Training High Performance Skills: Fallacies and Guidelines

Walter Schneider

Department of Psychology
University of Illinois
Champaign, IL 61820

Personnel & Training Research Programs
Office of Naval Research
Arlington, VA 22217

Approved for public release; distribution unlimited

Training, skilled performance, attention, workload, air traffic control, learning, information processing.

Performance skill acquisition are reviewed. These include long acquisition periods, heterogeneity of component learning, development of inappropriate strategies, and training of time-sharing skills. A tentative set of working guidelines for the acquisition of high performance skills is described. Original-supplied keywords include: (to 14736)
The following is a list of the published reports and abstracts supported by this contract. The basic research provides an interpretation of performance changes during skill acquisition. The enclosed report "Training High Performance Skills: Fallacies and Guidelines" is a synthesis of the basic results for application to applied training situations.

A List of Published Reports


A model for the development of automatic processing is briefly described. The model is a quasi-neural model in which information processing is done through the transmission of vectors between visual lexical, semantic, and motor processing units. Controlled processing involves gating of the output power of vectors to perform matches and to release response vectors. As subjects practice consistent tasks, associative learning enables an input vector to evoke an output vector and priority learning determines the power with which a vector is transmitted. Automatic processing involves a cascade of vector transmissions in which the output power of each transmission is determined by the priority learning. The transition from controlled to automatic processing takes place in four phases. Empirical illustrations of the transition are described.


This report discusses prediction of individual differences in task performance during and subsequent to task practice. Previous literature indicates that prepractice prediction of postpractice performance declines rapidly as time-on-task increases (for both simple and relatively complex tasks). Based on these effects, traditional conceptions equating general intelligence with learning ability are inconsistent with performance data. The present approach reviews practice effects from an information processing perspective. The distinction between two major types of practice effects is outlined and discussed with respect to the automatic and controlled processing framework. The thrust of the discussion of individual differences and practice is predicated on a theoretical organization which draws together theories of the structure of cognitive/intellectual abilities with aspects of resource theory and elements of automatic and controlled processing. A unified theory of practice is presented. The theory relates ability and performance individual differences to task component consistency characteristics. The supporting data of an experimental study of individual differences in initial, intermediate and final practiced performance stages are reported. Proposals regarding new assessment procedures and recommendations for the restructuring of both selection and training methods are described.


Briefly reviews divided attention, focused attention, attentional capacity, and effort from the perspective of automatic processing. Factors which affect automatization are discussed. The functions and limitations of automatic and controlled processing are briefly described.


The Broadbent Maltese Crosse model of attention is criticized. The potential for interference without the existence of a single central channel is discussed. The ability of practice reducing interference effects is presented as a serious problem for the Broadbent model.


The theory of automatic and controlled processing outlined in Schneider and Shiffrin (1977) and in Shiffrin and Schneider (1977) is defended in the present note. We argue that the criticism of Ryan (1983) range from irrelevant to incorrect, based on a brief review of data from the 1977 articles and on some more recent publications. The evidence Ryan discusses comes from the presemorized-list paradigm, a paradigm that undoubtedly involves automatic and controlled processes but probably not automatic detection and controlled search. We argue that a variety of mechanisms consistent with our general theory, some automatic and some controlled, could be operating in the presemorized-list paradigm and can explain the observed results.


The relationships between long-term memory (LTM) modification, attentional allocation, and type of processing are examined. Automatic/controlled processing theory (Schneider & Shiffrin, 1977) predicts that the nature and amount of controlled processing determines LTM storage and that stimuli can be automatically processed with no lasting long term memory (LTM) effect. Subjects performed the following: (a) an intentional learning, (b) a semantic categorization, (c) a graphic categorization, (d) a distracting digit-search while intentionally learning words, and (e) a distracting digit-search while ignoring words. Frequency judgments were more accurate in the semantic and intentional conditions than the graphic condition. Frequency judgments in the digit-search conditions were near chance. Experiment 2 extensively trained subjects to develop automatic categorization. Automatic categorization produced no frequency learning and little recognition. These results also disconfirm the Hasher and Zacks (1979) "automatic encoding" proposal regarding the nature of processing.

Experiments examined practice and transfer effects in consistently mapped (CM) and variably mapped (VM) semantic search. Experiment 1 examined improvements in reaction time in detecting words from a category as a function of exemplars (4-12) in the category. All CM conditions showed improvement, but there was no significant effect of the number of exemplars transferred to untrained members of the category. Experiment 1b examined the extent to which training on a subset of exemplars transferred to untrained members of the category. Results showed substantial positive transfer (60%-92%) to untrained exemplars from the trained category. The transfer was better if there were more exemplars on the training set. Experiment 2a showed practice reduced response sensitivity in CM category search but did not benefit VM category search. Experiment 2b showed that under high workload, untrained exemplars of the trained CM category were detected when first presented to exhibit substantial positive transfer (70%). We conclude that many of the practice effects observed for CM category search take place at either the category level or the category feature level. We suggest that practice results in context activation of the category node or category feature level. This context activation hypothesis is evaluated with respect to major phenomena relating to automatic and controlled processing.


Experimetal and statistical methods for examining individual differences in dual-task performance and time-sharing ability are reviewed and criticized. Previous data and analytic procedures are generally inadequate to evaluate a time-sharing ability. Errors resulting from unsophisticated use of correlational and factor analytic procedures are described. Four previous studies that concern time-sharing are considered in detail. The nature of task selection, scoring methods, and control of practice and reliability issues are discussed. Based on a reanalysis of available data, a time-sharing ability is not rejected. Simulation, incorporation of theory in planning models, and crucial tests of the hypothesis are proposed as methods for assessing the time-sharing ability.


The current research examined whether the total processing, from stimulus to response, must be consistent for automatic processing to improve task performance. Consistent versus inconsistent attending and responding (i.e., reponse translation) were factorially combined. The results showed that the consistency of attending produced substantial practice effects. However, consistency of the motor response translation component did not affect the asymptotic performance level. The results indicate that automatic processing develops for consistent components of a task even when the entire task is not consistent.


