the actual change in task difficulty may occur a few seconds after the change in score.

MR. STEWART: So the student is given a lead time warning of what's about to happen to him?

DR. CARO: Not necessarily. Since turbulence magnitude varies randomly within a given difficulty level and since there is likely to be some overlap between levels, the difficulty level per se could change before a perturbation typical of the new difficulty level occurs. An occasional effect will be a lead-time warning.

MR. WOOD: A rule which could be used is this: If an event or series of events is likely to occur in the criterion task, then the same can be used in training. Let's say, for example, that one is flying along in an aircraft and all of a sudden turbulence is encountered which in turn alerts the pilot to a new set of tracking strategies. Since these events are characteristic of the criterion task, why not employ them in training?

Mr. Gandelman: Let me try to word it a little differently. The pilot should be sensitive to changes or to cues that occur in the criterion task. If you make the change obvious in the training task, you're not developing in him any sensitivity or improving this skill.

". . . PILOT MODEL THEORY. . . ."

Mr. Sinacori

Mr. Sinacori: In the selection of the adaptive variables, was any consideration given to the work that has been done on handling qualities and on pilot modeling? For example, using control damping as an adaptive variable in the SFTS has the function of varying the effective time constant described by Guy Matheny (Matheny, 1969). The pilot model theory states that responses in a helicopter require the perception of angular rate before the pilot can adequately achieve stabilization. The theory also states that he will develop a lead equal to the basic lag of the system. Once achieved, he will then vary the gain to minimize error, and this depends on the forcing function. In other words, if the range of the forcing function exceeds the basic bandwidth of the vehicle, the pilot won't be able to do anything about it. The airframe lags would no longer be the dominant factors, and the error would simply increase while the pilot did nothing about it. However, if the forcing function is within the airframe response limits, he will attempt to adjust the gain to satisfy the error criteria.

So, what I'm saying is that the choice of control damping as an adaptive variable seems to be a good one, because the untrained pilot does not generate sufficient amounts of lead. By giving him increasing amounts of turbulence, you will force him to gradually increase his lead until it reaches that equal to the basic airframe lag. At this point, he can do no more. But, I wonder if the pilot lead does, in fact, vary from zero to the airframe basic lag in the SFTS and if it can be forced up there a lot faster with the use of adaptive logic. Would anybody care to comment on just what has been the utilization of pilot modeling in the choice of adaptive variables or in the adaptive logic?

DR. CARO: Well, in a general sense, the pilot-instructor model was used in selecting the adaptive variables. One of the Army's requirements was that training had to be as realistic as possible. Nothing would be put in that was purely artificial or did not contribute to the psychological illusion that the man was flying a real helicopter. The choice of control damping came about as we were looking at what happens in a real training helicopter when the student has difficulty controlling the aircraft. A good instructor will let the student have control, but he will keep his own hand on the control. As the student becomes more proficient, the instructor makes less and less input, until finally his hand is off the control altogether. If he does this smoothly enough, the student doesn't even know when the instructor's hand is on the control or when it comes off. This was the general model.

MR. SINACORI: The instructor's function was to ensure good response of the vehicle so that the student would see how it should respond under the conditions for which he was being trained.

DR. CARO: More importantly, the instructor's function was to reduce the real-world task to a level that the student can handle initially and then gradually remove this control assistance so that the student has the full task.

". . . EXPAND THE TOLERANCES. . . ."

DR. MATHENY: I share Julius Gandelman's concern about using a fixed-tolerance criterion to which the pilot must control the helicopter throughout an increasing amplitude of turbulence. This amounts to changing his control problem and his technique. The pilot must learn that, up to a certain level of turbulence, he can control the aircraft within a certain tolerance; beyond this level of turbulence, he must change his technique and expand the tolerances. In extreme turbulence during actual flight, you either change your control technique tremendously or tear the beast apart. On a less dramatic scale, the same thing happens in simulator training. Let's not be rigid about this. There is the problem of selecting the adaptive variable and difficulty levels, but we should also examine the need to alter the error tolerances in line with what we know about the requirements of actual flight.

Mr. Sinacori: I think you'll find that increasing turbulence up to the first break frequency of the aircraft that is encountered will be okay. But beyond that point, a strategy is required for the adoption of lead that will extend the bandwidth. So a relationship is implied between the turbulence level and the criteria used to establish the error tolerances.

Dr. Matheny: There should be some optimum tolerance.

Mr. Sinacori: Up to the basic, or second, lag of the airframe. Beyond that, there is no hope of controlling the aircraft, and error tolerances of any kind would be irrelevant.

Mr. Norman: You asked to what extent pilot modeling was being used in selecting adaptive variables or in establishing adaptive logic. We are planning an experiment in which we will use several adaptive variables. One is analogous to direct-lift control, another one is the effective time constant, and the third is the forcing function. In setting up the forcing function, we appear to have been able to demonstrate crossover frequency regression.

Dr. McGrath: What is crossover frequency regression?

Mr. Norman: It is a condition in which the pilot must insert lag, because he cannot follow all of the deviations, and starts averaging them.

Dr. Matheny: The pilot changes control techniques as a function of the changing task conditions.

Mr. Norman: We could see this happening as a function of his control inputs.

Dr. Matheny: Does the SFTS have a motion platform?

Dr. Caro: The SFTS has a five-axis motion platform: pitch, roll, yaw, heave, and sway.

Dr. Matheny: You see, the response characteristics of the motion platform are going to be very critical in their effect upon the crossover frequency of the pilot. Adding motion to the trainer changes the effective time constant of the pilot-machine system and, in most instances, will act to raise the pilot crossover frequency. This is just another complication in trainer design for which we don't have the answers.

Dr. Montemerlo: Well, maybe I just don't understand something. Julius Gandelman said a trainee should learn to fly his craft within certain tolerances at a low level of turbulence, but, at a higher level of turbulence, he should learn that he cannot use that same level of tolerance. Is that what he is supposed to be learning?

DR. MATHENY: That is what he could be required to learn.

DR. MONTEMERLO: I'm not concerned with what he *could* learn. Is that what he *should* learn?

DR. MATHENY: It depends on the level of turbulence he must learn to deal with. In very heavy turbulence, an experienced pilot does not try to hold altitude within ± 100 feet any more; he just rides it out and keeps the aircraft level.

DR. MONTEMERLO: Then improper response to turbulence is actually being trained here. You may have a case of negative transfer of training.

MR. POVENMIRE: It depends on how much turbulence there is.

MR. FLEXMAN: A few gusts and a little rain does not mean a turbulent thunderstorm.

DR. MATHENY: What does it mean? There is that question of difficulty levels again.

Mr. Povenmire

". . . THE STUDENT COULD PRACTICE INDEFINITELY. . . ."

MR. GANDELMAN: A very important related point is that the adaptive variables may not be changing at the individual's best rate. They are changing at some artificial rate. We are arbitrarily using 10-second adaptive-change intervals, and we don't know whether that interval is optimum for a particular individual.

DR. CARO: I think you are misinterpreting the adaptive logic. The performance-measurement interval is now set at 10 seconds, but the rate at which adaptive changes occur is a function of the individual student's performance. So task difficulty will be adapting to the individual's performance, but it will not necessarily change every 10 seconds. In the SFTS the student could practice indefinitely without any change of difficulty if he remained within the programmed tolerance. This is quite different from the way you approached it in your study, Milt.

MR. WOOD: One part of this problem gets back again to the measurement of subject performance and to the question of how much information you want to throw away. Time-out-of-tolerance is used in the SFTS, right? It does throw away information, and it would seem that a more sensitive measure would be desirable. I would want to look at some measure of central tendency of student performance, in combination with some measure of variance. It might be desirable to go a little further and try to come up with a measure of reliability. In other words, why not pull all the information that you can out of performance and use it to adjust time and difficulty level for adaptive practice?

Dr. Caro: I didn't understand your earlier discussion of central tendency. Now, in Chuck Kelley's method, there can be no information about central tendency, because error rate is held constant. There can be some variation in error rate, but information about central tendency has a limited value.

Mr. Wood: Its interest to me was primarily in terms of relative shifts in error between the two treatments and relative variations in error for any given trial. Because statements of error directly reflect the visual characteristics of the task (the location of pitch and roll indicators), they can be of some value in discussing task differences.

". . . A BROADER DEFINITION OF ADAPTIVE TRAINING. . . ."

Mr. Stewart: That raises an interesting problem that troubled me this morning. It derives from the definition of adaptive training that Chuck Kelley put on the board and his inclusion of that word *automatically*. I don't know why he put that word in parentheses, but I assume he is not dead sure that it really belongs in the definition. Automatically at what rate? I think a broader definition of adaptive training, one without the *automatic* in it, would do away with this problem. Is it enough to say that training is adaptive if you increase the difficulty of the task as the person performs it better, and that whether you change it rapidly or slowly is a decision you make in connection with the equipment you are using? I think this point is really central to understanding what the devil it is that adaptive training does. To enhance the rate of learning, maybe it needs to be automatic, instantly responsive, or as quickly responsive as you can make it. If you have a ballpark idea of the reasonable rate at which students learn the tasks, then you might simply change difficulty levels in pre-established incremental steps. I would characterize the method used by Milt Wood as incremental adaptive training and the other as continuous adaptive training.

". . . LEARNING TAKES PLACE SLOWLY. . . ."

DR. KELLEY: Whether a method is incremental or continuous depends on how adaptive changes are made. Let's not confuse adaptive changes with adaptive rates, which are defined by the periods of performance measurement. Adaptive changes can be made incrementally 100 times a second, or adaptive changes can be made continuously, with a time averager, so slowly that there is an imperceptible change in difficulty during a given training trial. So, *continuous* does not mean adapting quickly, and *incremental* does not mean adapting slowly.

Now, I've found many people make the system adapt much too quickly. This can cause wildly oscillatory variations of difficulty level within trials. Learning takes place slowly in the development of complex perceptual-motor skills, and there is no reason to make the task adapt quickly. But, a slower response does not mean that adaptive changes must be made incrementally; it means only that performance must be averaged over a long enough period of time to avoid an undesirable oscillation in the difficulty level of the task.

MR. SINACORI: Chuck, are you implying that there may be a basic time variable in any problem that is related to processing time?

DR. KELLEY: What do you mean by processing time? I'm not sure I understand.

MR. SINACORI: The amount of time in the averaging circuit over which error is measured.

DR. KELLEY: That is one way to do it, in an analog implementation. In a digital implementation, you would consider how many samples are averaged over what period of task performance. What you should never do is change the difficulty level as a function of the instantaneous performance of the student. When you do that, you permit all of the random error fluctuations to change the characteristics of the task. Because performance is usually measured unreliably—random elements are perturbing it—it is desirable to get a sample of performance averaged across a period of time. I am suggesting that people are in error much more often by making that period too short than they are by making that time too long.

". . . THE R-S LINK. . . ."

DR. MCGRATH: May I suggest a theoretical idea that might explain one of the dangers of making the adaptive interval too short? Most learning situations are concerned with establishing connections between stimuli and responses, S-R, so that an individual learns to take certain actions in the face of certain events. Now, the peculiar thing about adaptive training is that the response itself produces a new stimulus, S-R-S, which calls for a response, which in turn produces a new stimulus, which calls for a response, and so on, S-R-S-R-S-R-S-R. . . . Consequently, it is entirely possible to establish a bond between the response and the stimulus if the timing of the two is very close and if the R-S relationship is obvious and consistent. Perhaps this is why Julius Gandelman doesn't want the student to perceive the adaptive change. He may well be establishing an R-S connection in addition to, or instead of, the S-R connection that is the foundation of learning.

Mr. Flexman: Incidentally, this does happen when the student tries to beat the system. There are quite a number of examples in the literature showing how students respond to stimuli extraneous to the actual training task and learn that they can control the stimuli.

Mr. Stewart: Of course the answer to Jim McGrath's point is that you need a lag in the display. The student enters a new condition, and x seconds after he has been alerted to a change by the inherent cues in the trainer, the display verifies the change.

Mr. Wood: But, this R-S link is not all bad. It is all right if it is part of the criterion task being trained.

Dr. McGrath: Surely it is not a part of the criterion task in flight training. In fact, the paradox of adaptive training is that the R-S link in training is the reverse of any such link in the criterion task. In actual flight, a good response produces an improved situation, and a poor response leads to a worsening situation. On the other hand, adaptive training is based on an R-S logic in which good performance leads to a worsened situation, and poor performance leads to an improved situation.

Dr. Lauber: I'd like some clarification on the adaptive method in the SFTS. Is my understanding correct that damping on all axes will be maximum to start with and that undamping will take place only within one of those axes, say pitch? Does the student learn to control pitch while damping is held maximum on the other axes, or does adaptive damping take place on all axes simultaneously? I'm concerned that if adaptation doesn't take place simultaneously—that if you allow variation in one axis while holding others constant—there is fantastic opportunity for the trainee to pick up unrealistic modes of response because of all the interactions among the control axes.

MISS KNOOP: Are you increasing difficulty by adding turbulence, decreasing difficulty by removing turbulence, and then further decreasing difficulty by adding control damping?

MR. GANDELMAN: That is correct.

DR. KELLEY: To comment on John Lauber's point, sometimes it's necessary to have unrealistic conditions in training very complex skills. When you learn to play the piano, you start with one hand, then the other, and then learn to use both together. Although I'm not a helicopter expert, I understand that flying a helicopter is so difficult that a novice is not normally required to attempt to operate the controls in all degrees of freedom; that would be simply too much. So, learning to fly is a bit like learning to play a piano: the student learns to use first one hand, then the other, and then puts them together. That, of course, is very unrealistic in terms of what he finally has to do. Whether it is best to do it that way or to reduce all tasks to a difficulty level that would permit his learning them all together is a methodological point. Whatever method is used, the student is going to start out on an unrealistic task because the realistic task is too much for him to cope with at the start.

DR. LAUBER: I'm not sure that your example of the piano player was exactly analogous because, in a sense, learning to play with one hand and then the other is a simple kind of addition. In the case of flying a helicopter, there is an extremely complex interaction; introduction of the second task modifies the first task.

MR. WOOD: Fitts' theory concerning how we learn a perceptual-motor task suggests that we first become aware of changes in position, then we begin to discriminate rate changes, and finally, as skill increases with practice, we are able to discriminate second-order and even third-order characteristics of the tasks. If that is a reasonable theoretical basis, we can perhaps sequentially modify a task adaptively through these dimensions without causing too much confusion. Because, early in training, the student can perceive only these lower-order characteristics, only as he increases in ability can he discriminate the higher-order characteristics of the task.

". . . CALL IN A REAL LIVE INSTRUCTOR. . . ."

DR. CARO: There is an aspect of the SFTS adaptive logic that hasn't been mentioned. The desired error rate is set at 10%. If, at the end of a 10-second interval, the student is at a 20% or greater error rate, the task is made easier; if he is at a 5% or less error rate, the task is made more difficult. His performance is maintained within the range +10% or −5% of his desired error rate by making adaptive changes at the end of each 10-second interval. However, at the end of the task, a different set of standards is applied. If the student has gotten all the way through the task and no adaptive changes have occurred, that is, he has remained within these fairly broad tolerances, a tolerance of ± 2% is applied. So the SFTS will have two criteria for adapting: the short-term criterion applied at the end of a particular 10-second time interval and a more stringent criterion applied at the end of a much longer segment of performance.

Dr. Matheny: What happens if he doesn't meet that second criterion?

Dr. Caro: He continues at the same difficulty level.

Dr. Matheny: If you have to reduce difficulty, do you go back to the 10% criterion?

Dr. Caro: Your target error is always 10%. When it becomes necessary to reduce difficulty, in effect, you go back one step. You decrease the difficulty level, say from 8 to 7. You enter the next series at Level 7, with a criterion of 10% plus or minus the specified tolerances.

