1986 N87-26701

NASA/ASEE SUMMER FACULTY RESEARCH FELLOWSHIP PROGRAM

Johnson Space Center
University of Houston

# TRAINING FOR LONG DURATION SPACE MISSIONS

| | |
|---|---|
| Prepared by: | Joseph H. Goldberg, Ph.D. |
| Academic Rank: | Assistant Professor |
| University & Department: | The Pennsylvania State University, University Park, Pennsylvania |
| | Department of Industrial & Management Systems Engineering |
| NASA/JSC | |
| Directorate: | Engineering |
| Division: | Advanced Programs Office |
| Branch: | Systems Definition |
| JSC Colleague: | John W. Alred |

| | |
|---|---|
| Date: | July 23, 1986 |
| Contract Number: | NGT-44-005-803 (University of Houston) |

# TRAINING FOR LONG DURATION SPACE MISSIONS

Joseph H. Goldberg
Assistant Professor
Department of Industrial & Management Systems Engineering
The Pennsylvania State University
207 Hammond Building
University Park, PA 16802

The successful completion of an extended duration manned mission to Mars will require renewed research effort in the areas of crew training and skill retention techniques. The current estimate of in-flight transit time is about nine months each way, with a six month surface visit, an order of magnitude beyond previous U.S. space missions. Concerns arise when considering the level of skill retention required for highly critical, one-time operations such as an emergency procedure or a Mars orbit injection.

The objectives of this research project were to review the factors responsible for the level of complex skill retention, to suggest optimal ways of refreshing degraded skills, and to outline a conceptual crew training design for a Mars mission.

Currently proposed crew activities during a Mars mission were reviewed to identify the spectrum of skills which must be retained over a long time period. Skill retention literature was reviewed, to identify those factors which must be considered in deciding when and which tasks need retraining. Task, training, and retention interval factors were identified. These factors were then interpreted in light of the current state of spaceflight and adaptive training systems. Finally, the retention factors formed the basis for a conceptual design of Mars mission training requirements.

## TABLE OF CONTENTS

# TRAINING FOR LONG DURATION SPACE MISSIONS

It is human nature to forget highly learned information. Over time, psychomotor skills that may have been overlearned degrade into awkward movements at a later time, while ordered sequences and events rapidly become disordered. The study of human skill retention and degradation has been ongoing for many decades; useful information exists, but a comprehensive model of skill retention as a function of independent task and individual factors must still be developed.

This paper considers these skill retention factors in light of a long-duration spaceflight, such as a manned mission to Mars. Retention of finely tuned skills and knowledge is absolutely necessary for the successful completion of such a mission, yet man-machine system complexities are becomming more and more complex. These skill retention issues also have implications that go far beyond manned spaceflight. Industry must train workers, often for long periods of time and with concomitant losses in productivity. Colleges and universities are also in the business of training individuals with skills and knowledge for long-term retention. Clearly, accurate long-term retention of skills and knowledge is important for productivity and safety within the entire society.

This paper is divided into three major sections in its examination of long duration skill retention for manned spaceflight. (1) Currently proposed crew activities in a manned mission to Mars are reviewed. This information gave a concrete focus to the skill retention issues described here, and allowed bounds to be placed upon the duration and nature of training and retention. (2) Recent psychological and Human Factors literature (within the past 30 years) on factors influencing long-term retention was reviewed, concentrating on results and conclusions, rather than a critique of methodology. The purpose of this section was to provide a framework by which retention tactors could be studied, and to subsequently provide the necessary framework for a future model of skill retention time and quality. (3) A section on training for space missions was included. The intent of this section was to provide a foundation from which longer duration retention training could build, and to outline the required training elements for an advanced, long duration Mars mission. While much of such a conceptual design is merely an exercise in futurist guesswork, an attempt was made to logically build on the concepts presented in earlier sections of this report. A description of the research yet needed to develop a working skill retention model was also included. Such a model would greatly aid the conceptual design of Mars and other long duration missions, as well as industrial job design. Information for Sections 1 and 3, where not otherwise noted, was obtained from personal communication and experience at the Johnson Space Center, Houston, Texas. The author wishes to acknowledge Jack James, Andy Petro, and John Alred for their valuable assistance.