This research examines how the major phenomena of visual search for single characters generalize to word search when the target and distractor sets had a varied mapping (VM) across trials. Reaction time was a linear function of the number of comparisons with a positive slope of 48 msec per word, 92 msec per category. The VM reaction time data indicated a self-terminating comparison process, and there was little or no improvement with practice. Experiment 2 examined search with a consistent mapping (CM) between targets and distractors. Category search slopes dropped to 2 msec, but the function was still linear. Experiment 3 examined category detection carried out concurrently with serial recall of digits, allowing assessment of search performance under high workload. High workload caused a severe performance reduction in VM category search, and this decrement did not decrease with practice. High workload reduced initial performance in CM category search, but this decrement was eliminated with practice. The present category search results are similar to previous letter search results. Four principles of search are discussed in the context of a theory of automatic/control processing.


Current attentional research and theory are related to the development of skilled performance. Emphasis is given to how performance changes with practice. Dual process attention theory is reviewed examining the distinctions between automatic and controlled processing. The changing interactions between automatic and controlled processing in the development of skill are discussed. It is proposed that consistent practice produces automatic productions which perform consistent transformations in a heterarchical system. Automatic productions are proposed to be modular; show high transfer; become resource free; not be under direct control; and be fast, accurate, and coordinated. Controlled processing is assumed to develop automatic processing, maintain strategy and time varying information, and perform problem-solving activities. Perceptual data, some motor data, and several motor performance examples are presented to illustrate automatic/controlled processing effects. The relationship to current theories of motor skill are discussed. New research paradigms suggested by the current approach are discussed.


The effects of practice on processing speed and workload are illustrated. The need for theoretical predictions of practice effects is commented on. A quasi-neural
simulation model is described. The model quantitatively defines automatic and controlled processing in terms of activation levels and connection strengths between independent processing units. The model utilizes neurophysiological and communication theory concepts to illustrate why attention must be limited and how parallel processing can develop with consistent practice. Simulation predictions of a variety of attentional phenomena are described.


This talk presents four fallacies implicitly assumed in many training programs. These fallacies are: practice makes perfect; most training should occur in the final task; train for high accuracy performance; and skill learning is intrinsically enjoyable. Examples of these fallacies in training reading and air-traffic control are described. A theoretical perspective on automatic and controlled processing for skill acquisition is discussed briefly. The theoretical perspective suggests eleven guidelines for microprocessor-based skill training. The guidelines are described for computer-based training of air-traffic control skills.


A quasi-neural simulation model is described. The model quantitatively defines automatic and controlled processing in terms of activation levels and connection strengths between processing units. The model uses neurophysiological and communication theory concepts to illustrate: (1) why attention must be limited; (2) how parallel processing can develop with consistent practice; and (3) why automatic processing is limited by factors different from those affecting controlled processing. The fit of simulations to classic attention and search data is presented.


Previous research has shown substantial improvements in detection performance when subjects consistently detect a subset of stimuli. In contrast to conditions in which stimuli appear as both targets and distractors, there is little performance improvement with practice. The present experiments examine how varying degrees of consistency determine the improvement of detection accuracy with extended practice. The degree of consistency was varied by manipulating the frequency with which a letter was a distractor while holding the number of occurrences as a target constant. The experiments utilized a multiple-frame target-detection search paradigm in which subjects were to detect single-letter targets in a series of rapidly presented letters on four channels. Experiments showed that detection performance improvement with practice was a monotonic function of the function of the degree of consistency, decreasing to zero as the target-to-distractor ratio increased from 10:00 to 10:20. As consistency decreased, detection performance asymptoted earlier and at a lower level. A dual-task as a secondary task. Results showed that the previous target-to-distractor consistency had a marked effect on resource sensitivity of the detection task. The general issues of consistency in the development of skilled performance and in the development of automatic processing are discussed.


Guidelines for microprocessor-based skill trainers are presented. A training program for air traffic control (ATC) of rendezvous for in-flight refueling is described. The program seeks to optimize practice for developing automatic component skills. The program sequences the trainee through 10 stages to develop spatial skills for ATC. The resulting training program can develop fast, accurate, and reliable performance on the individual components with only a few hours of training per component. The proposed approach is contrasted with current training methods. The general applicability of the guidelines to microprocessor-based skill trainers is described.


The study tested guidelines for the use of microprocessors in training spatial skills for air traffic control. The central issue was the use of time-compressed simulation to aid the development of skill in identifying turn points and rollout headings for aircraft. Two groups of subjects were used. One group trained with a real-time simulation of the task, while the second group trained time-compressed version of the task running about 20 times as fast as real-time trials. Both groups were then tested in real-time trials. The results indicate that time compression can be a useful technique for increasing the efficiency of training.


Four fallacies concerning training for skilled performance that are often implicitly assumed in training programs are discussed. The fallacies are: 1) practice makes perfect; 2) training of the total skill is optimum; 3) the goal of training is to produce accurate performance; and 4) skill training is intrinsically enjoyable. Automatic/controlled processing theory, which emphasizes how training may be done to avoid these fallacies, is briefly discussed. Finally, 11 training guidelines are provided for the optimization of skill training.

The present paper outlines three major assumptions often implicitly assumed in dual task experiments conducted to assess operator workload. These assumptions are shown to be incorrect. Three criteria which should be met in dual task experiments that draw inferences from secondary task decrements are discussed. An experiment, meeting the proposed criteria, was conducted which demonstrated that when the criteria are met secondary task performance can be predictive of primary task difficulty. However, the data also indicate that a simple assessment of effort alone will not predict total task performance.


Can visual search tasks be combined without cost? To answer this question we had subjects search for one target character in a series of 12 rapidly presented frames. The type of processing, controlled or automatic, was manipulated by requiring search for variably mapped (VM) or consistently mapped (CM) target and distractor sets. Conditions included VM-only search (controlled processing), CM-only search (automatic processing), and simultaneous CM/VM search. Joint automatic and controlled search with emphasis on the controlled search task produced no loss of detection sensitivity in either task but did produce a large criterion shift in the automatic search task. Without instructional emphasis on the controlled search task, controlled search deteriorated. Subjects also showed a tendency to waste controlled processing resources when performing an automatic process. Automatic processing became less resource demanding with practice. However, controlled processing was always sensitive to resource reductions. The results show that subjects can sometimes perform dual search tasks without noticeable deficit when one of the tasks is automatic. The implications of these results are discussed.