Mr. Gandelman: Paul, perhaps you should mention what happens if the student goes all the way down to Level 1 of difficulty.

Dr. Caro: Yes. If he does, in fact, go through a complete task at Level 1, in other words, if he has regressed to the easiest task possible and has made no progress, the system automatically freezes, and the "red flag" comes up calling for an instructor. That is also something that won't happen in the aircraft; in the aircraft he would just crash and burn.

Mr. Green: You also have the possibility of freezing the problem and repeating the same routine?

Dr. Caro: We do not have that option automated. It is perfectly reasonable to have such a feature, but the decision was made to call in a real live instructor whenever a student goes all the way through a minimum difficulty problem with no progress. It may be that the student has fallen asleep and that someone has to take a look.

MR. GREEN: Does the problem have phases to it? In other words, are there integral parts which do not necessarily lock the student into one particular level of proficiency? For instance, if he is not doing well at landing and he is down to Level 1, can he go to another phase, such as tactical maneuvers, where a different form of flying is required?

DR. CARO: The tasks have been ordered in terms of judged difficulty. The student completes one task at Level 9 and enters another task at Level 3, but the second task at Level 3 is more difficult than the first task at Level 3. So, when the instructor sees that the student is making no progress, the instructor can opt to set him back to the preceding problem where he was able to proceed to Level 9. At this point, we are not even trying to automate that decision; we would like to, but we need additional information before attempting it.

MR. GREEN: I suspect such a decision may be warranted within short order after the device is delivered and operated.

DR. CARO: Yes. Well, this is a very flexible system. We hope that we have made some reasonable guesses and that the models we have selected are correct. But, after we get some data, we can revise the numbers that have been used in the initial program.

". . . SOME OPINIONS IN WRITING. . . ."

Let's end this discussion and get some opinions in writing. Will you all please answer these three questions: What are the proper guidelines for selecting adaptive variables? What problems do you see in the SFTS implementation of adaptive training? What high-priority research should be done?

The Proper Guidelines for Selecting Adaptive Variables

MISS KNOOP: I think the determination of guidelines for the selection of adaptive variables is a research study in itself. Furthermore, the experimental determination of guidelines must be made for different training objectives and different criterion tasks; the "correct" adaptive variables will be specific to unique training objectives and tasks. Ultimately the adaptive variables should themselves be selected adaptively, because the appropriate adaptive variables are likely to differ among trainees and for a single trainee at various points in the program.

Dr. McGrath: In the opening session, Chuck Kelley told us not to be reluctant to state the obvious, so I will: the paramount rule should be to select an adaptive variable that is clearly relevant to the operational task being trained. No other guideline matters if this one is violated. The other guidelines mainly insure that the selected variable is practicable. That is, the variations should be easily definable or measurable; the variations should have a significant, reliable, and known effect on the performance requirements; and that effect should be direct and uncomplicated.

Mr. Wood: I suggest the following three guidelines: (1) Identify possible distortions introduced into the task by adaptive techniques, and assure that any such distortions provide more positive transfer than negative transfer. Distortions can occur in terms of task dynamics, basic stimulus characteristics, time bases, and so on. Task distortion and, as a consequence, negative transfer could be minimized by relating elements in the training task closely to those in the criterion task. (2) The use of feedback variables in the training paradigm should be defined in terms of their impact on criterion performance. General guidelines are few; however, past research suggests several techniques for handling augmented feedback situactions. (3) Measures of performance should be carefully chosen and refined. Measures may need to be sensitive to both central tendency and variability. Also, the length of time over which measurement is obtained should provide for relatively slow adjustments of the adaptive variable.

Mr. Mansfield: The realism that has been apparent in simulation during the past few years should continue to be a strong dictator of the adaptive variables to be used. Although I have sensed, during this conference, that some of the participants don't consider realism to be that important in accomplishing training objectives, I continue to feel that it is an important guideline.

Mr. Stewart: The most important guidelines are: the relevance of the forcing function to the skill trained, the ease of varying task difficulty level, and the nature of the difficulty dimension itself—that is, whether the difficulty is increased by a physical challenge (turbulence), attention-dividing (bells, lights), or a design change (display size, control dynamics).

Mr. Sinacori: The following ingredients should be included in any set of guidelines for selecting adaptive variables: (1) sensitivity to the generation of errors regardless of what criteria are employed,

(2) possibility and practicality of adjustment in terms of stability and resemblance to the real world, and (3) predictability on the basis of some plausible theory.

DR. MATHENY: Guidelines required for selecting adaptive variables for training applications may differ from those required to select adaptive variables for research purposes. For training applications, the adaptive variables must be measurable and must be related to progress toward the training objective, that is, the task being trained. The appropriate performance measures, and so on, must be well defined before attempting to select the adaptive variable. For research purposes, the adaptive variable must, again, be measurable; however, it must also be adjustable over a relatively wide range of skill levels, with particularly good coverage at the lower end of the continuum.

DR. LAUBER: Adaptive variables should be selected to be consistent with the real world that the trainee will encounter when training is complete. The use of turbulence is probably a fine example: it increases the load on the trainee in a realistic way, and its implementation is relatively straightforward.

DR. KELLEY: Each adaptive variable should be selected in accordance with these guidelines: (1) its variation should be along dimensions relevant to the skill to be taught, (2) its variation should be closely related to measured performance so that a change in it reliably changes measured performance, and (3) it should be convenient to implement.

MR. NORMAN: I find it somewhat difficult to articulate guidelines; however, it seems clear that the adaptive variable must do no obvious violence to the training situation. I think probably the best approach at this point in time is to be empirical: try the variable, and if it works, use it.

MR. HALL: The variable should have some relevance to the operational task so that it won't be perceived as being artificial. Also, it should be capable of varying in magnitude and should be limited to changing the task along a difficulty continuum and no other.

DR. MONTEMERLO: I think the approach is quite straightforward. It should consist of the following steps: (1) behaviorally define the criterion performance, (2) complete a task analysis of this criterion behavior, and (3) identify the appropriate adaptive variable, if one is required, from the task analysis. I would never force the use of an adaptive variable where it could not be easily identified in the task analysis. For example, the selection of turbulence as an adaptive variable would be appropriate only if the task analysis clearly showed flying in turbulence to be an important criterion behavior.

Mr. Green: Of course, the ultimate guideline is to select adaptive variables that are most helpful in meeting specific training objectives. When a simulator is involved (and the early applications are likely to be in simulators), adaptive variables and their sequencing timetables will probably have to be assigned on an a priori basis. However, as more experience is gained and as more is learned about the relationship between trainee performance, the adaptive variables used, and the adaptive logic, more precise selection guidelines may be developed. For example, standard tables relating student performance to variation in adaptive variables may be prepared for a number of applications. If this approach turns out to be reasonable, tables probably would be developed initially for use with closed-loop systems.

Problems in the SFTS Implementation of Adaptive Training

Miss Knoop: Clearly, the major problem is in the performance measures. I question both the validity and the reliability of the measures proposed and strongly recommend that research using the SFTS be conducted on the performance measurement problem before very much reliance is placed on the adaptive training feature of the system. For example, percent of time-out-of-tolerance may be a useful measure for some of the training, but its usefulness is a function of the validity of the tolerances employed.

Dr. McGrath: The 10-second interval of performance measurement for determining adaptive changes will probably be too short. The available empirical evidence on adaptive training seems to favor the slowly adapting system, and so does theory. Moreover, on the grounds of psychophysics alone, I can think of few aspects of human performance that can be measured reliably in 10 seconds, and, if the performance measurement is unreliable, then of course the adaptive logic is reduced to a largely random process.

Mr. Wood: The time-out-of-tolerance measures of performance may have to be better defined. It is possible that much information of potential value may be lost during the "in-tolerance" performance. In addition, the very complexity of the SFTS model may obscure some of the important factors involved in the adaptive training process.

MR. MANSFIELD: The primary problems will be, first, to sell the SFTS concept to potential users of the system and, second, to motivate teachers and students who have been trained in the more traditional teacher-centered manner to use the system in a proper and beneficial way.

MR. STEWART: I suggest an investigation of the value of varying the sensitivity scales of the instrument displays to enable the student to fly within closer tolerances.

MR. SINACORI: I see five possible problems in the SFTS implementation. The program may be insufficiently flexible; the choice of adaptive variables may be inappropriate; the relationship between performance criteria and task difficulty may be improper; the model of the helicopter may provide an improper or irrelevant computational interface with the hardware or hardware dynamics; and the instrument dynamics may be improperly adjusted.

DR. LAUBER: The planned approach, which calls for the trainee to practice control on one axis, then on two axes, and so on, might possibly lead to difficulties in transferring from one task to another in the training sequence. The freezing of any axis or set of axes may cause the trainee to adopt strategies that are incompatible with strategies required when the axis or axes are unfrozen. I think this situation could result in significant negative, or at least neutral, transfer from one task to the next.

DR. KELLEY: The choice of adaptive variables seems to have been given a good deal of thought; I think a good start has been made. Of course, the adaptive logic requires empirical study with a great deal of attention given to the standards employed. The overriding problem in the SFTS implementation, however, is that of performance measurement. I doubt that the measurement approach presently planned is reliable and that it will result in effective adaptive task control.

MR. NORMAN: I see a couple of very practical problems. I think the appearance of the training situation is going to have to be structured in such a way as to be creditable in the eyes of the customer's acceptance pilots. Also, there's the possibility that the implementation may "bomb out" because the values assigned to the various parameters may be too far off the mark and not subject to a "quick fix."

DR. MONTEMERLO: The primary problem will be one of spelling out in sufficient detail the criterion tasks, that is, defining precisely the behavior required by a good helicopter pilot. A problem related to this, of course, is that of developing sufficiently valid and reliable measures of these behaviors in the training situation. Patty Knoop's discussion highlighted a number of the difficulties in the performance area.

MR. GREEN: I believe that, early in the SFTS implementation, a great deal of refinement is going to have to take place in the exit criteria and in the way a student is tracked through his training. I anticipate that the 1-to-9 scheme for scaling difficulty will have to be revised and that a system for accumulating milestone "misses" will have to be incorporated. Also, student motivation will probably not be handled very well by the current scheme.

High-Priority Research That Should Be Done

MISS KNOOP: I can think of four research areas to which I would attach a high priority. First, the performance measures and performance criteria required to implement the adaptive training feature must be developed and validated. Second, a transfer-of-training study is needed to compare nonautomated, instructor-managed training with the SFTS adaptive training concepts. Third, alternative adaptive variables, variables in addition to turbulence and control damping, should be investigated. A particularly interesting possibility would be the use of task structuring by means of various degrees of task loading, perhaps via malfunction insertion. Finally, I think we need research to determine the optimum relationship between the instructor and students in the SFTS. How can the instructor best be kept informed of the progress of the students? Is there an optimum instructor-student ratio?

DR. McGRATH: Without more detailed knowledge of the mission and capability of the SFTS, I would not want to guess at the priorities of its potential uses. However, as a matter of strategy, I think the first objective should be to accomplish that research which will most enhance the validity of the SFTS for its assigned mission, the training of helicopter pilots. The directors of the SFTS program should resist any immediate temptation to use the SFTS as a general research tool for exploring theoretical issues. The technology of adaptive training would be better served by a mission-oriented research program. When it has been demonstrated that the SFTS works—that it is an effective and practical training system—its versatility can be turned to the purposes of basic research. In other words, SFTS priority should go to the need to solve its own problems. After all, if the SFTS fails, adaptive training as an instructional technique fails, for it is billed (intentionally or not) as the epitome of adaptive training.

MR. WOOD: I think the whole area of adaptive training should be researched in the SFTS. My specific suggestions are:

1. Research on the kinds and effects of knowledge of results applicable to adaptive techniques, including how knowledge of results should be displayed to the student.
2. Determination of the relative effects of continuous versus discrete adjustment of task difficulty, including the development of dynamic techniques for setting time and level of fixed practice sessions.
3. The definition of effective criteria for adjusting task difficulty based upon student performance, for example, performance measures which combine central tendency with the variability inherent in performance.
4. Ascertaining the importance of varying levels of error during practice in terms of the effect on the stimulus characteristics of the task and on subsequent transfer to the criterion task.
5. Definition of the theory of adaptive training: How do theoretical concepts support, constrain, or direct the use of adaptive techniques?
6. Research on how adaptive techniques can be better used to improve the learning process, for example, when to increase task load, what task load to provide, and how to relate task load to skill.
7. Determination of whether or not adaptive techniques have any predictive validity for selection purposes.
8. Research on the effect of the subject pool (experienced versus nonexperienced) on the use of particular adaptive models.
9. Evaluation of the application of techniques used in perceptual-motor adaptive models to adaptive models for the training of verbal skills.

MR. MANSFIELD: In addition to the various problems and questions that have been raised during this conference—the selection of adaptive variables, the development of the training format, the value of feedback, and so on—the possible use of the SFTS for periodic refresher training is something I would like to see investigated.

Mr. Stewart: Many questions about adaptive training remain unanswered; these should be investigated, if possible, in the SFTS. For example, does SFTS adaptive training increase the rate of learning, improve retention, or increase student motivation in comparison with other methods? An experiment on the most appropriate ways of presenting feedback to the student regarding levels of difficulty also suggests itself. The following methods of presenting feedback could be evaluated: (1) warning the student of an impending change, (2) notifying the student concurrently with a change, (3) notifying the student only after a change in difficulty level has occurred, and (4) providing him no feedback at all. A second series of experiments might compare the simultaneous variation of adaptive variables to their one-at-a-time variation. Finally, I think the use of the SFTS as a predictor of subsequent pilot performance should be investigated. Measures of performance, or levels of difficulty attained, during early trials may be highly correlated with measures of final performance.


Dr. Roscoe

Dr. Roscoe: Research programs should be derived from real-world training problems and organized around available training devices having maximum capability for adaptive-variable manipulation. In the aviation field, the SFTS is presently the most advanced device available for this purpose, even though adaptive training research can be conducted using far less complicated equipment. Two questions having relatively high research priorities are:

1. What are the effects of presenting indications of current performance (true or untrue; continuous, periodic, or aperiodic; presented directly or inferentially) on speed of learning, terminal levels of performance, and transfer to criterion tasks?
2. What are the relative advantages of various representative adaptive-variable manipulation techniques as applied to various representative types of tasks in terms of the dependent variables listed?

Mr. Sinacori: I would like to see the initiation of research studies that would provide the following: proof of the validity of the adaptive techniques, guidelines for the selection of candidate adaptive variables, value of simulation versus flight methods for training, the contribution of major hardware subsystems in simulation, and psychological and engineering data that would be available for study by outside groups.

DR. MATHENY: The SFTS would be a good vehicle in which to explore how adaptive logic can be applied to the progression or regression from unit to unit within a task element. Such an effort could help answer some of the questions about measurement needs, the criteria to be used, the computer programming, and the hardware requirements in moving from automated demonstration to guided practice to adaptive practice and back again. In the absence of a detailed knowledge of the SFTS, I can make no further suggestions other than the obvious ones of exploring other adaptive variables, exploring rates of change, and so forth.

DR. LAUBER: I think parametric studies of all kinds should be conducted. For example, what are the optimal rates of adaptation? Our greatest need right now is for numbers. A related but independent problem concerns the efficacy of adaptive techniques relative to nonadaptive techniques. The SFTS could be the vehicle with which to conduct a full-blown transfer study involving real pilot trainees and real helicopters.

DR. KELLEY: The highest priority research should involve the improvement of the adaptive features of the SFTS itself. Although primary attention should go to the problem of performance measurement, both adaptive logic and adaptive variables should be studied. At the next level of priority, the SFTS should be used to explore a broad spectrum of different and more sophisticated areas relating to adaptive training. One example would be diagnostic programming: measuring special complex aspects of performance (pilot-related oscillation, engine-instrument monitoring) and using diagnostic measurements for appropriate instruction to remedy the faults or problems diagnosed. Another would be cross-adaptive training: permitting more and more of a student's skill to be freed for another task, as the primary task is learned.