## 1. MARS MISSION ACTIVITIES

A current scenario for an interplanetary mission to Mars includes several distinct phases:

1. Earth lift-off to low Earth orbit (possible at space station)
2. Taxi-transfer from space station to orbiting interplanetary vehicle
3. Transit to Mars
4. Docking with second orbital transfer vehicle at Mars

5. Land on surface
6. Reverse sequence for return mission

The time involved for such a mission is on the order of 5 to 9 months each way, depending on orbits, trajectories, etc., with a 6 month or longer surface visit. Using current and near-term technology, the transit times should not appreciably differ from this estimate, but the time spent on the surface could dramatically change with the addition of a permanent manned Martian base.

## 1.1. CREW ACTIVITIES

The activities to be carried out in a Mars mission are as varied as those in everyday life. They may be broken down into four areas, as shown below. Sources of information on crew activities and events include Oberg (1982), Oberg and Oberg (1986), Joels (1985), Conners et al. (1985), and National Commision on Space (1986).

Spacecraft Control and Maintenance. These critical activities will insure that mission success and crew safety will not be comprimised. Crew members, armed with automated equipment and extensive computer programs, must serve as diagnosticians, continuously asking new questions about the status of equipment. As discussed by others, maintenance will range from simple modular replacement of Lithium cannisters, to large-scale reconstruction or evacuation of space vehicle apparatus. Both dexterity and decision ability will be required.

Scientific Study. Much investigation will continue to be performed in space. Many accounts have indicated an even broader range of research topics than in previous space missions. These will include not only physical and hard science topics, but will be expanded to social and behavioral science issues including space habitability, behavioral interaction, and group power structures.

Crew Health Maintenance. The health of the crew will include both regular physiological and psychological screening. Many innovative diagnostic and treatment procedures will be developed for long duration space missions, based on prior space station experience. Intelligent computer programs, in the form of expert systems, will likely be extensively used as guides for diagnosing and studying new forms of illnesses.

Recreation. Many have stressed the importance of recreation in an interplanetary mission (e.g., Fraser, 1968). Alternative cognitive and physical activites during off-duty time will be important to maintaining a healthy crew.

## 1.2. FACTORS IN A LONG DURATION MISSION

Skill training and retention requirements for a Mars mission will necessarily differ from that required by all previous missions, as determined by criticality and duration of events. Those skills that quickly degrade must be refreshed often or continuously, while better retained skills need only be refreshed periodically. A training program for this mission must consider several factors that are unique to a manned interplanetary journey, as listed below.

Skill Retention Duration. The required skill retention interval, between training and actual performance, may be 6 months or more. This is an order of magnitude greater than previous U.S. missions, and presents many unknowns for complex skill and procedures retention. Special on-board refresher training will be required for spme of these degraded skills.

Crew Autonomy. As one-way communication lags of 10 to 30 minutes will be encountered near Mars, Earth-based mission control will be of little use. Instead, ground-control will serve as an independent opinion source and coach for an autonomous crew. The crew of 6 to 8 will function as a team, with each member contributing complementary expertise. Crew training thus must focus on enhancing those traits that increase this autonomy, and counter the negative effects of group thinking.

Crew Confinement. The adverse effects of long-term confinement must be well understood before undertaking this mission. Training for long-term confinement must be considered, and techniques of countering confinement, such as projecting video landscapes, may be necessary. Study of analog confinement environments, such as prison or arctic stations, will aid in this definition. .

Criticality of Skills. Some required skills, such as orbital docking, will have a criticality beyond all other skills. Many of these will be performed only once or twice in a mission, after a long no-practice duration. The effects of real and perceived skill criticality on performance and training must be understood before undertaking a Mars mission.

Automation. Extensive use of artificial intelligence and automated sensing and diagnosing apparatus will be used for routine spacecraft control and maintenance. The crew will be responsible for monitoring this equipment, and factors determining crew monitoring or vigilance performance must be understood. A useful human-machine allocation model must be developed, and training for this will be required. NASA has already taken a first step in defining this model (von Tiesenhausen, 1982).

Workload. The effects of mental and physical workload must be modeled before initiating a long mission, to allow a constant performance level within an autonomous crew. The choice of how many crew members to allocate to tasks should be determined via a generic workload modeling computer program.

Environment. The adverse effects of vibration, noise, radiation, ion concentrations, and carbon dioxide are among the many environmental factors whose effects will be felt over the entire mission. The effects of these factors on health and skill retention must be considered in the design of the Mars mission.