Generally speaking, performance declines when humans ingest alcohol; however, there is little precision in predicting alcohol effects. The confusion regarding when and in what situations performance will decline because of alcohol intoxication appears to be due to inappropriate classification of information processes involved in task performance. The present research utilized principles of automatic/control processing theory to examine alcohol effects. Performance on tasks performed via automatic processing showed little decrement due to alcohol but large decrements occurred on controlled processing tasks. Results indicate that the type of practice and therefore the type of information processing predicts performance decrements due to alcohol. Alcohol produces a general reduction in resources and ability to share resources both within and between tasks. The results allow a delineation of alcohol effects and provide a theoretical framework for prediction of alcohol effects.


The present program provides the user with the capability of computing a parametric sensitivity measure, d' and the more distribution-free measures A' and Ag (see Pollack, Norman, & Galanter, 1964).


The NEST program provides the user with a means of checking for nesting errors in RAPTOR programs. It gives users the ability to verify not only the normal open and closed bracket symbols, but, more important, their own bracket-defined symbols (such as ENDIF for closed bracket).
Training High Performance Skills: Fallacies and Guidelines

Walter Schneider

Abstract

A high performance skill is defined as one which, over 100 hours of training are required, substantial numbers of individuals fail to develop proficiency, and the performance of the expert is qualitatively different from that of the novice. Training programs for developing high performance skills are often based on assumptions that may be appropriate for simple skills. These assumptions can be fallacious when extended to high performance skills. Six fallacies of training are described. Empirical characteristics of high performance skill acquisition are reviewed. These include long acquisition periods, heterogeneity of component learning, development of inappropriate strategies, and training of time-sharing skills. A tentative set of working guidelines for the acquisition of high performance skills is described. The report also includes a preface providing abstacts of research carried out during this period.

Fallacies and Guidelines

Introduction

This article examines special considerations and problems associated with high performance skill acquisition. Much of skill learning experience and most skill learning research relate to learning simple skills (e.g., lever positioning). Generalizations based on improvements over short training periods can produce fallacious training assumptions. These assumptions are frequently implicitly assumed in training programs. This paper explicitly identifies some of these fallacies and presents more prevalent assumptions. The section on fallacies is written from a devil's advocate position. It is intended to cause the training program designer to question frequently held implicit assumptions. The next section presents empirical results illustrating special considerations for training highly skilled performance.

The recent increased use of microprocessor-based training emphasizes the need to develop explicit guidelines for skill training. Microcomputers can provide feedback, graphic illustrations, and drill on many components of critical tasks. In the past, training amounted to a combination of classroom instruction and laboratory or on-the-job experience in the final work environment. Now microcomputer programs can be easily modified to train individual component skills, graphically represent the problem, provide augmented cues, sequence the training, and so on. If training program developers blindly make computers perform the same type of simulation activities that were previously done with simulators, there is no reason to expect training efficiency to improve (with the exception of possibly decreasing the number of trainers). For example, an Advanced Controller Exerciser (McCauley, Boot, & Muckler, 1982) was developed to replace the traditional multi-person simulation system for training air intercept control. The microprocessor-based system resulted in poorer trainee performance than the original system. Greater awareness of the special considerations of high performance skill acquisition may enable better use of the flexibility provided by microprocessors.

For the purposes of this paper, high performance skills will be defined as having three characteristics. First, the trainees must expend considerable time and effort to acquire a high performance level (i.e., greater than 100 h). Second, the training programs which produce such skill levels will characterize experience substantial failure rates even among individuals motivated to acquire the skill (i.e., greater than 20%). Third, there will be substantial qualitative differences in performance between a novice and an expert.

Military air traffic and air weapons control provide examples of high performance skills. To develop proficiency requires one to two years of training. Washout rates for training programs vary from 25% to 70%, with 50% being typical. Novices and experts show very different performance characteristics. For example, when performing a two aircraft live intercept, novices have difficulty estimating the turn radius of the aircraft. The novice (13 weeks of training) continues to watch the display for minutes to determine whether the specified turn maneuver will produce the desired effect. In contrast, an expert watching the intercept could specify after only 20 s (two scope sweeps) that the one aircraft is coming in too
Fallacies and Guidelines

34, 60

The training of a fighter pilot provides another example of a high performance skill. This training typically requires 150 flight hours over two years, and washout rates range in excess of 30% (Griffin & Mosko, 1977).

A more mundane example of a high performance skill is that shown in professional-level typing. Typical training time necessary to develop a 50 word-per-minute typing speed is in excess of 200 hours (Deighton, 1971). Most of the people who try to develop typing skill never obtain that level. A highly skilled typist independently moves his/her fingers to different keys simultaneously, whereas the novice makes individual movements to keys sequentially (Runelhart & Norman, 1982).

It is difficult to generalize research for the training of highly skilled performance. First, there are few parametric empirical studies. The studies that do exist often confound effects of training procedure, trainers, and subject criterion differences (Eberts, Smith, Dray, & Vestewig, 1982). Also, since performance changes qualitatively over time, training techniques which may be quite useful for initial acquisition, may be very ineffective for later skill development. Finally, our theoretical understanding of the nature of performance change with practice is very limited. There are some theoretical perspectives that predict qualitative changes in performance (e.g., Lew, 1974; Shiffrin & Schneider, 1977). However, the theoretical development does not yet specify which training techniques would be best at different stages of skill acquisition.

Training Fallacies

Many training programs are based on implicitly assumed fallacies. These fallacies in the sense that they are misleading and are based on unsound generalizations. The next section provides empirical evidence for the unsoundness of the assumptions. All of these fallacies have some truth. However, when taken to extremes, they often produce inefficient training programs. Examples are provided from the training of military air traffic controllers. Examples could easily be drawn from tasks such as typing, pilot training, or reading. The reader is encouraged to assess whether these generalizations can be seen in training programs familiar to the reader. These generalizations are described from a devil's advocate position to encourage the reader to critically consider some commonly held training assumptions.