DR. CARO: It is difficult to imagine an area of application for adaptive training where we know enough to use the technique efficiently. Much of what we need to know relates to what we intend to teach and why we should consider adaptive techniques for that training in the first place. Until we have a reasonable approximation of an optimum curriculum, any research related specifically to adaptive training will be risky at best. How can we conduct research comparing, for example, adaptive training with some other technique unless we know something about the interaction of each technique with the content of the experimental training material?

So far as research directed more specifically to adaptive training itself is concerned, there appears to be much disagreement about the role of feedback. The requirement for a score and for the form that score should take should be researched.

The effects upon performance of adaptation rates, the time sample upon which adaptive decisions should be based, and whether or not deadbands are needed is another area for research. Also, where deadbands are used, the width of the bands, that is, performance tolerances, needs to be determined.

The related areas of performance scoring, such as Patty Knoop described, and the specification of criterion behavior also require research.

Mr. Norman: It should be used to help identify the broad underlying concepts of adaptive training and the particular characteristics that make it superior to nonadaptive training. For initial exploration, my candidate is an investigation of the effect of rate of progression-regression performance criteria and performance evaluation interval on rate of skill acquisition, retention, and criterion task performance.

Mr. Gandelman: I have prepared the following list of research questions and issues:

1. Difficulty functions must be determined for each of the adaptive variables.
2. Distributions of pilot performance for a basic set of tasks should be developed.
3. Research on performance measurement should be undertaken to determine the informational content of alternative measures and correlations among measures.
4. A scheme for "matching" adjustment factors to the individual is needed to more truly adapt changes to an individual's performance.

Mr. Hall: I think the highest priority should be given to experimentation involving the following variables: task difficulty increments and decrements, length of the performance-sampling periods, types of performance measures, and characteristics of the adaptive logic.

Session 3. Related Fields of Research

DR. W. G. MATHENY, *Chairman*

". . . EMPHASIS ON THE *AND SO FORTH*. . . ."

DR. MATHENY: This session will be concerned with fields related to adaptive training, such as computer-aided instruction (CAI), adaptive methods applied to testing and psychophysics, and so forth. The emphasis, I think, will be on the *and so forth*.

Patricia Knoop has been examining the application of adaptive notions to performance measurement and has graciously agreed to discuss this work with us. Also, Chuck Kelley will discuss his use of adaptive techniques in research. In addition, I think we can profitably spend some time in this session on further defining the concept of adaptive training. Chuck, will you please start us off?

DR. KELLEY: I think that a lot of our cousins have been working on adaptive training problems under different labels and that it might do us a lot of good to know about the work being done in self-organizing systems, computer-aided instruction, and whatever. I, myself, must admit to having worked less in adaptive training than in adaptive systems for other purposes.

The first study I ever did in this area concerned what I called self-adjusting vehicle simulators; that was before the word *adaptive* had, itself, gained popular use. The study involved submarine control and illustrated the use of an adaptive variable in design research. The type of submarine under study was a big beast, about as long as the length of a football field, and had to be controlled to within a few inches in depth, particularly when the periscope was up. (A periscope must at times be raised for only very brief looks around and pulled back down again or risk being picked up by somebody's radar.) One of the control problems involved the motion of the horizontal rudders, the planes. These were huge control surfaces, like the rudder of a ship, but located on the sides of the submarine, stern and bow. They were used to control the depth.

One of the critical problems in designing a new submarine is to determine how fast the planes should move. Moving them quickly makes control easier but requires a large motor and high power consumption, whereas moving them more slowly makes control more difficult but permits the use of a smaller motor and lower power consumption. This problem presented an ideal application for self-adjusting simulation. With speed of the plane motor as the adaptive variable, variable plane speed could be simulated and allowed to adjust adaptively. We established a threshold of performance and a rule that, whenever the subject was controlling depth to better than threshold, we would make the plane rate slower; whenever he was controlling to worse than threshold, we would speed up the plane motor, reducing the lag and tightening up the control loop. In this way we could find out what plane rate he needed to control to any performance threshold. The Navy wanted the boat controlled to between six and 12 inches.

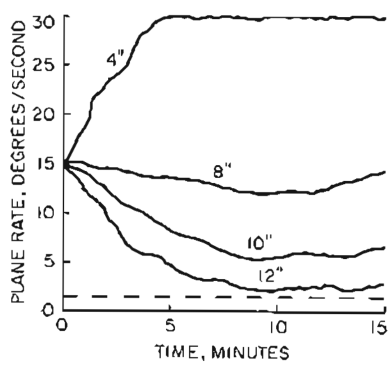


Exhibit 18. Plane rate of a self-adjusting submarine simulator for four different error thresholds. The plane rate adjusted automatically, increasing when error was above threshold, decreasing when it was below. Rates were limited to the range of 2 degrees per second to 30 degrees per second.

Four runs were made by a highly skilled subject under four different performance thresholds (levels of RMS error). These are shown in Exhibit 18. We found that the subject was unable to control to within a four-inch threshold, even when he was given all the speed on the planes that he wanted. However, he could control to within an eight-inch threshold at a plane rate of 15° per second, within a 10-inch threshold at a plane rate of 6° per second, and within a 12-inch threshold at a plane rate of 4° per second. So, the study showed how sensitive submarine control was to plane rate and provided the designer with the tradeoffs that would be made between plane rate and control error.

Exhibit 19 shows the results of letting the error threshold vary slowly and letting the plane rate, our adaptive variable, be the dependent variable. We started at the four-inch threshold and very gradually increased it over a period of about a half an hour to the 12-inch threshold. As the error threshold increased, the plane rate needed to meet the threshold was gradually decreased. This curve, therefore, shows the important functional relationship between accuracy of control and rate of motion of the plane. Adaptive techniques permit you to get a direct plot of this functional relationship.

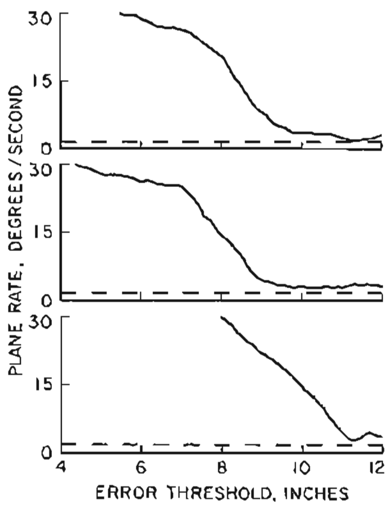


Exhibit 19. Plane rate versus error threshold on a self-adjusting submarine simulator. Error threshold gradually increased, and rate of motion of the planes adjusted automatically from the initial value of 30 degrees per second. The first two records are on a highly skilled subject; the last was a novice.

". . . AN EFFICIENT WAY TO GATHER DATA. . . ."

The next experiment concerned the head-up display in a Navy TFX aircraft. Designers wanted to know how bright to make the display. Because of the inherent limitations of CRT phosphors, generating a bright display is a real problem, so it is important to know the exact brightness requirement. Also, when you fly over bright clouds, you have a background brightness in the neighborhood of 10,000 foot-lamberts. Can you see your head-up display with this kind of background? If you can't, when can you see it? If you want to make it bright enough to see, how bright do you have to make it? The experiment was conducted to tell the designers how bright they had to make this display when using various combining glasses.

For those who are not familiar with it, let me take a moment to illustrate the head-up display concept. A very bright CRT generates the artificial horizon and the other display symbology. The display is projected through optics onto a transparent combining glass placed in the forward line of sight of the pilot and then reflected back to the pilot's eyes from the same position as objects viewed ahead of the aircraft. The pilot looks out of the aircraft through the clear combining glass and sees the information on the glass at infinity; that is, he sees collimated images, displayed as if they were out there, far ahead of him in space. Only a small amount of the energy hitting this glass is reflected—10% would be a pretty high reflectance for a lot of these combiners—because of the requirement to view outside objects through the combining glass.

The adaptive experiment involved a discrimination task in several simulated displays. Subjects were in a simulator containing a very bright visual background generated by a xenon lamp. Then, as we displayed information on the head-up display, every time the subject made a correct discrimination, we made the image dimmer; every time he made an incorrect discrimination, we made the image brighter. The display was made brighter by increments nine times as large as the decrements used in making it dimmer. This ratio was used because we were interested in a 90% threshold; that is, we wanted to know the display that he would see 90% of the time, and this ratio of increment to decrement stabilizes at the 90% threshold in an adaptive system. We had many different combining glasses to evaluate—trichroic combiners, neutral-density filters, and so forth. Exhibit 20 shows some data from one typical run.

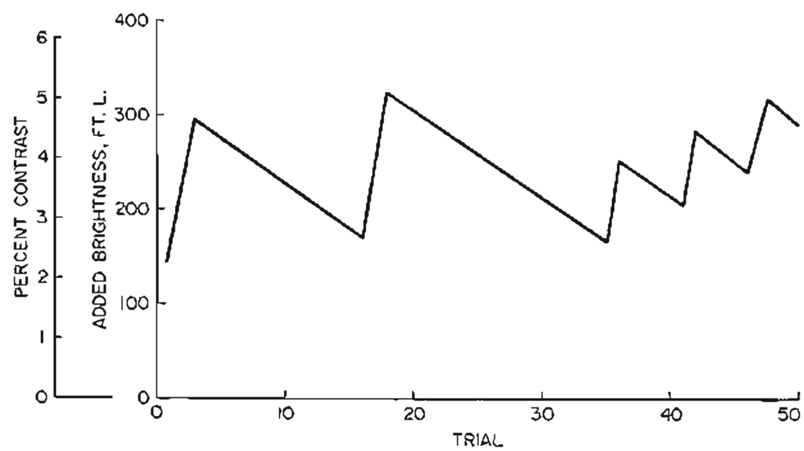


Exhibit 20. Raw data from a typical run in an adaptive head-up display brightness experiment. Brightness was decreased a fixed amount after each correct display reading and increased by nine times that amount after each incorrect reading.

I was pleased with this experiment because it showed that the adaptive technique was an efficient way to gather data and it provided data we could have confidence in. I felt that these data did indeed reflect the differences among the combining glasses and that we were able to supply the Navy with information that was of real value within the context of a very limited experimental budget.

DR. ROSCOE: Chuck, the technique you described seems to be very similar to one of the classic psychophysical methods.

DR. KELLEY: The method of limits.

DR. ROSCOE: I don't see that it is any different.

DR. KELLEY: It is essentially an automation of the method of limits. In the method of limits, the stimulus given the subject is a function of how he performs on a previous trial, and that is adaptation.

DR. LAUBER: I think it's interesting to note that Blough used an adaptive technique to obtain dark-adaptation curves from pigeons. The "running" absolute threshold was determined by having the bird peck at one key if there were no stimulus light present (or one below threshold) and at another key if the light were above threshold. If the bird responded to the *stimulus present* key, the brightness of the stimulus was slowly decreased. If the bird responded to the *no stimulus present* key, brightness was increased, hence tending to clamp the brightness of the stimulus at the instantaneous absolute threshold of the bird.

DR. McGRATH: The method is also useful in studies of signal detection where the detection rate must be the same for all subjects. You see, the number of signals a subject has detected will often influence his vigilance for subsequent signals. In some research problems, you must measure individual differences in signal detection, but you must expose all subjects to the same number of detected signals. This can be done by making signals progressively easier for subjects who are missing them and progressively more difficult for subjects who are detecting them. In this way, a constant signal-detection rate can be maintained, and the intensity of the signal can be the performance measure. It is an exact model of Chuck Kelley's constant-performance-standard technique of adaptive training.

DR. KELLEY: That is a good illustration of a nontraining application. When should it be used?

DR. McGRATH: Whenever you must control the subject's expectancy for signals. You can do this with adaptive methods by controlling the number of signals he actually experiences.

DR. KELLEY: We did a study of spacecraft control in which we compared control in one versus two versus three axes and in which we compared continuous versus on-off control. The adaptive variable was amplitude of the disturbance under control. The results, illustrated in Exhibit 21, showed that control was better with one axis than with two and better with two than with three. Also, continuous control was found to be better than on-off control. All main effects were statistically significant.

This was a very brief experiment. We had to run only a couple of subjects in balanced sequences of five-minute trials. All the data were gathered in a day, after the experimental rig was set up. The brevity of this experiment is all the more significant because people had been studying this set of variables for years without any really conclusive results. The study was a good illustration of the experimental efficiency of the adaptive technique; only a modest amount of data is needed because the data are gathered with precision.

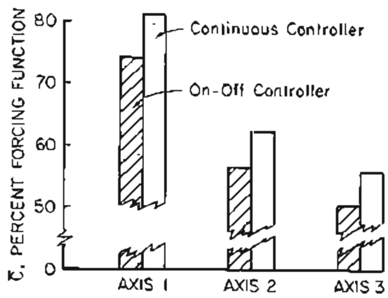


Exhibit 21. Continuous versus on-off control in 1, 2, and 3 axes.

". . . AN ADAPTIVE LOADING TASK. . . ."

There is just one other technique I want to throw at you, and then we can get back to the general discussion. In an experiment on cross-adaptive control, as we called it, the primary task was two-axis tracking in which RMS vector error was the measure of performance. The independent variable was display size—100%, 75%, and 50% of full scale. First, we measured RMS error for each of the three conditions, balancing sequence effects, and so forth. Then we introduced a secondary "warning-light" task consisting of two lights in the periphery. When one of them came on, the subject had to operate a switch in the appropriate way to turn the light off. In the course of a trial, the light came on at frequent, unpredictable times, and the subject kept turning it off. The subject would take longer to turn the light off when he was busy than when he was not. Thus, the total number of lights turned off was used as a score to indicate the level of loading.

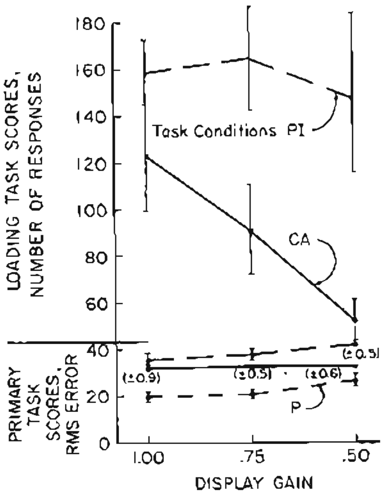


Exhibit 22. Primary and loading task means ± 1 SD for three task conditions: P = Primary task only; PI = Primary plus Independent loading task; and CA = Primary plus Cross-Adaptive loading task.

As shown in Exhibit 22, we obtained no effect on the secondary task score (upper dashed line) at the 75% size, but we got a little at the 50% size—a few more warning lights were turned off when the display was full size compared with when it was half size. These differences were not statistically significant. However, performance on the loaded primary task was significantly worse (lower long-dashed line) than on the unloaded primary task (short-dashed line).

Then we applied the cross-adaptive technique. Whenever the RMS error was less than we wanted it to be, we turned the loading task on; whenever the error was greater than we wanted it to be, we turned the loading task off. Thus, the loading task came on or off depending on how well the subject was performing. The two solid lines of Exhibit 22 show the cross-adaptive task data. An important finding, in terms of the concept of adaptation, was the lack of any variation in RMS error on the primary task (lower line). The variation of the RMS error score was so small that we couldn't even plot it—under each condition it was less than a unit when RMS error itself was 33 units. So, by the use of an adaptive loading task, we stabilized RMS error precisely. This illustrates why an error score cannot be used as the dependent variable in an adaptive task. If the adaptive task is working right, the error score remains constant; the variance is thrown into the adaptive variable.

The upper solid line of Exhibit 22 shows the cross-adaptive loading-task score. Notice how beautifully it reflects the three conditions of primary-task display gain. These score differences are statistically significant and show the cross-adaptive task to be much more sensitive to the independent variable than either the primary task alone or the primary-plus-loading task.

Well, these are the experiments we have done that illustrate the use of adaptive variables. One general recommendation I would make is to spend a lot of time tuning the adaptive task—choosing and testing the adaptive variable, the performance measure, and the error threshold—before starting to gather experimental data.

". . . THESE INTERACTIVE VARIABLES. . . ."

MR. FLEXMAN: Chuck, has this adaptive concept been used in trying to determine an optimum control system for an airplane, in determining flight stability derivatives, for example?

DR. KELLEY: Not very much. Although it has been discussed in meetings like this, I don't think the adaptive concept has been used in handling qualities work.

DR. MATHENY: Suppose you are interested in the interaction between several variables in a trainer—the characteristics of the motion platform, the characteristics of the visual display, and the characteristics of the auditory cues. Instead of conducting a series of experiments with these interactive variables, could you set them up to vary adaptively while an experienced pilot flies the system?

MR. FLEXMAN: Do you mean to vary them simultaneously?

DR. MATHENY: Simultaneously. You can envision that there might be an optimum relationship between motion and audition.

MR. FLEXMAN: Are you going to try it?

DR. MATHENY: Well, we are toying with this notion right now. There is an enormous problem with respect to these interactive variables.

MR. FLEXMAN: Wouldn't that be the same as in a control system where several functions are continuously varied?

DR. MATHENY: There are so many coefficients and equations to vary in effecting a control system that it gets difficult.

DR. KELLEY: Maybe you have to vary them one at a time using a successive iteration approach. You know, change this one awhile, then change that one awhile, and, when you know there's been an interaction, cycle back and forth to get a successive approximation.

Dr. Matheny: It is a challenging and interesting approach. I shall now call upon Patricia Knoop to talk a little about her work on performance measurement, an essential ingredient in this whole adaptive training area.

". . . TWO VIEWS OF PERFORMANCE MEASUREMENT. . . ."

Miss Knoop: I will restrict my discussion to the automated measurement of pilot performance. Also, I will consider the performance measurement system as consisting of three ingredients—the measure that describes the individual's performance, the criteria by which performance is evaluated, and the logic that relates the two.

I think there are two opposing views of performance measurement; their point of disagreement is primarily in the derivation of performance criteria. One view holds that it is sufficient to base the criteria on what experts consider to be good performance. For example, we might base criteria on information in the flight training guides prepared by Air Training Command that specify standards of pilot performance. The other view holds that criteria should be based on what good pilots actually do during operations. To establish criteria on this basis, normative data must be obtained by sampling actual performance.

We have been developing performance measurement techniques for the Air Force T-37 undergraduate pilot training program by pursuing approaches suggested by each of these two views. The T-37 is a twin-jet trainer, the aircraft in which student pilots receive their first jet training. In the current research study, a digital recorder on the aircraft records 32 different flight variables during training missions. After quite a bit of data reduction, a performance evaluation printout and a series of debriefing charts are produced on the ground for use in evaluating the student's performance. The work is being done at Wright-Patterson AFB by the Training Research Division of the Air Force Human Resources Laboratory. Results so far are very preliminary, but I can illustrate the type of student evaluation aid we are developing.

We have found that the relationship between bank and pitch in the lazy-8 maneuver is a good measure of pilot skill, and we have developed techniques for plotting this relationship automatically, so that, at a glance, the instructor and student can recognize performance errors following an in-flight mission. Bank angle is plotted against pitch angle, and this profile is compared with a standard profile. From the example shown in Exhibit 23, it is easy to see that the student has pitched up too fast at the start of each half of the maneuver. Also, from the actual and standard airspeed profiles shown in Exhibit 24, it is easy to see that the student's airspeed dropped way below the criterion values at these points.

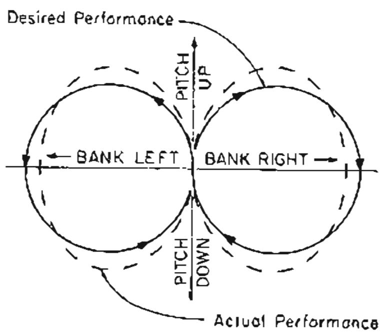


Exhibit 23. Pitch versus bank in a lazy-8 maneuver.

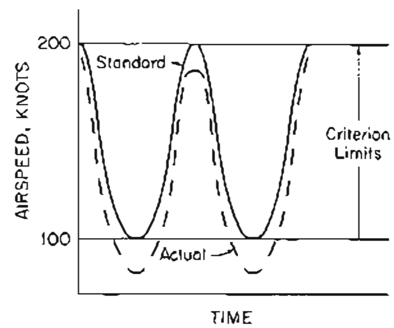


Exhibit 24. Actual versus standard airspeed profiles in a lazy-8 maneuver.

MR. STEWART: Are you measuring against an ideal established mathematically or against normative data?

". . . VARIATIONS ON A SINE WAVE. . . ."

MISS KNOOP: I will discuss that right now. I mentioned earlier two different views of measurement. One evaluated performance against a preconceived notion of ideal performance. Information from Air Training Command, from flight manuals, and from our own research findings, as they developed, provided our initial basis for measurement. The next task, and one on which we are about to embark in May, is to collect normative data at Williams AFB. The idea is to alter the existing measurement programs on the basis of normative data; in doing that, we will have a comparison of what can be accomplished with each of the two approaches, plus data for some work I have yet to describe.

DR. ROSCOE: Patty, it seems to me that it might help if you were to plot a large number of these lazy-8 records on the same piece of paper and then submit them to a Rorschach expert.

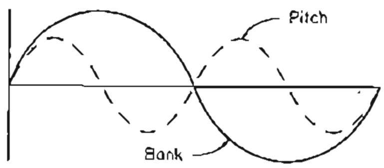


Exhibit 25. Bank and pitch angle plots for a lazy-8 maneuver.

Miss Knoop: That may not be such a bad idea after all. Actually we have obtained some interesting insights from personally analyzing the lazy-8 data. For example, the bank angle plot and the pitch angle plot in a lazy-8 both closely resemble a sine wave (Exhibit 25). These different variations on a sine wave immediately suggest a Fourier analysis, or a harmonic analysis, and have given us some new ideas on techniques for modeling normative data.

Mr. Stewart: It is not surprising, is it, that lazy-8 performance is rhythmic and symmetrical?

Dr. Roscoe: That's the definition of the maneuver.

Miss Knoop: That may be so, but we also found that our preconceived notion of symmetry—that the extent of pitch up should equal the extent of pitch down—was not correct. The pitch down turns out to be much less than the pitch up.

". . . WHY A LAZY-8. . . ?"

Mr. Flexman: I just can't resist getting into this a little bit, Patty. I'm a little concerned with why a lazy-8 maneuver is used as a basis for the standard of pilot performance. If you get your very best pilots, who, in terms of your standards, are supposed to be well standardized and the model of performance, I dare say you will find exactly what we have found in the lazy-8, in the chandelle, and in every other maneuver in the book. The variance among top standardized pilots is greater than that among the students.

Dr. Roscoe: From one pilot to another or the same pilot at different times?

Mr. Flexman: Between pilots. We found, for example, when we required the experienced pilots to make a 60° turn by simply looking outside like the students, their actual bank varied from 35° to 75°. We also had them do lazy-8's while we photographically recorded their pitch, bank, and heading instruments throughout the maneuver. Instead of doing a symmetrical maneuver by coordinating pitch and bank, they got all their bank in within the first 45° of turn. From there on it appeared to them as if they were changing bank because of the pitch change. Also, they took all their bank out during the last 45° of turn. In the past, performance on these maneuvers has been the model that we used to determine the adequacy of training. But, as I have said many, many times, it is the various psychological, if you will, ingredients that have transference to other applications of flight skill. Go ahead and use the maneuver for now, but we should start measuring the psychological variables. Paul, I think you have found the same kind of unreliability with your check pilots.

DR. CARO: Yes, we have, Ralph.

MR. STEWART: The performance measurement program will generate the data needed to give you a range of acceptable performance—everything outside of that range would be excluded. Then you will have an Orwellian kind of automatic check pilot. If the student flies the machine within these limits, he becomes a pilot.

DR. CARO: You are doing the same thing, to an extent, that we are in the SFTS. We are measuring what we can define and what we have had some experience measuring with other techniques.

MISS KNOOP: Even that is difficult enough.

DR. CARO: Yes. We are not as operationally oriented as we would like to be.

MISS KNOOP: Because we do not as yet know very much about performance measurement, we have emphasized the development of hard-copy debriefing aids in addition to the several performance measures that we can validate and compute automatically. We are better at determining what is relevant to performance evaluation than we are at making implications of why and how it is relevant. Thus, we seek a realistic (within state-of-the-art) blend of automated objective measures and instructor interpretation of these measures along with his personal observations of performance to produce an improved performance-measurement technique. As our research progresses, we will be able to expand the role of automation in measurement by interpreting as well as describing salient features of the performance.

The approach with which we have begun our studies is very classical and involves making a priori judgments about the measures that should be computed. We have based these judgments, so far, on existing knowledge and literature about the maneuvers. However, in an attempt to obtain face validity of the measures, we have examined, as well, subjective ratings of each performance and the rater's comments about why he rated as he did. Airspeed is one measure that appears to have high face validity. This, however, does not validate airspeed as a measure for the lazy-8; we have yet to accomplish formal validation, or un-validation, using normative data. In addition, our task is not to develop the training curriculum but to work with the existing curriculum to develop improved measurement techniques.

". . . UNDERNEATH THE LAMPPOST. . . ."

Dr. Matheny: Unless we set up a huge training research program in which we can introduce a whole new curriculum and validate it, we are caught in this circular business of expert judgments.

Dr. Kelley: I wish we could break out of this circle. This has been the problem in research on pilot selection. Without a real-world criterion, selection tests are built to predict the ability to get through training school. Because all of the combat criteria programs have foundered on the rocks, we have never really been able to establish independent criteria of what makes a good pilot. I feel we are like the drunk underneath the lamppost who is looking for the money he lost. Somebody starts helping him only to discover that he lost his money back in the dark, but he is not looking back there because the light isn't good enough. Thus, we measure what we *can* measure although we are not at all sure that it has much to do with being a good pilot.

Dr. Matheny: I think we are on the topic of another seminar, Chuck.

Dr. Roscoe: I disagree. I think that should be the topic of this seminar. I think that it is time that we pay some attention to what makes a good pilot. It seems to me that there are four aspects of pilot performance. One is the ability to manipulate the airplane, the perceptual-motor skills. A second one is procedural; it has to do with communicating, planning, navigating, and cockpit housekeeping. A third is what we've chosen to call residual attention: How much attention does the pilot have left over to take care of the unexpected emergency? Can he maintain his housekeeping and manipulative performance while attending to the emergency? The fourth one, and the one that probably really makes the difference in combat, is what we have chosen to call risk-taking behavior: What kind of ability to predict the outcome of various alternative courses of action does the man bring to the situation? How well can he re-order priorities when the situation deteriorates or improves? Now, those four aspects of pilot performance, although not directly on the topic of adaptive training, cannot be avoided in any discussion of adaptive training. We should look at what it is we are trying to train the man to do in the long run, whether he's going to be a combat pilot or an airline pilot.

DR. MATHENY: I agree, Stan. These are important problems that you can attack by many training techniques, perhaps one of which is adaptive. If I may, I will defer that to the session on "Where Do We Go From Here?" this afternoon.

". . . AN ADAPTIVE MATHEMATICAL MODEL. . . ."

MISS KNOOP: I want to finish my discussion by describing an effort we launched recently to develop a computer-aided technique for deriving performance measures and criteria. Recognizing the size of the effort that would be involved in applying the approach we are using with the T-37 to all aircraft in the Air Force inventory, we want to use the T-37 data as the basis for developing a computer program that can examine data from other aircraft or simulators and, by adapting itself to the critical parameters, can evaluate performance automatically. An assumption of the effort is that a good instructor pilot can do a pretty good job of subjectively evaluating the ability of a student. We hope to refine this evaluative ability through the use of more sophisticated rating methods and through better control of the setting in which the evaluations are made. Also we hope to increase the reliability of these subjective ratings through the use of several instructors rather than just a single instructor.

Subjective evaluations of performance will be used as a guide by computer programs for automatically deriving performance measures from data recorded on the T-37 aircraft. The result will be the automatic development of a measurement program which generates scores that correlate, to the greatest extent, with the subjective expert evaluations. This is accomplished through the development of an

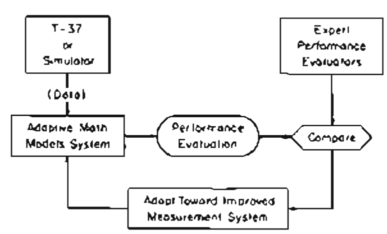


Exhibit 26. Adaptive mathematical models system.

adaptive mathematical model that effectively *learns* to evaluate performance by examining actual data. The adaptive mathematical model that is generated can then be utilized in other aircraft programs. This system is diagrammed in Exhibit 26.

We now have an adaptive mathematical model implemented; it has been tested on a hypothetical task evaluated with a canned scoring method. The task is an ILS approach-procedure-turn and landing. The model converged very quickly to a unique scoring system that produced results closely approximating those of the canned method that it knew nothing about. Of course, this preliminary study provides little substantial proof of the value of the model; we are now testing it with the T-37 data and the expert evaluations.

Mr. Wood: Is it an automatic stochastic modeling technique?

Miss Knoop: Yes.

Dr. Caro: I wonder what happens if you are using, for your instructor evaluations, a pilot who rates everything on the basis of airspeed. It would be very easy for your model to come up with something that correlates very highly with his evaluation, but how useful would it be?

Miss Knoop: That is a good question. Ideally, this approach should be validated experimentally with many different instructors or evaluators giving the best ratings they could.

Dr. Caro: I wonder if you are not compounding bias, rather than eliminating it, when you combine instructors.

Miss Knoop: How do you mean?

Dr. Caro: By using evaluations from more than one instructor, you are not necessarily getting a more valid performance measure.

Miss Knoop: I think you have to admit that it is an improvement.

Dr. Caro: No, I don't have to admit that, because there may be some biases that cause instructors to tend to rate alike. You might be able to eliminate variance among instructors, but you still might not be getting any nearer to an objective representation of what takes place in the air.

Miss Knoop: Perhaps we can use the scores generated by a group of instructors only as initial guidelines, placing more reliance on actual learning data and the ability of the math models to extract information from it.

Dr. Caro: Yes.

". . . FLYING PATTERNS INSTEAD OF MANEUVERS. . . ."

MR. FLEXMAN: Patty, I invite your attention to the Rand report that will be coming out as soon as Bill Stewart completes all his work; I am sure that it will contain an account of the Canadian system. They have looked at performance measurement systems for many, many years, and their findings have led them to be interested in performance in flying airport traffic patterns instead of maneuvers. The guys take off, fly the pattern, and land; this performance has correlated highest with completion-of-training criteria. This is also the direction that the airlines are going because the patterns are closer to what pilots actually do in an airplane; it is very seldom that an airline captain flies a lazy-8, and it is very seldom that a fighter pilot ever does a chandelle.

DR. MATHENY: How do they measure performance in the pattern?

MR. FLEXMAN: A group of evaluative pilots, not instructor pilots, is trained to use a very complex rating form. They categorize performance and use normative data, the percentage of people who do this, that, and the other thing, against which to score performance. This is all they record. On the basis of this score, plus their subjective evaluation, they reach a pass-fail decision.