Fallacy 4 -- Practice Makes Perfect

"Practice makes perfect" assumes that if individuals continue to perform a task, their performance will improve, reaching near optimal levels. For learning simple tasks such as memorizing a phone number, practice makes perfect. However, this assumption does not prove to be a valid generalization for high performance training. For example, in air traffic control training, a large portion of the training time is occupied with the student simply practicing the task. However, many students show only very slow acquisition rates by practicing the task and do not obtain acceptable performance levels by the end of the training program (recall the 50% washout rate).

The statement that practice makes perfect is an over-generalization. Not only does practice often fail to make perfect, it sometimes produces no improvement in performance at all. For example, if subjects practice a digit span task for weeks, subjects who do not consistently group the digits show little improvement in their ability to maintain information in memory (see Chase & Ericsson, 1981). Practice on consistent component tasks does improve component skills (see below). Consistent components are those elements of the task where the subject can make the same response to the stimulus whenever it occurs. When given explicit training on using a strategy to consistently encode the incoming stream, subjects' digit span recall can increase substantially (Chase & Ericsson, 1981).

Fallacy 2 -- Training of the Total Skill

The second fallacy is that it is best to train a skill in a form similar to the final execution of the skill. Total task training is necessary since the final performance is in the target task. However, belief in this fallacy tends to shift the focus of the training into a task target format. A belief in this fallacy seduces one to maximize fidelity even when it yields little training benefit (e.g., see Hopkins, 1975). Training an air traffic controller to perceive turn points at which an aircraft should be radially on track illustrates the inefficiency of total task training. A normal aircraft requires 4 min to sweep out a complete turn on the radar screen. It is difficult to learn to perceive turn radii from such observations. First, because of percent of the task where the subject can make the same response to the stimulus whenever it occurs. When given explicit training on using a strategy to consistently encode the incoming stream, subjects' digit span recall can increase substantially (Chase & Ericsson, 1981).

Fallacy 3 -- Training the Total Skill

The second fallacy is that it is best to train a skill in a form similar to the final execution of the skill. Total task training is necessary since the final performance is in the target task. However, belief in this fallacy tends to shift the focus of the training into a task target format. A belief in this fallacy seduces one to maximize fidelity even when it yields little training benefit (e.g., see Hopkins, 1975). Training an air traffic controller to perceive turn points at which an aircraft should be radially on track illustrates the inefficiency of total task training. A normal aircraft requires 4 min to sweep out a complete turn on the radar screen. It is difficult to learn to perceive turn radii from such observations. First, because of percent of the task where the subject can make the same response to the stimulus whenever it occurs. When given explicit training on using a strategy to consistently encode the incoming stream, subjects' digit span recall can increase substantially (Chase & Ericsson, 1981).

Fallacy 2 -- Training of the Total Skill

The second fallacy is that it is best to train a skill in a form similar to the final execution of the skill. Total task training is necessary since the final performance is in the target task. However, belief in this fallacy tends to shift the focus of the training into a task target format. A belief in this fallacy seduces one to maximize fidelity even when it yields little training benefit (e.g., see Hopkins, 1975). Training an air traffic controller to perceive turn points at which an aircraft should be radially on track illustrates the inefficiency of total task training. A normal aircraft requires 4 min to sweep out a complete turn on the radar screen. It is difficult to learn to perceive turn radii from such observations. First, because of percent of the task where the subject can make the same response to the stimulus whenever it occurs. When given explicit training on using a strategy to consistently encode the incoming stream, subjects' digit span recall can increase substantially (Chase & Ericsson, 1981).
Fallacies and Guidelines

B) One assumes that the real world optimally orders the sequence of events for training. Again one is assuming that the world is fortuitously constructed such that the spacing of practice at consistent task elements is optimal for learning. C) It is best to train a task when attentional capacities are overloaded. If one wants to learn to drive and converse at the same time, one should begin practicing doing both tasks together. D) It is acceptable to be confused about how errors influence performance. Whenever one performs a complex task, it is often difficult to tell which errors caused poor performance or whether those errors were caused due to lapses of attention or an inability to perform the given task. E) One assumes that frustration due to errors and poor performance does not reduce student effort. And finally, F) one assumes that there is little transfer from component task training to total performance. If this last assumption is true, training the total task is the only way to substantially improve performance. However, in many situations there is substantial transfer of component training. For example, training with cardboard models of the cockpit can produce substantial savings in performing the task in the aircraft (Cero, 1973).

Fallacy 1 -- Skill Learning Is Intrinsically Enjoyable

The third fallacy is that skill training is intrinsically satisfying and thus, adding extrinsic motivators is inappropriate. One example is air traffic control training. Air traffic controllers are professionals. Their futures depend on how well they do in the training program. Hence, one would expect little benefit from providing extrinsic motivators. However, being in a darkened room controlling simulated aircraft for 9 hours can get boring. As the training day wears on, it can become difficult to concentrate on one’s work. The problem with the belief in this fallacy is that it can justify a training program designer’s lack of concern about motivating the learner. The problem of motivating the learner is left to the training personnel.

In my laboratory (Human Attention Research Laboratory, University of Illinois) about 1,000 subject training hours are executed per year. Probably the most cost-effective piece of equipment in my laboratory is a noise synthesis chip. This $15 chip can be programmed to emit interesting noises when important events occur. In an air traffic control task, for example, whenever a subject identifies the correct turn point, the aircraft flies the appropriate trajectory with an interesting frequency sweep auditory shot. Before the addition of extrinsic motivators, about 30% of the subjects failed to develop sufficient accuracy in our skill acquisition experiments. After adding extrinsic motivators (e.g., interesting sound effects, interesting visual display patterns, providing criterion-based feedback), failure rates were reduced to less than 5%.

In many training programs the most important determinant of performance is how long a learner spends actively practicing the task. When designing a training program, one must include motivational elements to maintain active participation.