DR. CARO: We also have found that, by using independent evaluators who do not instruct but who are trained to evaluate, we get more reliable measures.

DR. ROSCOE: Our recent experience has been similar. King Povenmire and the flight instructors at the Institute have developed a descriptive five-point scale; I hesitate to say that it is an objective scale, although it is as objective as they can make it. At least it is quantitatively descriptive of each of five proficiency levels for each aspect of performance on which a student is rated for his private pilot ticket. From observer-observer flight checks, where we can have a check pilot in the right seat and the student's instructor in the back seat while the student is flying, we have found the reliability of these independent ratings, for most items, to be in excess of .70.

DR. MATHENY: Patty, would you like to summarize?

". . . ALL THE APPROACHES WE CAN THINK OF. . . ."

MISS KNOOP: Yes. I wish to reiterate that I am not really researching the area of adaptive training, except where performance measurement plays a role; of course, I think it plays a very critical role. I think that we should not be premature in our adaptive training research; we should have a better capability in the area of measurement before going very far because so much about adaptive training depends on how well we are able to measure performance. In the T-37 program, we are taking all the approaches we can think of to performance measurement. One is a very classical approach, using as a guide expert judgments of good performance; in another, we are relying more on normative data than on expert judgment; in a third, we are attempting to develop an adaptive mathematical model to facilitate the measurement of performance on a larger scale.

Mr. Flexman: Patty, I would like to give my enthusiastic endorsement of what you are doing. I personally feel that the success of the performance measurement system will come only when we can control the conditions under which a pilot is flying; that means using a simulator. However, I do not share your concern about the lack of adequate performance measures for adaptive training. Performance measures are used in adaptive training as a basis for increasing or decreasing task difficulty. It is really very easy to measure time-out-of-tolerance on altitude, airspeed, or the other parameters that define the necessary performance levels that have to be achieved at the end of the program.

Dr. Matheny: Thank you very much, Patty. Perhaps we can address ourselves now to some of the questions that you all have indicated an interest in discussing. What is the evidence that adaptive training is more efficient, produces a better quality product, or improves retention? Is there anything inherent in this technique that gives us an advantage in any of these areas? What are your comments?

". . . ALMOST AS AN AXIOM. . . ."

Dr. Kelley: It seems to me that the assumption underlying adaptive training underlies a lot of other training as well; that is, there is an appropriate level of difficulty for the tasks you give students, and level of difficulty can be too high or too low, making the task too difficult or too easy. I accept that as a reasonable assumption; almost as an axiom. If training tasks are much too easy, not much training is going on; if they are much too difficult, not much training is going on. If one can accept that as an assumption, then adaptive training by its nature has the advantage of adjusting the level of difficulty to some predefined level. I don't think any other assumption needs to be made to justify the adaptive training approach.

MR. STEWART: Of course, the level of difficulty is controlled in all training. In adaptive training, the difficulty level is adjusted more immediately to the performance of the student. But the question remains: If you start ten students in a good adaptive training program and ten students in a good conventional program, which group will arrive at the desired performance quicker?

DR. MATHENY: What is the empirical evidence, if any, that they learn any given task faster?

". . . A HANGUP IN THINKING. . . ."

DR. MCGRATH: May I first question Chuck Kelley's basic assumption or, at least, the language in which it was put? I would not say that the basic assumption of adaptive training is that there is an optimum level of difficulty for different stages of learning but, rather, that there are certain stages of learned skills that need to be achieved before the total task can be performed. In other words, it is not necessarily a question of arraying a series of tasks from easy to difficult. It may be that certain skills must be acquired before other skills can be acquired, and one may be no more difficult to acquire than the other. In such a case, the order of prerequisite skills, not the order of task difficulty, would determine the training regimen. For example, to apply adaptive techniques to cognitive tasks, one would not necessarily manipulate task difficulty but would stage an orderly process of concept acquisition. I believe the use of the term *difficulty* produces a hangup in thinking about adaptive training because one starts looking for variables that degrade or enhance the performance measure, yet these may not be the variables that are the keys to learning.

DR. KELLEY: I think I agree with just about everything you said. However, I did not say, "An optimum difficulty level;" I expressly tried to avoid saying that. I did say that the effectiveness of a training program is reduced when the task is too easy or too difficult. And I grant your last point that *too easy* and *too difficult* ought to be further defined. Still I do not believe that we can get away from the concept of easy-difficult in any training program where we evaluate performance. Some people score better and some people score worse, and better-worse corresponds exactly to the dimension of easy-difficult.

DR. ROSCOE: I would like to get back to Jim McGrath's comment of a moment ago. As I understood it, his point was that the adaptive variable does not necessarily have to be metered in terms of difficulty. It can be, as in the case of concept formation, metered in terms of the amount of information learned. When you have learned certain things, you proceed on to other things.

DR. MCGRATH: Much like programmed learning of verbal materials, where certain branches take you back to a review of previous material or ahead to new material.

DR. ROSCOE: Exactly the point.

DR. KELLEY: But, do you ever really get away from the concept of difficulty? Performance is always measured in terms of accuracy or speed or some concept of better or worse—items that are right or wrong, training materials that are more or less difficult.

DR. ROSCOE: I don't think there is any implication that the twelfth week of a course is more difficult than the fourth week of a course.

DR. KELLEY: That is not what I mean. When a variable adapts, in the sense that I have used the term, there is a functional relationship between the measure of performance and the change in the adaptive variable. As the subject performs better, meaning that his score begins to improve, the adaptive variable moves the task to a higher level of difficulty. If there isn't any relationship, any feedback, between the change in the task and the performance score, there is not going to be any adaptation. The many training situations that do not contain any measured differences in task difficulty simply do not lend themselves to the concept of adaptation.

DR. MCGRATH: But, you see, the emphasis upon making the task easier or more difficult as being the key to producing an adaptive system can lead to misconceptions in the design of adaptive trainers. It is a simplistic notion which might lead the system designer to concentrate on making the task manipulatable in difficulty while ignoring the elementary principles of staging the learning experience. For example, in your *Human Factors* article, Chuck, you said, "The only essential requirement is that an adaptive variable systematically affect the difficulty of the task . . ." (Kelley, 1969, p. 552). That statement can easily give the impression that it does not matter much what adaptive variable is used as long as it can be manipulated to vary task difficulty.

DR. KELLEY: It is the only essential requirement for making a system *adaptive*, and that is true. But it says nothing about the training validity of the system. In the same article, I said that a training device could be made less rather than more effective by being made adaptive in the wrong way. Jim, you could clarify something in my mind if you could explain to me how you could make the system adapt if you could not adjust a variable that had something to do with difficulty. Just explain how it might be mechanized.

DR. MCGRATH: I can see how you could adjust a variable that did influence difficulty but was irrelevant to the operational performance you are trying to train.

DR. KELLEY: I can see that; there is no question about it, but you have begged the question. Can you take a variable that does not affect difficulty, as reflected in the performance measurement, and make it adaptive? I cannot see how it could be done.

DR. MCGRATH: I am thinking about applications beyond the training of system-control tasks, applications in which the adaptive function is in the problem presented to the student rather than in the state of some forcing function. For example, an electronics maintenance trainer might adaptively present a series of malfunction-locating or circuit-tracing problems to the student. When he can solve one kind of problem, he is given another. The sequence of problems could be arranged according to relative difficulty level, or the sequence could be arranged so as to lead the student through the logic of the circuitry without regard to which points in the logic are intrinsically more difficult than others. In the end, all problems would be presented to the student by either method, and the mechanization of the trainer would be the same by either method, but one method might develop the required skill and knowledge more effectively than the other.

I did not mean to imply that manipulating task difficulty is irrelevant but only that it should not be the central focus of adaptive training. Not everyone thinks of task difficulty in the broad sense that you do, Chuck, nor is everyone so aware of the pitfalls in selecting a valid adaptive variable. So, my concern is that an emphasis on varying task difficulty might be misplaced and misunderstood. In fact, I would not be greatly surprised if someone designed an "adaptive" trainer for typists in which the typewriter stand was systematically vibrated.

MR. WOOD: I have found that I have to distinguish between objective difficulty and subjective difficulty; they do tie together. We attempt to measure objective task difficulty by means of some average amplitude of a forcing function through our models to maintain constant some level of subjective difficulty. I liked Stan Roscoe's term, *residual attention.* What we are doing in the adaptive process is maintaining constant some level of residual attention or task load.

Dr. Roscoe: Like that measured in Chuck Kelley's experiment involving the side task of the lights.

Mr. Wood: Right. Now, I go along with Chuck Kelley that you do have to choose some objective measure of task difficulty and vary this as the student learns. However, his subjective level of difficulty is going to remain constant.

". . . UNWISELY RESTRICTING THE USE OF ADAPTATION. . . ."

Dr. Roscoe: I agree with Jim McGrath rather than Chuck Kelley on this matter. I believe that we are unnecessarily and unwisely restricting the use of adaptation to some sort of a fixed process, like tracking, or some repetitive operation where the task remains essentially the same. Jim's point is that we can use the adaptive concept where the task does not remain the same—where new elements of performance are added, where new items of information must be mastered, where new concepts must be formed, and where the person progresses from part-tasks to a whole complex task. Progression of this sort is not readily metered in terms of difficulty. It involves the transition from one subtask to another, readily and on a time-shared basis. By restricting ourselves to performance on a tracking task or some other simple, repetitive, steady-state type of operation and varying only difficulty, we are throwing the baby out with the bathwater.

Dr. Kelley: We all concede that we should not restrict ourselves to tracking tasks. However, I am not yet at all clear on how we are going to get rid of the concept of difficulty.

Miss Knoop: You are talking about programmed learning, Stan.

Dr. Roscoe: I certainly hope so! You mean that is not adaptive?

Miss Knoop: Do we include programmed learning in adaptive training?

Dr. Kelley: I would.

Dr. Roscoe: That is the point I am trying to make. By adaptive training we should mean some automatic variation in what happens next as a consequence of the immediately preceding performance of the student. That is as far as I would limit the definition.

DR. MONTEMERLO: I don't think you need to restrict the definition, either. You asked for an example of an adaptive variable that could be changed along a dimension other than easy-to-difficult. I did a task analysis on what it takes to learn Lebesque integration. I found that you have to know some real-variable analysis, some algebra, and some calculus. In teaching someone Lebesque integration, it may take a year to get him proficient in algebra. Teaching algebra may be more difficult than teaching how algebra and calculus are combined in performing Lebesque integration. That may take only three weeks. Thus, you have the difficult part coming first. When you get to the criterion task, Lebesque integration, all these very difficult parts can be put together very simply. You do not have a linear procedure because you have tasks that must be learned earlier but are not necessarily less difficult than tasks that are learned later.

DR. MCGRATH: I would prefer to view the training process as a progression of behavior, rather than a handling of easy-to-difficult tasks. Some progression of behavior has got to be defined to model the training process, and that progression may or may not be sensibly defined as easy-to-difficult. It may all be semantics, but semantics are important in stating principles.

DR. MATHENY: Let me exercise the Chair. We are not getting any closure on this problem.

DR. KELLEY: May I make a very quick reply to Mel Montemerlo? In Lebesque integration, or any other example that you care to name, if the student gets all problems right under all circumstances, the task is too easy; if he gets all problems wrong, the task is too difficult. I don't mean anything more than this by the terms *easy* and *difficult*. Choose different words if you want, but performance is measured, and people are scored, along *some* dimension. With some forms of learning, perhaps difficulty is not the appropriate dimension of measurement, but if it is not, there must be another dimension of measurement that plays an analogous role in the adaptive system.

". . . UNENCUMBERED BY EVIDENCE. . . ."

DR. MATHENY: Bill Stewart has asked the question, does the introduction of adaptive training get anyone anywhere quicker in terms of reaching a proficiency level? What evidence do we have that it does?

Dr. Roscoe: We can discuss this at length, because we are unencumbered by any evidence one way or the other.

Mr. Stewart: The application of adaptive principles to tracking has been done for a perfectly good reason—it makes a nice, neat, elegant, controlled experiment. But, in the process of conducting these experiments, I was hoping somebody would draw off a little information as to whether or not learning was faster. Milt Wood has obtained some evidence to the contrary. He found that a staged instructional program resulted in better or quicker learning than a continuously adaptive program. So where are we?

Mr. Flexman: I think that we have enough history on the adaptive concept to say that it facilitates learning. We use it in all our educational processes. We start a student out on easy tasks and, as he progresses, give him more difficult tasks.

Dr. Montemerlo: I disagree with that. I think most often in schools you go from the difficult to the simple. That was the message I tried to bring forth in discussing my analysis of Lebesque integration. The difficult parts were learned first; putting them together was easy.

Mr. Flexman: Well, difficult and simple involve the factor of time too. Indeed, it is easier to teach algebra to a college student than to a first grader.

Mr. Sinacori: Can't we simply be optimistic about it and, by accepting the fact that all instruction up to now has been adaptive in nature, view our efforts as looking for more optimal techniques?

Mr. Wood: I think it is important to distinguish between verbal skills and perceptual-motor skills. Generally speaking, adaptive training has addressed itself only to the training of perceptual-motor skills.

Dr. Roscoe: To the contrary, there has probably been more adaptive training done on verbal skills than there has on perceptual-motor skills.

Mr. Wood: Yes, but we call it programmed instruction.

Dr. Roscoe: I don't care what we call it.

Dr. Matheny: We should, I suppose, clarify the distinction between adaptive training and computer-aided instruction. Milt Wood has pointed out that we generally associate adaptive training with psychomotor behavior, closed-loop tracking, for example. On the other hand, most of the evidence being used to support the value of adaptive training comes from the literature of other areas. To justify drawing upon this literature, the notion of adaptive training has to be expanded beyond the realm of automatically adjusted, closed-loop, tracking behavior.

DR. ROSCOE: Add "Mahoney" to the list of converts.

". . . THE DEFINITIONAL HANGUP. . . ."

DR. CARO: May I get into this? Using an adaptive training model, we use a computer to modify the material presented to the trainee in some fashion, based on how the trainee is responding. *Computer-aided instruction* might be a better term to describe this. We can get hung up for a long time on what is adaptive training and what is not. I propose that we form a new club called, the Society for the Elimination of the Term *Adaptive Training.* This might get us over the definitional hangup and on to discussing how we might use various tools, such as a computer, to increase training efficiency and to advance the technology of training. Use adaptive techniques where appropriate, but do not be so concerned about staying inside some boundaries that we devise to limit their scope. Part of our trouble is treating adaptive training as a training technique rather than as an approach to the adjustment of training difficulty.

DR. KELLEY: I would like to speak to that. I hope that we don't change the name from *adaptive training* to *computer-aided instruction.* We should continue to apply ourselves to human learning broadly considered, encompassing perceptual-motor tasks and cognitive tasks as well. Furthermore, we should not tie ourselves to a specific application—a computer may not necessarily be involved in adaptive training. It was just this need for a fundamental definition that prompted me to write out, first thing, what I meant by adaptive training. Maybe someone would like to take exception to that definition or reword it.

DR. LAUBER: Yes, I would like to. I think inclusion of the word *automatic* has thrown us off. Aren't we simply concerned with formalizing a decision logic and a system for performance evaluation that has always been used in educational contexts by a good instructor?

DR. KELLEY: We are, as long as that logic is conceived to include changing the nature of the task as a consequence of a performance measurement.

Dr. Lauber: An educational setting has always involved feedback, even when the teacher makes the evaluation. That is, the result of the teacher's evaluation changes the task or the subject matter or whatever. It does not matter whether the loop is closed through an instructor or through a machine. But, in order to close the loop through hardware, we have to be able to formalize the logic that the instructors have been using to make these evaluations.

Dr. Roscoe: Well, I guess we are waiting for you to drop the other shoe.

Dr. Lauber: And, we don't know how.

Dr. Roscoe: I think the definition must include some delimitation of how; otherwise, it is no definition at all.

Dr. Lauber: I was not attempting a formal definition.

Dr. Roscoe: I think that maybe you were on the edge of one, though.

Dr. McGrath: John, you have defined precisely what the developers of adaptive training are doing: formalizing a long-standing educational concept. However, by the practical experience that has been gained, some key variables that have never been brought to light before have come out. For example, time variables prove to be extremely important in the practical application of adaptive methods. The period over which the performance is measured and the rate at which the adaptive variable changes are concepts you will never find discussed in textbooks on educational practice.

Dr. Lauber: I agree. That is what I see as being the formalization of it. We are looking closely at the logic, and all of a sudden we are finding all these relevant variables. Instructors have certainly not been aware of these; they have been making intuitive evaluations based on who knows what.

Dr. McGrath: There is another concept in adaptive training research that is different from classical educational research. In the applications that have been described here, backtracking, making the task easier, takes on just as much importance as making the task more difficult. In classical teaching methods, the emphasis is almost always upon staging up to another level, very rarely on staging back again.

Dr. Lauber: Probably because the rate of progression is usually so slow that backtracking rarely becomes necessary.

Dr. Kelley: It is because in training you are normally working with independent variables in which there are successive changes. Normally, as a person learns, he moves along some dimension from (if you'll pardon the expression) easy to more difficult, or simple to more complex. But some independent variables, such as those that affect visual acuity, do not progress in this manner. It is important to be able to make the target smaller for people who have good visual acuity or larger for people who have poor visual acuity. With these variables it is important to go in both directions.

". . . THIS IS OUR MICROSCOPE. . . ."

DR. CARO: I think in adaptive training, we are trying to simulate a competent instructor and to make his decisions more rational.

DR. MONTEMERLO: Do you mean, simulate the entire instructor or just one or a few aspects of the instructor?

DR. CARO: Well, we are really simulating the instructional process. In the case of the SFTS, we are using a computer to handle a few of the decision-making functions that the instructor usually performs. We hope, of course, that through use of the computer we will make the process more reliable and rational. As we gain experience with the SFTS, we should be able to more nearly accomplish this.

MR. WEEKES: Perhaps Patty Knoop's 32 measurements, when she is able to evaluate them, will help us.

DR. CARO: She will be able to help us make our computer program more rational.

MR. WEEKES: Yes, and make it better able to do what the instructor now does.

DR. CARO: Perhaps we will even be able to improve upon the competence of the most competent instructor.

MR. FLEXMAN: I would like to retain the concept of adaptive training as something new and different because, as was mentioned by Jim McGrath, it is leading to new information on learning theory. I have not been particularly impressed with what the learning theorists have given us thus far to help us structure training or educational programs better.

As in any laboratory work, we have tried to enlarge some of the critical elements so we can see them better. This is our microscope for looking at a learning process. I am very enthusiastic about our attention being focused even more minutely on some of these variables associated with what we call adaptive training.

MR. WOOD: You were talking about starting a club for the elimination of *adaptive training*, Paul. We could take an opposite view and start a club for using the term *adaptive*. *Programmed instruction* is the older term, but I really think the more global concept is adaptive training. It provides the good things of programmed instruction —self-paced and individualized instruction—and also provides control over residual attention.

MR. FLEXMAN: I think the role of the computer is excellent for assisting us in developing our understanding, Paul, but I hope principles and methods that can be applied to environments free of the restrictiveness of a computer will come out of this work as well. For example, using what we've learned thus far about the sampling period and the rates with which we can increase or decrease difficulty, perhaps an observer can adjust difficulty against observed performance without the need of a computer.

DR. CARO: Let me define *computer* as some device for facilitating rational decisions.

MR. FLEXMAN: It might even be a human.

DR. CARO: That would make it real tough, but I will concede that possibility.

MR. STEWART: I am concerned that a purely adaptive program may drive the learner harder than is desirable. When he begins to err, and you degrade the task rather sharply, it seems to me that you take him down to a low plateau from which he builds up again very slowly. Adaptive training may work against the operation of natural learning plateaus in which the learner solidifies his learning before progressing further.

MR. FLEXMAN: Great subject for research.

DR. KELLEY: You could also say that an adaptive task may drive him easier than is natural. Is there anything about the task being adaptive that would necessarily drive him harder? It depends on the configuration of the task.

MR. STEWART: It does.

". . . WE ARE MAKING SOME MISTAKES. . . ."

MR. POVENMIRE: I'd like to say something with reference to Paul Caro's remarks. When you talk about doing something the way a good instructor would do it, you seem to be considering just one function of a good instructor, that is, telling the student when he is out of limits and adjusting his task. But, adapting to an individual learning problem may be important also. Consider a naive subject trying to hold altitude in a flight trainer. He is told that he is below or above altitude, and the turbulence is reduced until he can get within tolerance. Then the turbulence is increased a little, and he is told that he is above or below altitude, and so on. Proceeding in this way, he would not get past a certain level of performance. On the other hand, if he were told that he was fighting something that the airplane would do by itself and if he were shown this sinusoid oscillation, he would have the insight needed to progress further. He could go on to adjust his control pressures to coincide with the natural oscillation.

DR. CARO: Telling the student when he is out of tolerance and adjusting his task are those instructor functions most relevant to adaptive training. In the SFTS, adaptive training is not the only automated training technique. In fact, it is a relatively minor one. Functions, such as verbal coaching, also are at least partially automated. I agree that there are many aspects of the learning situation that we're not yet smart enough to automate. To paraphrase what Chuck Kelley said yesterday morning, at this point in our ignorance, we are making some mistakes and only doing part of a job, but we are learning.

DR. MATHENY: Although the Chair is supposed to summarize at this point, the best I can do is to provide a couple of observations. The introduction of simulators has caused us to become concerned about what we can do with training simulators, and that has got us to be a little more objective in the training objectives. Adaptive and computer-aided instruction and all the rest have caused us to be very specific about the techniques we use, what we are trying to teach, and how we are going to measure what is taught. I think the adaptive logic and the mechanization of that adaptive logic are somewhat incidental. You do that part of it in the best way you can within the hardware restrictions and the state-of-the-art.

Session 4. Where Do We Go From Here?

MR. RALPH E. FLEXMAN, *Chairman*

". . . WHAT EVIDENCE DO WE HAVE. . . ?"

MR. FLEXMAN: As we structured the conference initially, we thought it appropriate first to review the basic issues of adaptive training, then examine some specific applications, and finally to make some recommendations as to where we go from here. However, it hasn't worked out exactly that way. We have been identifying and discussing basic issues during most of the conference. Perhaps this is for the best because we cannot begin to realize the potentials of adaptive training until we address ourselves to these issues. So, although I hope that I will be able to summarize where we should go from here at the very end of our session this afternoon, I would like to get to that point by first reviewing the questions that have been raised. First, what evidence do we have on the advantages of adaptive training over other training concepts? Do any of you have any specific evidence that you might bring up at this time? Chuck, can you point to any specifics that suggest unique advantages?

DR. KELLEY: Adaptive training should be better because it has theoretical advantages, but empirical evidence that it truly is better than nonadaptive training does not really exist. We do not know for sure that adaptive training will shorten learning time, although we hope it will. Also, we do not know for sure that training will be more effective in terms of some ultimate criteria because we very seldom have any ultimate criteria with which to measure training effectiveness.

MR. NORMAN: I have some suggestive evidence that adaptive training improves retention and accelerates learning. The performance of a fixed-progression group was compared with the performance of a continuously adaptive group in learning to fly through turbulence in the UDOFTT. Two training sessions were conducted with each group; the two sessions for each group were separated by about one week. The continuously adaptive group reached criterion performance by the end of the first session and continued at this level in the second session. The fixed-progression group, on the other hand, did not reach criterion performance in the first session and actually performed poorer at the start of the second session.

MR. FLEXMAN: In applying the adaptive concept to the SFTS, we have found many problem areas. I asked Julius Gandelman to list some of the problems that we have run into; he gave me about four pages of problems. It is obvious that there are too many research issues for us to take up here; however, I did ask Julius to discuss briefly some of the more important problems.

". . . IF WE KNEW THESE RELATIONSHIPS. . . ."

MR. GANDELMAN: I am at a loss to know where to start; all of these problems seem important to me. Furthermore, I am frustrated because the adaptive process has not been well stated analytically and because neither theory nor experiment shows an explicit relationship between task difficulty level and the performance of the trainee. In Chuck Kelley's scheme, a constant performance standard is maintained, and difficulty is related to some ancillary variable. He says that difficulty level represents student performance, but he does not say exactly how it is represented. Is that correct?

DR. KELLEY: I think not. The relation between the difficulty level (the adaptive variable) and the performance measure is precisely and explicitly stated by the adaptive logic. I have employed a logic that changes the derivative of the performance measure so as to maintain a constant performance over the period of learning or across the conditions that form the independent variable of the task. However, I don't say that this is the way all adaptive tasks should be structured or the way that all adaptive training devices should be built; there are other considerations.

MR. GANDELMAN: All performance change is reflected in the adaptive variable; so one infers that the adaptive variable reflects performance. I would like to see an equation expressing difficulty as a function of learning or performance. I find it somewhat disturbing that one has not been forthcoming.

DR. KELLEY: What kind of an equation would satisfy you? I don't quite understand your point.

MR. GANDELMAN: Can you show me, for example, distributions of performance measures for various levels of turbulence and how these distributions will change with trials over time? This is what I would like to see. If we knew these relationships, maybe we could be a lot more efficient in choosing the adaptive variables and selecting their ranges of variation.

DR. CARO: Julius, you want to see the relation between the intensity of the adaptive variable and performance. The problem is that performance adapts to each adaptive-variable intensity level.

DR. MATHENY: Nevertheless, I sympathize with Julius Gandelman. The fact is that, if you are going to introduce turbulence as the variable in an adaptive training situation, you have to decide how to adjust it as a function of some limits of performance. Should the range be plus or minus 2,000 feet or plus or minus two feet? Obviously, the task is going to differ as a function of the range chosen. You have to have some information about these relationships just to apply the adaptive concept.

DR. KELLEY: You even need that general kind of information to design a nonadaptive training system.

DR. ROSCOE: Well, maybe it is time to start talking about where we go from here; it seems that there are several questions that we could answer experimentally. Before listing these, I think we need to agree on the way we will eventually evaluate different systems. It seems to me that the only sensible way of evaluating any training system, or technique, is in terms of its transfer to some operational criterion task. For example, helicopter training has to be evaluated in terms of the effectiveness of helicopter pilots in operational situations. Of course, this approach is difficult because we have a hard time getting that kind of information. However, if we can agree that a measure of transfer to eventual performance on an operational task is what we are striving for, some of these questions are very simple to state and perhaps not even too difficult to test, particularly if you have versatile and flexible and powerful instruments like the SFTS.

The question of the importance of knowledge of results in adaptive training situations has come up repeatedly. Chuck Kelley feels that it is very important although he has asserted that he has no evidence, from any experiment, to support his feeling. Others here have conducted experiments in which they have not bothered to show to the trainee the momentary status of the adaptive variables. So here is a simple experiment that can be done. In the course of doing adaptive training research, we should include a control condition to give us a direct answer as to the relative advantage of knowledge of results.

Next, there are a number of questions concerning the relative advantages of different ways of manipulating one or more adaptive variables. The answers are not theoretically important, but I think they are of considerable practical importance to anyone who is trying to develop an adaptive system. Different experimenters have already used different means; these could very well be the subject of some experimentation.

A third question concerns the spectrum of tasks that are appropriate for adaptive applications. We have talked about applications in the learning of tracking tasks, procedural skills, concepts, and so on. We need some idea of the spectrum of applications of adaptive techniques in training. Although not profound research efforts, I think the results would have considerable practical value.

MR. FLEXMAN: Most of the formal work to date has been done in the perceptual-motor skills; I think more effort should go into applying adaptive techniques to the acquisition of cognitive skills and then comparing the results to those obtained from the application of the branching logic of programmed learning. Considering the problems that we have in adaptive training, we can have as vast an experimental program as you have in learning. If we understood all the essential variables in adaptive training, we would understand the learning process too.

Gene Hall, would you like to take this opportunity to summarize or direct our attention to a particular problem?

". . . SELLING IT TO THE MANAGEMENT. . . ."

MR. HALL: Here is what I think is happening. We all implicitly believe that there is some value in adaptive training, but none of us really knows what this value is. We have faith that a general transfer will result from adaptive training; furthermore, we hope that we are doing more than just teaching the trainee to cope with the adaptive variable itself. Julius Gandelman's concern has been that you may not be teaching the trainee to fly better but only to cope with increasing levels of turbulence. In the process, you are forcing on him a special flying technique, one that may not necessarily be related to the general flying technique you are trying to develop. If all we are doing in the SFTS is automating the presentation of turbulence, adaptive training would not seem to be very worthwhile.

Mr. Flexman: I think your point is well made. However, most pilots would probably agree that general flying skills are being developed while the student learns to handle a greater level of turbulence—his response rate becomes quicker, and his sensing capabilities and residual attention are expanded.

Let me call upon Jack Mansfield to give us a layman's view. Jack, what kind of an impression do you have as to where we are and where we ought to be going to end up with something useful for the applied end of the business that you represent?

Mr. Mansfield: I underscore your use of the term *layman's view.* My orientation is certainly more toward practical operations than toward training theory. We at American Airlines are hopeful that we will benefit from what is learned with the SFTS and that we will get some guidance on how to more effectively utilize the training equipment that we have already purchased. American has just recently purchased the DC-10 simulator with programmed instruction capability; right now we are concerned with how to fully utilize it. We are concerned that our instructions to the programmer will not be sufficiently complete and that, when we do get the ability and knowledge to put together an effective training program, we will have left some critical requirements out of the contractual agreement. One of the primary problems that I foresee with something like the SFTS is selling it to the management people in our business in a way that will lead them to give it a fair trial, to put it to test.

". . . STILL AT THAT MELANCHOLY OCCUPATION. . . ."

Mr. Flexman: I am going to call next on a guy who has been applying adaptive concepts over many years of pilot training, Dick Weekes.

Mr. Weekes: My function at this conference, as I view it, is to seek answers. Therefore, I do not have any answers for you, only questions.

My first question concerns a discontinuity between my understanding of adaptive practice and that which appears to be the view of some of the other participants. This two-day symposium was opened with a discussion of the definition of adaptive practice, and now, at the end of the conference, we find that we are still at that melancholy occupation. I am reminded of the play characteristic of golfers; the group progresses around the course as a team, each man taking his turn at play, yet each man addresses himself only to his own ball and to his own difficulties.

Before I started writing the SFTS program, I thought I understood what constituted adaptive practice, and my conferees nodded their agreement as I expounded my views. My understanding was essentially this: The complexity or the difficulty of the task is adjusted to the student's ability to cope with it; specifically, the student is relieved of as much of the total task workload as necessary to allow him to concentrate on some subtask or some component sequence. As the student becomes more expert and more experienced, he progressively assumes responsibility for greater proportions of the total task. These total task requirements are automatically added to or subtracted from his current workload, as necessary, to assure that each trainee will be working on a problem at a difficulty level optimally suited to his stage of training and relative skill.

With my conferees nodding their agreement, I went to work without a worry. As I began construction of the course of study, I was careful to introduce course content in small increments of advance, with detailed presentation of concepts, and with extended practice in manual and perceptual skills. At the same time, I made considerable effort to so structure each lesson as to permit automatic manipulation of the sequence of presentation of course content as necessary to accommodate to a specific unique difficulty of one student. But after I completed the scenario with a satisfied sigh, my advisors said, "That is wonderful. Now let's exercise the student in adaptive practice." My reaction to this suggestion was, "Heavenly bodies! If what I have been laboring to accomplish is not an application of the principle of adaptive practice, then whatever in this world *is* adaptive practice?" That spotlights my lack of understanding.