Fallacy 4 -- Train for Accurate Performance

The fourth fallacy is that the primary goal of training a skill is to produce highly accurate performance. In air traffic control, controllers are trained to maintain optimal separation between the aircraft. Training for maximal performance accuracy can be counterproductive. In many skill training programs the goal should be to obtain acceptable accuracy on a component skill while allowing attention to be allocated to other components of the task. In air traffic control, an operator who can maintain optimal separation of only two aircraft would not be an acceptable controller. What is desired is an operator who can maintain a safe separation between 10 aircraft.

Training programs following this fallacy tend to produce operators who can perform individual component skills well but cannot operate well in high workload situations. Specialized training may be necessary to develop skills which will operate well under high workload (see below). Also, in order to achieve reliable performance under high workload, substantial overtraining may be necessary. For example, LaBerge (1976) has shown that training subjects to compare symbols when not directly attending to those symbols requires about six times more training than does training them to compare symbols while attending to the task.

Fallacy 5 -- Initial Performance is a Good Predictor of Trainee and Training Program Success

Belief in this fallacy suggests that if one measures a learner’s performance in the first few hours of training, one can predict performance after hundreds of hours of training (for detailed discussion of this issue, see Ackerman & Schneider, in press). In reality, initial performance of complex skills is very unstable and often provides a poor prediction of final performance. For example, the correlation between the first and fifteenth hour of performing a simple grammatical reasoning task was only 0.31 (Kennedy, Jones, & Harbison, 1980). As the skill becomes more complex, more novel, and requires longer training times, correlations between initial and final performance decrease. Note also that performance may not be a good measure of learning. For example, augmented training may greatly facilitate performance but slow learning.

This fallacy presents a particular problem in evaluating training programs. Certain techniques may work very poorly in a short training program but be very beneficial in a long training program. For example, in a 6 month air traffic control training program, it might be very beneficial to have a 6 hour training module on identifying the heading angles of aircraft. However, in a pilot project, the researcher may have to demonstrate the benefit of a particular module with a simulated training program that is only 6 hours in length. It is likely that whole task training would be the most effective training in a 6 hour time scale even though a combination of part- and whole-task training would be better for a 6 month training program.

Fallacy 6 -- Once the Learner Has a Conceptual Understanding of the System, Proficiency Will Develop in the Operational Setting

This fallacy leads to training programs which present technical information in a classroom setting and provide minimal instruction on how to use this information in performing the task. For example, in air traffic control the classroom teaching describes the aircraft performance characteristics. However, the student may not be shown explicitly what those performance characteristics look like on the radar scope. Often operators need a great deal of experience with the system even
after they have learned to conceptualize it accurately. For example, it is relatively easy to visualize a mental model for a manual transmission. However, many hours have to be spent in a car before gear shifting becomes proficient.

**Fallacy Summary**

Most training programs regard one or more of the above fallacies as true principles. If one rejects these fallacies, training design becomes much more difficult but potentially more successful. First, instead of assuming that practice is sufficient to produce high skill levels (Fallacy 1), one might emphasize practicing consistent components of the task. Second, as opposed to training the learner in the total task (Fallacy 2), one should break down the task and re-represent the task to maximize the learning rate on each component. Then one should sequence the various components to maximize the integration of the task. Third, one cannot assume that the material itself is sufficiently motivating for the learner (Fallacy 3). One may need to design extrinsic motivators into the task. One should find extrinsic methods for motivating the trainee without interfering with the actual teaching process. Fourth, one cannot train for perfect performance (Fallacy 4); one should train to what would be considered an acceptable performance level but also train so that the learner can perform the task with little or no attention allocated to consistent components of the task. Fifth, one should be very cautious when extrapolating results from short training periods to predicting either successful training procedures or successful trainees (Fallacy 5). Sixth, one should recognize that providing the learner a theoretical understanding is often only the first stage in developing a high performance skill (Fallacy 6).

**Empirical Characteristics**

To develop effective training programs, it is useful to know the prominent features of high performance skill acquisition. The characteristics of developing a high performance skill are quite different than those of learning declarative information. Teaching declarative information generally includes presenting the information once in a classroom-type setting. In contrast, developing a skill entails presenting comparatively fewer "facts" per unit of time but requires the learner to devote a great deal of effort in developing and practicing component skills. The training program designer must be cautious not to use the academic training model as a model for high performance skill acquisition. The training program designer should be aware of the nature of skill acquisition functions, the heterogeneity of component tasks, the need to discourage poor strategies, and the need to train time-sharing skills.

**Extended Practice Function.** The first class of training problems relates to acquisition functions. High performance skill acquisition is characterized by log-log acquisition functions, improvement over long periods of training, initial instability, and false asymptotes. Practice curves are fit with a variety of functions. Reaction time data are most commonly fit with power functions and exponential functions (see Newell & Rosenbloom, 1981). Performance rating scale data are generally fit by logistic and exponential functions (see Spears, 1982). These various curve fitting procedures show very high correlations (see Newell & Rosenbloom, 1981). The following discussion uses the power function for illustration, but none of the arguments would change if any of the other curves were utilized, since all show fast initial acquisition rates with gradual approach to an asymptote. The speed of responding is well fit as a power function of the number of trials (see Newell & Rosenbloom, 1981). The power law predicts that the log of the time to complete a response will be a linear function of the log of the number of executions of that particular response (see Figure 1). The power law is stated in the form of:

$$T = B N^{\alpha}$$

$$\log(T) = \log(B) + \alpha \log(N)$$

$T$ is the time to respond, $N$ is the number of trials and $B$ and $\alpha$ are constants (with $\alpha < 1$). The power law predicts that if reaction time to perform a response decreased from 10 s to 5 s over the first 100 trials of training; at 440 trials, response time will be 4 s; at 1,978 trials, response time will be 3 s; and a total of 10,000 trials of training would be necessary to reduce reaction time to 2.5 s.

Newell and Rosenbloom (1981) refer to the power law as the "ubiquitous law of practice." The power law holds for a wide range of response time tasks. Predictions from the law fit data including: operating a cigar rolling machine (Crossman, 1959), adding digits (Crossman, 1959), editing text (Moran, 1980), playing card games (Newell & Rosenbloom, 1981), learning a choice reaction time task (Seibel, 1963), detecting letter targets (Neisser, Novick, & Lazar, 1963), and performing geometry proofs (Neves & Anderson, 1981). The training system designer can use the power law to predict performance improvement of component skills as a function of practice (for an excellent example of this type of prediction, see Card, Moran, & Newell, 1983).