Historically speaking, teachers have been utilizing the principle of adaptive practice throughout the ages to prevent instant failure. We let a toddler clutch our finger as he attempts his first steps, relinquishing control to him as he becomes able to manage his own support and balance. We employ the same procedure financially when he gets his first job but continues to live at home and eat at the family table. The swimming teacher gives as much support to a nonswimmer as the student requires until he learns how to keep his head above water. We put safety wheels on bicycles. Similarly, the flying instructor assumes as much of the total flying task as necessary to prevent the student from being overwhelmed by the diverse requirements of the total task and to permit him to concentrate on the specific subtask under scrutiny.

Mr. Flexman: Welcome to the club, Dick. Let me turn to someone who is in the business of building trainers. Sandy, what is your view of this problem?

Mr. Green: I have a few general comments about adaptive training. At the present time the primary emphasis on the use of adaptive techniques is in the area of flight simulators. However, I would like to see more interest in their potential application to operational, real-time systems. I make that statement because these systems, such as new military weapons, appear prior to any formal training, and there will be an increasing need to train qualified people to operate one-of-a-kind systems.

Previously, systems were more generalized in use. In the weapons systems realm, this is not true any more. Even in the piloting realm, there are so many differences between the aircraft that adaptive training is going to be a form of on-line, real-time implementation. What format they will have, no one can really delineate at this point. However, I foresee a heavy emphasis in that area. For instance, I think the determinants of transfer functions are going to be derivatives of highly specific system designs and applications. I think the development of adaptive training systems and the development of adaptive operational systems are going to have to go down parallel roads, and at some point they will be joining and, of course, correlating information.

". . . A VAST FIELD FOR OUR EFFORTS. . . ."

Dr. Roscoe: I would like to amplify my earlier remarks about the desirability of studying the effect of knowledge of results in the adaptive situation. Guy Matheny winced when I brought the subject up, and then Paul Caro immediately attacked me on the subject during the coffee break. I don't think there is any question about the efficacy of knowledge of results, but I think there are some serious questions about what kind of information should be presented. For example, should you present correct or incorrect information? Guy Matheny did an experiment in which he presented incorrect information and got beautiful results. So it is not necessarily best to let the student know exactly how well he is doing. At least, it is a good experimental question. Should knowledge of results be continuous, periodic, or aperiodic? That is a practical question when you are setting up an adaptive training program.

Should knowledge of results be presented directly, by inference, or in some other way? There are many ways in which you can imply what the performance is to the student. I don't know the answers, but I suggest that no one else does either. So it is a legitimate subject for experimental investigation and one of some practical importance in designing a training program.

Now, one other line of discussion, and I will be through. I mentioned the desirability of establishing operational criteria for the evaluation of any training system or program. I realize this is hard to come by. Assuming that we have to establish some training standards that we hope are related to operational criteria without any proof thereof, I still believe that the proper evaluation of the training program is in terms of the transfer to operational tasks.

Now, there are some intermediate measures that could be taken along the way, and these should not be overlooked. One obvious measure is the speed of learning to reach the performance level that is specified by the training program. Another obvious measure, and a corollary of the first, is the terminal level of the performance achieved in a given period of training time. Through adaptive training, perhaps we can achieve a higher level of performance in x number of hours than we can through nonadaptive training. We should know whether or not that is so.

Another criterion of great importance is retention over periods of disuse of a skill or knowledge. The question of overlearning comes in here, but it is possible for adaptive training to result in either better or worse retention. Perhaps it will result in achieving a given level of performance more quickly but will not yield the same degree of retention.

Finally, all of these concepts should be evaluated in terms of their transfer to as nearly operational criteria as we can devise. These are good experimental questions, and nobody here knows the answers.

MR. FLEXMAN: It would be nice if we could form an organization to systematically explore some of these critical problems. Unfortunately, we cannot expect the Naval Training Device Center and the Air Training Command to provide all of the support for the necessary research in this field. I think this is a relevant question for the Department of Health, Education, and Welfare, the Department of Transportation, and many other agencies. Perhaps the publication of the views expressed here will stimulate greater research interest in this area.

Research is needed not only for training but for the extension of the adaptive concept to completely unexplored applications in many other areas—as a prediction device, as a new approach to personnel selection, as a means for operating on-the-job equipment, as a means for controlling the priority of information presented to a task-loaded operator, as an extension of programmed learning, and so forth. There is a vast field for our efforts, and the only way we will make progress is by the right presentation of the method and its potential to the profession. Now, by the profession, I mean the behavioral scientists, the design engineers, the educators, and the users.

". . . ONLY TWO TRULY UNIQUE IDEAS. . . ."

Dr. McGrath: Adaptive training could become recognized as just a special branch of flight simulation. That is a distinct possibility, and it must be avoided. The only way these broader applications can be obtained is by broadcasting the principles of adaptive training that can be generalized. These are very difficult to identify, as we have discovered during this conference.

Nevertheless, adaptive training research and development efforts have produced several benefits for general education and training. One benefit is the emphasis on near-continuous performance measurement during the training process. Another benefit is a more searching analysis of the variables that can influence student performance or task difficulty, because the adaptive training paradigm requires that all of these variables be considered in the choice of the adaptive variables and the development of the adaptive logic. I think further work on adaptive training will result in a quantitative mapping of the human learning process because one of the tangible payoffs of adaptive training is a detailed record of the development of skills and knowledge.

However, there are not many original ideas in adaptive training when you come right down to it. Many of the concepts discussed in this conference are basic to the traditional approaches to teaching, and as nearly as I can tell there are only two truly unique ideas attributable to the adaptive approach. One is the formalizing, or even computerizing, of the link between student performance and the presented task. Traditional educational methods have provided such a link, but in adaptive training it is far more structured and requires a deliberate logic. This logic is central to the adaptive training paradigm, whereas it seems to be incidental to most educational theory. The other unique idea is Chuck Kelley's method of maintaining performance constant. I think this characteristic sets adaptive training aside from all other kinds of training. I am not convinced of the benefits, but it is nevertheless unique to adaptive methods of training.

Now, if we really want to encourage others to think about how to apply adaptive methods, we must identify all such generalizable principles that characterize adaptive training or are used in designing an adaptive system. These should be principles that can be applied to training problems that have nothing to do with aircraft control or tracking performance. For example, have you ever considered how adaptive training methods might be applied to the training of athletes, such as pro football players? Then, consider the training of typists, letter sorters, and other business machine operators. Think about the training of surveillance crews in target detection and recognition. Think about the uses of adaptive methods in teaching languages or in studying risk-taking and gaming behavior. Adaptive techniques might even be useful in the broad field of social behavior. For instance, a law enforcement officer, these days, must be trained to make flexible responses to very subtle cues. He must learn to perform under stress, much like a pilot. He may also need desensitization training so that he does not overreact to provocation. Possibly such training problems would yield to imaginative applications of adaptive training principles.

But it is not for us, who have no expertise in these areas, to say how adaptive training techniques can be applied. That can be done only by those who thoroughly understand the operational problems and the training environment in these other areas. So it is necessary to get the information on adaptive training technology to experts in these other fields, if we are to extend its applications. Those experts are the key people who will have to apply the adaptive training method, but they can't do it unless they have a good understanding of the generalizable principles of adaptive training.

DR. CARO: I think Jim McGrath made a very significant statement. He wound up by saying we need to convey techniques of training to people with subject-matter knowledge. Adaptive training is nothing but a technique for presenting certain types of information to a student. It is going to have no impact unless it is accompanied by a solid understanding of the task and a critical look at the curriculum. I am reminded of a comment I read recently about closed-circuit television. CCTV has had a large impact on education in the past decade or so but only because it has brought about a very critical look at the content and organization of the curriculum. I think if we omit that step in our use of adaptive training, adaptive training won't be around in a few years.

". . . RUNNING INTO A *ZEITGEIST*. . . ."

Mr. Flexman: Thank you, Paul. I'm going to continue on around the room to get everybody's final statement concerning their views on the subject of this meeting.

Mr. Norman: I think adaptive training may be running into a *zeitgeist* against depersonalized instruction; this exists now particularly in the university. There may be a great deal of opposition to a fully automated SFTS in which the student flies and flies and relates to nothing more personal than a stack of IBM cards.

Dr. Caro: I wish we had the capability of causing that to be a real issue in the foreseeable future. We are not going to have a fully automated SFTS. I wish we could because there are many other questions that could be addressed then.

Mr. Flexman: The PLATO System, as some of you will see tomorrow, has had some spectacular successes in the training of nurses. Yet, there has always been a very small minority of student nurses who react vigorously against the dehumanization of that training process. However, the increased learning rate and the fantastic increase in retention obtained suggests that it is well worthwhile. Perhaps we need to present these training processes differently.

Dr. Caro: Charley Slack did some research a few years back in which he conducted psychotherapy without a therapist. The patient entered a box, then talked into a microphone, and his verbal behavior was reinforced. The results were fascinating. They suggested that depersonalization is not necessarily a consequence of getting the therapist out of the system. It should be possible to make instruction very personal by having it conducted by a computer rather than a human instructor. The instructional logic is the key.

DR. KELLEY: There are two issues in the dehumanization question. First, the numerous people involved in the very powerful union of educators are threatened because they may be replaced. This consideration has certainly limited the use of instructional television, for example. Everyone would agree that it is far better to have a good lecturer, who really knows his subject, lecture to you through the medium of television than it is to sit in a hall and listen to a second-rater. Nevertheless, the latter method still prevails by and large in university systems. Second, there is the real question of what qualities in instruction are, in fact, lost when the human instructor is removed. Unions aside, some things are changed when you computerize, and I think we should be very sensitive to these changes. The first issue is not one we should concern ourselves with as scientists, but the second one certainly is.

DR. MATHENY: That reminds me of the response of a British administrator to the invention of a young scientist. He said, "The brilliance of the ingenuity of your invention almost blinds me to its utter uselessness."

DR. MCGRATH: In the tradition of science, we may not know what we're doing, but we're doing it very carefully.

MR. FLEXMAN: Or the airline captain who addressed his passengers over the speaker system, "Our navigation equipment has gone out on this trip, and we're now completely lost, but we're making a wonderful groundspeed."

DR. MATHENY: It is extremely important to remember that we are dealing with the complete individual, even though we have been restricting ourselves to thinking of a man as the pilot in an aircraft. The research that has been done in this area has been mostly unimodal, primarily with visual inputs. Well, a lot of other inputs are involved, and this multimodal problem may give you fits. For example, if the characteristics of the motion platform are not appropriate, some very weird things may happen as a function of increased turbulence.

Regarding the broader question of how we get anyplace in this business, we have to face the facts and look to the United States Government for research funding. There is so much to be done. At this point, we cannot show anyone that we can save them money, or even propose to, as an argument for getting money for research and development. The SFTS does offer an opportunity to get some information. Exploit it to the utmost. Now, I pass to Milt Wood.

". . . IN A SOMEWHAT CONFUSED STATE. . . ."

Mr. Wood: There are several items on my list of issues that I wish to mention in summary.

First, I certainly agree with Jim McGrath that general principles which define adaptive training need to be identified and stated. Perhaps it is time to start forming a theoretical basis for adaptive training. We could look to modern learning principles and to the interesting area of information processing to see to what extent they support, constrain, or direct the concepts of adaptive training. Also, I have been looking at adaptive training specifically in terms of perceptual-motor skills; on the basis of our discussion here, I now feel that the concept should be more encompassing.

Second, we need to look into the potential of an adaptive approach in providing us a rationale for when to increase task complexity. We may be able, for example, to define an index of available attention that can be used for increasing task complexity.

Third, there is the question of the subject pool. In talking with Don Norman and trying to decide why the results of our two efforts were different, we agreed that the subject pool may have played a part—he was dealing with trained pilots, and I was dealing with naive students. Perhaps different adaptive models will have to be employed with different subject pools.

Fourth, there are basic questions concerning overall adaptive design. My study pointed up some of the issues involved; so did Chuck Kelley when he emphasized the need for models that handle learning as the gradual process that it is. I suggest that we look at different ways for using the best of several worlds, perhaps the use of fixed levels of performance, adaptively modified. We have a long way to go, and we are just starting, but I guess that is quite obvious.

Mr. Flexman: Thank you, Milt. John, I know that you have not been closely associated with this area but have been exposed to it.

Dr. Lauber: I feel much relieved. Yesterday morning when we started, I was convinced that I was the only person who was confused about some of the issues, definitions, and methods of this area. Now, as this conference nears the end, I realize that I have company and that adaptive training has not advanced to the point where there are no issues to be resolved. Hopefully, at the "Second Annual Workshop on Adaptive Training," some of these questions will have been answered.

Mr. Flexman: Mel?

DR. MONTEMERLO: I, like John Lauber, came here in a somewhat confused state. However, I now see some specific questions for research, one of the foremost being the question of where adaptive training is more effective. I would propose a study along the lines of what Cronbach calls an aptitude-treatment-interaction study: high-aptitude and low-aptitude pilots would perform high-, medium-, and low-difficulty tasks with adaptive and nonadaptive training. This would be one way of determining where adaptive training would have its highest effectiveness. Earlier, when I mentioned that we should do some studies on adaptive training versus the nonadaptive training provided in existing systems, you indicated that we were not ready for that type of thing. What do you think we are ready for? Do we need implementation of adaptive training, more realistic training tasks, or more research?

MR. FLEXMAN: Well, had I not been involved in the SFTS program, I would have said that we are ready now. But, you know, it is one thing to talk about doing something and another to do it.

DR. CARO: I think you would have to say that, if we started now on a second SFTS, it would be much better than the first. We have learned a good deal.

MR. FLEXMAN: Yes, we have.

DR. CARO: But, I hope that what we have learned is not that we should not have built the first one.

MR. FLEXMAN: I don't think so at all. Bill, what do you have to say in summary?

". . . RIGHT DOWN THE GROOVE. . . ."

MR. STEWART: I agree with everybody. Instead of summarizing, I wish to throw in one final detail which specifically addresses the SFTS work. A possible adaptive variable, and an interesting one to me, might be the sensitivity of instrument display. Ask the man to fly to very, very close tolerances, and give him a variable instrument so that, as he gets better, the instrument becomes more sensitive. The level of sensitivity of the instrument could be displayed to him. At first, a tolerance of 5° off course might be established, and then gradually, as his in-tolerance time increases, the instruments would become more sensitive, providing him with more precise flight deviation information. Pretty soon he is flying right down the groove. This approach might be most applicable for improving performance in very stable conditions.

Mr. Flexman: That is certainly an interesting avenue to pursue. John, may we hear from you next?

Mr. Sinacori: I wrote down a list of possible contributions to the adaptive scheme that the control analyst may make. I will go through it briefly. In general, I think the adaptive concept has potential, and I am most anxious to see experimental results that verify its value. I see the control analyst continuing to play a supporting role in this area; this role may never change. Remember, engineers tend to see a human being as an aggregate of wires, logic elements, sensors, computers, and so forth—a view that may be quite useful only in this role.

Here are some of the considerations. First, the identification of possible adaptive variables and the determination of their sensitivity are certainly within the realm of the pilot modeling theory. Many of us did not believe in the pilot modeling technique, when it was first advanced by McRuer, and tended to pooh-pooh it. He said that, even if you don't accept it, you should at least accept the chance it offers for structuring experiments so that you can prove or disprove it yourself. It may have some similar value on the adaptive scene as well. It may help us to better understand some of the principles and to be able to make predictions of some sort.