A major feature of high performance skills is that they show improvement over extended periods of time. For example, Crossman's (1959, see Figure 1) subjects showed improvement in operating a cigar rolling machine over 3 million trials and 2 years.
The continued improvement after extended practice has two implications. First, high performance requires a great deal of practice even after the trainee understands the nature of the task and can perform the task accurately. Second, it may be beneficial to design training procedures which will allow the trainee many trials of performing critical component tasks.

The third characteristic of acquisition functions is that initial performance (e.g., the first 100 trials) is likely to be an unstable predictor of later performance. For example, Kennedy, Jones, and Harbison (1980) found that the correlation between Day 1 and Day 15 of performing a grammatical reasoning task was 0.31. The subjects showed different initial performance levels, acquisition rates, and final asymptotes.

This instability of early acquisition stems from at least four sources (see Ackerman & Schneider, in press). First, the rate of improvement during the first few hundred trials is very rapid, causing a large within-subjects variation. Second, subjects with differential experience with related tasks start out at different performance levels. For example, assume one wanted to assess an individual's ability to perform an air traffic control task by running a simulated control session. Assume also that subjects who had played 20 hours of a particular videogame would start at this task with the equivalent of 2 hours of training. In a 1 hour test of performance these videogame-wise subjects are likely to perform substantially better than non-videogame-wise subjects. However, since the actual training program would require hundreds of hours of training, an individual with a faster learning rate will surpass someone with a 2 hour head start in training.

A third source of initial instability is that individuals vary in their rate of acquisition of skills (e.g., see Kennedy et al., 1980). It typically requires hundreds of trials to reliably assess learning rate, and hence it is difficult to get a quick estimate of this parameter. A fourth source of initial instability is that different abilities appear to limit performance at different stages of practice. Early in the training program, general cognitive abilities are critical for following instructions. Later, the basic psychomotor abilities may become the limiting factor (Fleishman & Rich, 1963).

A fourth characteristic of acquisition functions is that the willingness of the learner to continue to practice the task strongly influences final performance level. Most individuals can be motivated to perform even very boring tasks for a few hours. However, many individuals cannot maintain motivated performance when practicing a skill for tens or hundreds of hours. Most of the people who purchase a musical instrument never practice long enough to achieve even basic levels of proficiency with the instrument. In dual task studies (e.g., Schneider & Fisk, 1984) substantial numbers of college student subjects fail to continue to put their full effort into learning the task after about 6 hours of training (even when they risk loss of bonus pay). This often does not result in a decrement in performance but rather a plateau or lack of improvement in performance. Bryan and Harter (1899) commented about the difficulties of overcoming plateaus in the development of skill. A training program designer must help the trainee continue to expend the effort to improve performance over very long practice periods. Note, with proper motivation, performance plateaus seem less likely (see Keller, 1958).

Figure 1. Time taken to make cigars as a function of practice (Crossman, 1959).

Fallacies and Guidelines
Heterogeneity of Component Improvement Rates. The second class of problems in skill development relates to the heterogeneity of component improvement rates. Performance on most complex tasks is determined by a variety of component skills. For example, in air traffic control performance is determined by perceptual skills for identifying the locations and trajectories of the aircraft, cognitive skills to predict events and schedule traffic, and output skills related to handling communications and operating equipment. Improvement rates for these different task components may vary widely. For example, the keying time may decrease by 5% over thousands of trials of training, whereas the time to decide how to schedule traffic may be reduced by 30% in the same number of trials.

Consistent task components show large improvements with practice, whereas varied components do not. A consistent component is defined as one in which the subject can make the same response to a particular stimulus situation every time the stimulus situation occurs. For example, in a consistent letter search task the subject would push a specific button every time the letter "E" appears on the display. In contrast, a varied component is one in which the mapping between the stimulus and response varies across trials. For example, in a varied letter search task, a subject might search for and respond to "E's" on one trial, but on the next trial, the subject might search for "T's" and not respond to the letter "E." Figure 2 shows the data in a letter search experiment manipulating consistent and varied practice at the task. In the consistently mapped condition subjects responded to target stimuli letters whenever they occurred. Detection accuracy improved substantially over some 860 training trials. In contrast, in the variably mapped condition in which subjects responded to different stimuli on different trials, there was no improvement over trials (Schneider & Fisk, 1982b).

Similar to letter search, motor responding tasks show improvement primarily on consistent sequences. In a motor output sequential responding task subjects reproduced sequences of eight button pushes. After the first 10 trials, execution of varied response sequences showed no improvement in speed or accuracy. The execution with consistent sequences improved in accuracy (30%).

Figure 2. Detection accuracy as a function of training trials in a letter search task (from Schneider & Fisk, 1982b). CM refers to a consistently mapped letter search, VM to variably mapped.
Fallacies and Guidelines

speed (22%), and response variability (50% relative to the varied sequences) during 50 training trials (see Schneider & Fisk, 1983). Chase and Ericson (1981) found that subjects who varied in their grouping of digits showed little improvement with practice in a digit span task. However, when they consistently grouped the digits and associated them with salient classes of events, there was substantial improvement with practice. With a year of practice, one subject was able to increase his digit span to 80 digits.

Consistent task components show substantial improvement in processing speed with practice, whereas varied components do not. In a category search experiment, Fisk and Schneider (1983) determined the increase in reaction time as a function of the number of category judgments made. The slope relating comparison time to number of comparisons was 200 ms per comparison in the varied search condition and only 2 ms per comparison in the consistent search condition. In this case there was an increase of two orders of magnitude in processing speed for the consistent component relative to the varied component.

Consistent task components can be performed with little or no attention, whereas varied task components are resource sensitive even after extended practice. Figure 3 shows the data from a dual-task experiment (Schneider & Fisk, 1984). The primary task required comparing two digits in memory to two digits on the display that changed every 400 ms. The secondary task required detecting words from a given semantic category. If the category task was consistently mapped, subjects could perform the category task equally well whether they performed the digit task or not. In contrast, if the category task was variably mapped, extended training produced no improvement in the ability of the subjects to time-share the category and digit search tasks. Kennedy and Bittner (1980) provide another example of lack of improvement in time-sharing ability in a task requiring varied responding. They found that when subjects had to count tones on two channels, there was no performance improvement over 15 days of practice. Most time-sharing studies have consistent components and show substantial improvement with practice (e.g., Demos & Wickens, 1980; Schneider & Fisk, 1982).

Heterogeneity of component skills complicate the problem of assessing trainee performance and providing knowledge of results. For example, in air intercept control of a stern attack the final performance measure is how accurately the fighter is placed behind the enemy aircraft. This single score is determined by perceptual errors in assessing the heading, speed, and turn point of the aircraft; cognitive errors in planning the strategy or deciding the turn point; or motor errors in controlling the equipment and giving commands to the pilot. A single performance score provides the learner with little detail about how to improve performance. In heterogeneous tasks it is important to provide the learner knowledge of results about the performance of individual component skills (see Newell, 1976). By developing component skill tests, an operator's weaknesses in various phases of a complex task can be illustrated.

Heterogeneity of component skill learning rates may require differential training of components. In the air traffic control training modules improvement rates range from 23% improvement during the first 100 trials for identifying the turn point of an aircraft to only a 4% improvement over the first 100 trials at identifying heading angles. After the first 1,000 trials at training to identify

![Figure 3. Single (filled symbols) and dual (open symbols) task category search detection accuracy as a function of sessions of practice (Schneider & Fisk, 1984). Note the elimination of the dual task deficit in the CM condition (solid lines for consistent category search) and stability of the dual task deficit for the WM condition (dashed line for varied category search).](image-url)
turn point, inaccuracy in this component no longer limited total performance. In contrast, heading identification would improve for thousands of trials. To optimize final trainee performance, training time should be allocated so as to maximize total performance improvement per unit of training time.

Eliminating Poor Strategies. A third class of problems in high performance skill training is the need for special training to eliminate poor strategies. In many skills a variety of strategies may be used to perform a particular task component. However, one strategy usually allows better development of more advanced skills. For example, in driving an automobile one can drive by lining up the hood ornament to the road stripe or by estimating the angle of the curve and turning the steering wheel appropriately for that angle. The former strategy is easier to learn but requires many more decisions and results in higher workload.

Trainees may be resistant to discontinuing the use of a strategy that is easy to learn but results in high workload. In laboratory search tasks some subjects require specialized training in order to enable a low workload strategy. Most subjects can perform a consistent category search task in combination with a high workload digit task without deficit (Figure 3, also Schneider & Fisk, 1983). However, occasionally a subject seems unable to perform a category search task under high workload. Figure 4 illustrates data from two such subjects. In dual-task conditions, the subjects' category detection rate was only about 30% of what it was in the single task conditions. During 4 hours of testing there was no evidence that these subjects were improving. This lack of improvement in consistent category search contrasted to the data of most of the subjects (see Figure 3). At the point of Session 4, these subjects would be considered washerout subjects. They could not adequately perform this task under high workload. These subjects were then trained to perform an easier semantic search and digit detection task. When the subjects were successful at learning easier categories, they returned to the original category condition in which they were having difficulty. The subjects' dual task performance increased from the previous 30% to 84% accuracy even though they had had no training on the specific category detection task. The subjects reported that during the interim training they had learned to "let go" and to respond to the words without thinking about them. Once these subjects had learned to "let go," they could perform the category task with high accuracy even while performing a high workload task (for details, see Schneider & Fisk, 1983).

Figure 4. Subjects' single and dual task consistently mapped category search detection accuracy. After Session 4, subjects were given alternative training to facilitate "letting go" of attentionally processing the words (from Schneider & Fisk, 1983).
The acquisition of reading skill provides interesting illustrations of the need to break bad habits. LaBerge and Samuels (1974) report that some readers become overly concerned about word encoding accuracy. These readers focus too much effort at doing the word encoding skill and have few attentional resources available for semantic integration of the task. Frederiksen (Frederiksen, Weaver, Warren, Gillette, Rosebery, Freeman, & Goodman, 1983) reports that poor readers develop a strategy of looking at the first and last letters of the word and guessing what the word would be. These readers can continue to use this strategy for years without substantial improvement in reading. One method which seems to successfully break this habit is to have the reader identify word units within computer-presented words (push the button any time the letter pattern “min” appears). Such training enables the learner to focus on elements of the language which do produce accurate performance and hence show substantial practice effects.

Developing Time-Sharing Skills. High performance tasks often require the development of specialized time-sharing skills. In order to perform two tasks simultaneously, it is critical that the two tasks be practiced together (e.g., Damas & Wickens, 1980). Training with different task priorities may be necessary to help subjects determine the optimal allocation of attention for maximal skill performance (Gopher & North, 1977). It is difficult to train subjects to respond to multiple channels simultaneously (see Duncan, 1980). Schneider and Fisk (1983) have found that practice in conditions requiring a high frequency of simultaneous responding can greatly improve the operator’s ability to deal with occasional situations requiring simultaneous responding. It may also be necessary to train subjects so that they can determine how to trade-off speed for accuracy. In reading, for example, with full attention allocated to word encoding, word encoding accuracy may be 98%; however, only a small amount of attentional resources are available for semantic processing. In contrast, if word encoding is done with minimal attention to the word encoding, encoding accuracy may be 98%, but the majority of attentional resources will be available for semantic processing. The latter strategy may result in far better comprehension. Finally, in very high workload situations, the operator must often triage through the list of priorities deciding which tasks must be left undone. Munro, Cooper, and Toms (1982) have shown that as workload increases in a simulated air traffic control task, the operators intentionally ignore information in order to cope with the critical aspects of the situation.

This review illustrates salient features for training in high performance environments. This review is not intended to be exhaustive (for general reviews on skill acquisition, see Anderson, 1981; Welford, 1968, 1976). The emphasis here has been on identifying those features of the training program which enable the novice to become a high performance practitioner.