The second consideration on my list is the identification of specific operator techniques for task performance, with emphasis on schemes for shaping and eliciting appropriate behavior. An example is the human interface with motion that I have studied. The last one is a better definition of time structuring and a determination of what is involved in making this aspect of training more effective. In other words, training has to devote much time to each particular task, and obviously there are many tasks typically involved. Henry Jex, using an eye-tracking device, was able to measure instrument scan rates, dwell times, and so on for a 707 ILS approach task. Using these data he was able to verify, with a good deal of accuracy, a number of contentions about when exactly a pilot will be saturated and how much time he has available to him for various tasks. This type of work might provide us some insights into the time-structuring question.

". . . THE INSTRUCTOR'S CONTRIBUTION. . . ."

Mr. Flexman: Thank you, John. Patty, I think it's about your turn.

Miss Knoop: Well, for one last and final time, I am of the feeling that once we have solved the performance-measurement problem, we will know much more about what needs to be done in adaptive training. Everybody probably feels pretty much the same about that. Another point I want to make was stimulated by all the talk of replacing the instructor in flight simulator training. I think it should be in the record that, while we can do a great deal through automation to improve the effectiveness of training, there will always be a place for an instructor and his observations in a training system. So, rather than trying to replace the instructor, we should think more in terms of aiding him to the point where he has more freedom to exercise the unique skills that he brings into the training system and cannot now fully utilize in the simulation environment.

MR. FLEXMAN: Admittedly, many of us have the objective of replacing the instructor; however, I think that we also realize that this may be an unrealistic objective and that the instructor, with his managerial and diagnostic skills, should be in the system. With good adaptive systems as tools, maybe he can become a better manager and diagnostician by using the information from the computer and prescribing remedial training situations. Perhaps he will become more sensitive to motivation. Even if our efforts were to completely replace him, this would not necessarily be bad. The impetus to take over more of his functions and see if we can do them better makes us work harder and assess our results more rigorously. On the other hand, I think there should be equal efforts going on to enhance the instructor's contribution to the learning process, and maybe these two efforts can be mutually supportive.

MISS KNOOP: I'm thinking more of the importance of the instructor in performance evaluation than in training operations.

MR. FLEXMAN: Yes, I doubt if a computer is going to be able to assess that this guy almost grabbed the wrong thing, that he is very nervous and unsure of himself, and that he is right on the precipice of pure panic behavior. Paul, do you have any summary statements?

DR. CARO: I agree with what each of you says about research needs individually, but putting it all together, I hear too much emphasis on the techniques of adaptive training. Adaptive training is a technique that will allow us to manipulate the training environment under machine control. But we need to be concerned about the definition of that environment and the curriculum, something more than just the mechanization of the adaptive model. The emphasis should be on better system analysis, curriculum engineering, and performance assessment. We need to pay more attention to what and how well we are teaching than to the particular technique of teaching it. The tail must not wag the dog. We must guard against any application where adaptive training could become an end rather than a means to more efficient training.

". . . CLOSING UP SHOP. . . ."

Mr. Flexman: Chuck, you have had more experience and background in the area than all of us put together. I think that it would be very appropriate for you to wind this thing up.

Dr. Kelley: I will attempt to keep my remarks pointed, because many of the things I would say in summary have already been said by others. I would like to amplify the point that Patricia Knoop made. The problem of performance measurement in the context of an adaptive system must be brought to the forefront. In the SFTS case, it is very easy to forget that you are taking three adaptive variables that you don't know very much about and manipulating them on the basis of performance tolerances that are quite arbitrary and that have relationships that are unspecified.

Performance measurement is the key to these considerations. Now, if we—the people who know as much about this problem as anybody else—tend to lose sight of the problem of performance measurement, what is going to happen when adaptive systems are used as routine training instruments? So we have to keep coming back to the same point: How do you measure performance? What exactly are you measuring? Is the measurement reliable, sensitive, and, above all, valid? You can't ask those questions too often, and if you ever forget them, they will come back and kick you in the head.

When you talk seriously about adaptive training or try to mechanize it, then you have to look at training in a way that you may never have looked at it before. No training concept can be taken for granted; all the assumptions and techniques must be examined in detail. I believe the more significant the skill or knowledge we seek to impart, the more difficult it is to even quantify what it is we are really trying to train, much less make it adaptive.

In many areas, adaptive training cannot do much beyond helping people get the basic initial skills. Adaptive methods can be used to train music students, but what it takes to make a truly fine musician, for example, is something that we cannot reasonably expect a computer to recognize. When skills that are very difficult to acquire and that really matter in the world are the object of training, we must never lose sight of the many complex criteria that are involved in properly assessing those skills. We must never let the simplification that is involved in adaptive training distort in our minds the complexity of the real task that we are dealing with.

I have learned through some hard experience how little we know about ultimate criteria of performance in difficult real-world tasks. Yet, this is the direction we have to go if we are really going to make significant progress. Just a little bit of progress in that direction, just one new way of measuring what a pilot really does in flying an airliner, will prove immensely valuable in structuring the training program; we can make our training program more effective because we know something important about flying the aircraft that we never knew before. We have to find, in a sense, new performance measures in the real world before training benefits can fall out.

MR. FLEXMAN: Thank you, Chuck. We are closing up shop; does anyone have a final word?

MR. HASLER: Does anyone need a ride to Vandalia?

Mr. Hasler

DR. CARO: I would like to say something. You have done a marvelous job, Ralph, and we all thank you for putting on this conference. It has been a valuable experience.

MR. FLEXMAN: It has been a surprisingly compatible group. I hope we can do it again. Whether our time was well spent will be proven in our new perspective of the problem area and in our future reactions to that perspective.

SELECTED REFERENCES

BAHRICK, H. P., NOBEL, M., and FITTS, P. M. Extra-task performance as a measure of learning a primary task. *Journal of Experimental Psychology*, 1954, **48**, 298–302.

BIRMINGHAM, H. P. The instantaneous measurement of human bandwidth. Naval Research Laboratory, Washington, D. C. Paper presented at the 8th Annual Conference on Manual Control, Dunlap and Associates, Inc., Stamford, Conn., May 1959.

BIRMINGHAM, H. P., CHERNIKOFF, R., and ZIEGLER, P. N. The design and use of "equalization" teaching machines. Naval Research Laboratory, Washington, D. C. Paper presented at the IRE International Congress on Human Factors in Electronics, Long Beach, Calif., May 1962.

BRIGGS, G. E. On the scheduling of training conditions for the acquisition and transfer of perceptual-motor skills. Port Washington, N. Y.: Naval Training Device Center, Contract N61339-836, Technical Report NAVTRADEVCEN 836-1, December 1961.

CARO, P. W., Jr. Adaptive training—an application to flight simulation. *Human Factors*, 1969, 11, 569–575.

CHERNIKOFF, R. Human equalization: Concept, machine, and measurement. In *Complex vehicular controls: Proceedings of Office of Naval Research symposium, Farnborough, England.* Long Beach, Calif.: Douglas Aircraft Co., Inc., May 1962.

CLAUSEN, J. T., CURTIN, J. G., and EGLER, J. F. Concept formulation report, synthetic flight training system Device 2B24. Orlando, Fla.: Naval Training Device Center, Contract N61339-68-C-0108, Technical Report NAVTRADEVCEN-68-C-0108-1, July 1968. (AD 673982)

CONNELLY, E., SCHULER, A., and KNOOP, PATRICIA A. Study of adaptive mathematical models for deriving automated pilot performance measurement techniques. Wright-Patterson Air Force Base, Ohio: Air Force Human Resources Laboratory, Contract F33615-68-C-1278, Technical Report AFHRL-TR-69-7, 1969.

FACONTI, V., MORTIMER, C. P. L., and SIMPSON, D. W. Automated instruction and performance monitoring in flight simulator training. Wright-Patterson Air Force Base, Ohio: Air Force Human Resources Laboratory, Contract F33615-69-C-1159, Technical Report AFHRL-TR-69-29, February 1970.

GAINES, B. R. Teaching machines for perceptual-motor skills. In D. Unwin and J. Leedham (Eds.) *Aspects of educational technology: The proceedings of the programmed learning conference held at Loughborough, 15–18 April 1966.* London, England: Methuen, 1967, 337–358.

HUDSON, E. M. An adaptive tracking simulator. Otis Elevator Company, Brooklyn, N. Y. Paper presented at the IRE First International Congress on Human Factors in Electronics, Long Beach, Calif., May 1962.

HUDSON, E. M. Adaptive training and nonverbal behavior. Port Washington, N. Y.: Naval Training Device Center, Contract N61339-1395, Technical Report 1395-1, July 1964.

HUDSON, E. M. Adaptive training device. U. S. Patent 3311995, 4 April 1967. In U. S. Commerce Department Patent Office. *Official Gazette*. Washington, D. C.: Government Printing Office, 1967, p. 41.

HUDSON, E. M. Adaptive techniques in multiparameter problems. *Human Factors,* 1969, 11, 561–568.

JAMESON, W. P., WALSH, J. M., FLEXMAN, R. E., and COHEN, E. Synthetic flight training system (SFTS): Concept formulation report. Orlando, Fla.: Naval Training Device Center, Contract N61339-68-C-0106, Technical Report NAVTRADEVCEN-68-C-0106-1, April 1968.

KELLEY, C. R. Self-adjusting vehicle simulators. In C. R. Kelley (Ed.) *The predictor instrument: Final report and summary of project activities during 1961.* Arlington, Va.: Office of Naval Research, Contract Nonr 2822 (00), Technical Report, January 1962. Also in C. R. Kelley (Ed.) *Adaptive simulation.* Arlington, Va.: Office of Naval Research, Contract Nonr 4986(00), Final Report, August 1966. (AD-637658)

KELLEY, C. R. Testing the effects of extended space missions on piloting skill. In E. Burgess (Ed.) *AAS science and technology series.* Vol. 5. P. Horowitz (Ed.) *Physiological and performance determinants of manned space system design.* Baltimore, Md.: American Astronautical Society, 1965, 167–186.

KELLEY, C. R. Further research with adaptive tasks. Arlington, Va.: Office of Naval Research, Contract Nonr-4986(00), Final Report, August 1967. (AD-657-343)

KELLEY, C. R. Adaptive and automated research techniques from engineering psychology. *American Psychologist,* 1969, 24, 293–297.

KELLEY, C. R. What is adaptive training? *Human Factors,* 1969, 11, 547–556.

KELLEY, C. R. and KELLEY, ERICA J. A manual for adaptive techniques. Arlington, Va.: Office of Naval Research, Contract Nonr-4986(00), Final Report, August 1970. (AD-711-985)

KELLEY, C. R. and PROSIN, D. J. Adaptive performance measurement. Arlington, Va.: Office of Naval Research, Contract Nonr-4986(00), Final Report, August 1968. (AD-677-049)

KELLEY, C. R. and PROSIN, D. J. Adaptive techniques in measuring complex performance. Arlington, Va.: Office of Naval Research, Contract Nonr-4986(00), Final Report, September 1969. (AD-694-509)

KELLEY, C. R. and WARGO, M. J. Cross-adaptive operator loading tasks. *Human Factors,* 1967, 9, 395–404.

KELLEY, C. R. and WARGO, M. J. Adaptive techniques for synthetic flight training systems. Orlando, Fla.: Naval Training Device Center, Contract N61339-68-C-0136, Technical Report NAVTRADEVCEN 68-C-0136-1, October 1968. Also in C. R. Kelley (Ed.) *Adaptive simulation*. Arlington, Va.: Office of Naval Research, Contract Nonr 4986(00), Final Report, August 1966. (AD-637-658)

Knoop, Patricia A. Programming techniques for the automatic monitoring of human performance. Wright-Patterson Air Force Base, Ohio: Aerospace Medical Research Laboratory, Project 6114, Task 611412, Technical Report AMRL-TR-66-16, April 1966.

Knoop, Patricia A. Automated aids for flight simulator instructors. *USAF Instructor's Journal,* 1967, 4(4), 61–64.

Knoop, Patricia A. Development and evaluation of a digital computer program for automatic human performance monitoring in flight simulator training. Wright-Patterson Air Force Base, Ohio: Aerospace Medical Research Laboratories, Project 6114, Task 611412, Technical Report AMRL-TR-67-97, August 1968.

Kopstein, F. F. and Seidel, R. J. The computer as adaptive instructional decision maker. Alexandria, Va.: Human Resources Research Organization, Contract DAHC 19-70-C-0012, HUMRRO Professional Paper-1-70, January 1970. (AD-703-597)

Lewis, B. N. and Pask, G. The theory and practice of adaptive teaching systems. In R. Glaser (Ed.) *Teaching machines and programed learning*. Vol. II. *Data and directions*. Washington, D. C.: National Education Association, 1965, 213–266.

Lowes, A. L., Ellis, N. C., Norman, D. A., and Matheny, W. G. Improving piloting skills in turbulent air using a self-adaptive technique for a digital operational flight trainer. Orlando, Fla.: Naval Training Device Center, Contract N61339-67-C-0034, Technical Report NAVTRADEVCEN 67-C-0034-2, August 1968. (AD-675-805)

Matheny, W. G. The effective time constant—a new technique for adaptive training. *Human Factors,* 1969, **11**, 557–560.

Matheny, W. G. and Norman, D. A. The effective time constant in tracking behavior. Orlando, Fla.: Naval Training Device Center, Contract N61339-67-C-0034, Technical Report NAVTRADEVCEN 67-C-0034-3, August 1968.

Mirabella, A. and Lamb, J. C. Computer based adaptive training applied to symbolic displays. *Perceptual and Motor Skills,* 1966, 23, 647–661.

Pask, G. Adaptive teaching with adaptive machines. In A. A. Lumsdaine and R. Glaser (Eds.) *Teaching machines and programmed learning: A source book*. Washington, D. C.: National Education Association, 1960, 349–366.

Pask, G. Electronic keyboard teaching machines. In A. A. Lumsdaine and R. Glaser (Eds.) *Teaching machines and programmed learning: A source book*. Washington, D. C.: National Education Association, 1960, 336–348.

Pask, G. The cybernetics of evolutionary processes and of self organizing systems. System Research Ltd., Richmond, Surrey, England. Paper delivered at the 3rd Congress of the International Association of Cybernetics, Namur, Belgium, September 1961.

PASK, G. The logic and behavior of self organizing systems as illustrated by the interaction between men and adaptive machines. System Research Ltd., Richmond, Surrey, England. Paper delivered at the International Symposium on Information Theory, Brussels, Belgium, 1962.

PASK, G. and LEWIS, B. N. An adaptive automaton for teaching small groups. *Perceptual and Motor Skills*, 1962, 14, 183–188.

PASK, G. and VON FOERSTER, H. A predictive model for self-organizing systems. *Cybernetica*, 1961, 4, 20–55.

REGAN, J. J. Preface. *Human Factors*, 1969, 11, 545.

STEWART, W. A. Rand symposium on pilot training and the pilot career: Recollections of the chairman. Santa Monica, Calif.: Rand Corporation, Report RM-6282-1-PR, March 1970.

STOLUROW, L. M. Some factors in the design of systems for computer-assisted instruction. Arlington, Va.: Office of Naval Research, Contract N00014-67-A-0298-0003, Report 7, May 1968.

WALSH, J. M. and FLEXMAN, R. E. Preliminary adaptive training design report for synthetic flight training system Device 2B24. Orlando, Fla.: Naval Training Device Center, Contract N61339-69-C-0200, Technical Report 69-C-0200-15, April 1970.

WOOD, M. E. Continuously adaptive versus discrete changes of task difficulty in the training of a complex perceptual-motor skill. In *Proceedings of the 77th Annual Convention of the American Psychological Association*, 1969, 4, 757–758.

YOUNG, F. E. and HALL, E. R. Synthetic flight trainer system (2B24): Concept formulation report. Orlando, Fla.: Naval Training Device Center, Contract N61339-68-C-0107, Technical Report NAVTRADEVCEN-68-C-0107-1, July 1968.