Working Guidelines

The training program designer implicitly or explicitly develops the training program in accordance with certain rules. The six fallacies discussed in this paper indicate the assumptions the designer should not make. In the last 5 years at the Human Attention Research Laboratory subjects have been trained for over 10,000 hours in a variety of skill acquisition experiments. Current laboratory training programs include air traffic control and electronic troubleshooting tasks. The following is a list of rules used in the design effort. These were developed out of basic research on developing visual search tasks (see Schneider, 1982). These rules should be treated as an initial set of working guidelines and are provided to help focus discussion and research in order to develop an explicit set of training rules. The laboratory is engaged in a long range research project to evaluate the effectiveness of these guidelines in applied training programs. (For a more detailed discussion of these guidelines, see Schneider, 1982).

These rules are based on the proposition that human performance results from the interaction of two qualitatively different forms of processing (James, 1890; LaBerge, 1976; Norman, 1976; Posner & Snyder, 1975; Schneider & Fisk, 1983; Shiffrin & Schneider, 1977). These two forms are referred to as automatic and controlled processing. Automatic processing is a fast, parallel, fairly effortless process which is not limited by short-term memory capacity, is not under direct subject control, and performs well-developed skilled behaviors. Automatic processing typically develops when subjects deal with the stimulus consistently over many trials. Controlled processing is characterized as a slow, generally serial, effortful, capacity-limited, subject-controlled processing mode that must be used to deal with novel or inconsistent information (see Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Controlled processing is expected when the subject’s response to the stimulus varies from trial to trial. From the automatic/controlled processing perspective, training should develop automatic component skills to perform consistent task components and develop strategies to allow limited controlled processing resources to inconsistent or poorly-developed task components (see Schneider & Fisk, 1983).

Rule 1 -- Present information to promote consistent processing by the operator. In order to train fast, low workload processing, the operator must perceive and deal with situations consistently. Developing consistent processing can be done in a variety of ways, including the use of analogies, providing specialized representations of the problem, and adaptive training. In teaching electronic circuitry, circuitry may be beneficial if the operator can develop an analogy of a water-flow process for a particular logical element. As the operators visualize the analogy, they consistently make the same response to a given situation, developing automatic component processes. After several hundred trials, the operator can specify the output with little effort and without the use of the analogy. To make a judgment of where an aircraft should turn, a maneuver which normally takes about 5 min, would take place in about 0.5 s. By compressing time in this way, we can provide the trainee more trials at executing this particular component
in a single day of training than he/she could get in a year of training with conventional methods. In order to offer extended training experience it may be necessary to compromise on simulation fidelity to increase the number of training trials that can occur per hour of training.

Rule 3 -- Do not overload temporary memory and do minimize memory decay. In training the air traffic control task, during initial acquisition the flight path is drawn on the display so that operators need not retain that information in memory. This facilitates the maintenance of a consistent memory representation and speeds the development of automatic component processes to identify turn radii. The number of new tasks to be performed concurrently should be limited to minimize attentional overload.

Rule 4 -- Vary aspects of the task that vary in the operational situation. When developing automatic components, the components must generalize to the entire class of situations to which they are appropriate. For example, when training the operator to identify the turn point for an aircraft, the test intercepts occur at all possible locations on the screen. If all turn identification occurs with the aircraft in the center of the screen, the skill may not generalize well to other positions.

Rule 5 -- Maintain active participation throughout training. Active participation is enhanced if subjects need to respond every few seconds. For example, to train subjects to visualize solution spaces, subjects observe the solution space going through a range of intercept angles and then are presented a test vector. The subject must identify whether that test vector is appropriate for that intercept angle. Without these frequent tests, subjects’ observation becomes passive, and there is little improvement with practice.

Rule 6 -- Maintain high motivation throughout the training period. Provide the trainee extrinsic motivation to maintain high levels of effort. When subjects respond incorrectly, a simulated crash can occur. Adaptive training can sequence subjects to ever more difficult training conditions but still allow them to experience a high degree of success throughout. In order not to significantly reduce training time, the motivational feedback should be limited to a small portion of the training period (e.g., less than 5%).

Rule 7 -- Present the information in a context which illustrates more than the to-be-learned task. For example, when subjects identify proper intercept points, feedback shows how the planes would fly to that intercept point, thus illustrating the trajectories of the flight path while the operator tries to perceive the final rendezvous point. Caro (1973) recommends training flight skills within a functional mission context. The trainer must, however, be careful to not overload the subject (Rule 3) and efficiently train component skills (Rule 2).

Rule 8 -- Intermix component training. Intermix the training on various component skills rather than train each component individually before proceeding to the next component. This intermix training distributes the practice and facilitates perception of the interrelationships of the components. The proportion of component and total task training time should be allocated so as to maximize final total task performance.

Rule 9 -- Train under mild speed stress. Automatic components are fast processes, probably occurring in less than half a second. Speed stressing subjects improves the development rate. When not speed stressed, subjects tend to use slow, controlled processes which may not be acceptable in the operational environment. For example, in the turn point identification task, air traffic control operators are expected to make a response in less than 2 s.

Rule 10 -- Train strategies which minimize operator workload. In many tasks there are multiple strategies which involve differential workload. In air traffic control the operator is allowed only one decision to get the airplanes together in a rendezvous. If operators develop a strategy of many small corrections during training, the workload imposed by that strategy makes it difficult to handle sets of five aircraft.

Rule 11 -- Train time-sharing skills for dealing with high workload environments. Train operators in situations requiring: using different speed and accuracy trade-offs; frequent simultaneous responding; and triaging through task priorities. Structure the training so the expert can perform reliably even during those rare occasions when his/her skills are pushed beyond reasonable limits.

Conclusion

There are special problems associated with training high performance skills. It is difficult to get an appropriate perspective of the changes that occur during months and years of training. Certain assumptions which work well in short-term training programs may be fallacious when extended to long-term training programs. The training program designer needs to understand the assumptions underlying each given training procedure. The trainer should be aware of the special problems of acquiring high performance skills. The research community must work to develop and test an adequate set of guidelines for high performance skill acquisition. With appropriate perspective, research, and guidelines the current computer revolutions can flow into a training revolution.