These important factors must all be considered when designing a training program for a long duration mission. While shorter duration mission crews have tolerated and even performed well under some of these factors, their effects will be exacerbated by long-term confinement. Since a Mars mission is an order of magnitude beyond current missions in duration and complexity, its training program cannot be evolutionarily developed. Instead, a rethinking of training is required; a model specifying training needs by type of skill and degradation level must be developed. The purpose of this paper is to take an initial step towards such a model, by indicating those factors that affect skill retention, and thus training requirements.

## 2. SKILL RETENTION

The duration and quality of skill retention should necessarily determine the training requirements of a long duration space mission. Skills that quickly degrade must often be refreshed, whereas better retained skills may be neglected for a longer time. Before considering this literature, several qualifications must be made, however. (1) Reports of studies in this area are often not readily obtainable. This may be due to the fact that much of the training research has been conducted in the private and military sectors, which have little impetus to publish in widely distributed publications. Also, much of this research is very task specific, and investigators may have have felt that their research would have low utility outside their immediate scope. (2) The major retention factors are covered below as discrete topics, but all are intimately intertwined and confounded. Differences in the length of a post-training retention interval, for example, are confounded with the type and duration of initial training. Conclusions drawn here must clearly be interpreted with a great deal of caution. To gain a better understanding of these factors, however, they are discussed separately, ignoring conjoint and interactive effects. (3) In some cases, conclusions were necessarily drawn from very few studies, clearly scientifically inappropriate. This was pragmatically done to at least provide a direction for future research needs and developments.

Naylor and Briggs (1961) reviewed over 60 years of literature, and created the first categorization of retention-influencing variables. (1) Task variables included the procedural/tracking task dichotomy introduced below. They raised the important issue that the difficulty and organization of a task is likely responsible for observed retention differences. (2) training variables included three subclasses of factors: the amount of initial training, distribution of training sessions, and transfer effects from other tasks. (3) Retention interval variables included those factors present within this period. (4) Recall variables consisted of other retention-influencing factors, such as the training fidelity, or the presence of any warmup activity prior to retention testing. The present review drew heavily on this work, and extended their factor categories. A subsequent review (Gardlin and Sitterley, 1972) covered many skill retention studies, under contract to NASA. These investigators provided annotated reviews of many studies that were directly applicable to the piloting of space vehicles. The present review also drew on this paper, but was broader in coverage.

Selected skill retention studies cited below are summarized in Table 1, which presents the following information: (1) Investigator(s), (2) Retention: time interval between end of training and initial retest, (3) Task: type of performance task. P: procedural (discrete), T: tracking and control (continuous), (4) Indp. Var.: independent or manipulated variable(s); D: duration of training, R: retention interval, S: structure of training, F: fidelity of training, O: organization of task, RR: retention interpolated activity or rehearsal, (5) Task Description, (6) Cmplxty: complexity of the task(s), subjectively estimated by the number and type of simultanous activities that had to be performed, (7) Training: method of training; duration or criterion, (8) #Ss: number of subjects tested across entire study, (9) Exper: subjective subject experience at task; all subjects were inexperienced in abstract tracking tasks, whereas some aircraft control studies utilized experienced pilots; I: inexperienced, E: experienced, (10) Dependent Var.: dependent or measured performance variable(s)

The results of these studies, referred to throughout this report, are summarized in Figures 1, 2, and 3. Manipulations of retention interval, training duration, and training organization are shown, and every attempt was made to combine similar studies with identical dependent parameters into one figure. Figure 1 shows performance in procedural tasks, or those requiring cognitive control or sequencing over many procedural steps. Note that three time scales, to allow sufficient resolution, were used on the retention time axes: 0-24 months, 0-6 months, and 0-4 weeks. Dependent variables here included both errors and time to complete procedures. Figure 2 shows performance in simple tracking tasks over the same retention intervals as the procedural tasks. Dependent measures here included integrated error in volts, inches, or arbitrary numbers, measuring the deviation between a target and one's ability to follow it. Other measures included the acqusition time to capture a target, or the percentage of total time on a target. All of these parameters generally required some form of continuous sampling by the experimental apparatus. Figure 3 also shows continuous tracking, but only for studies which presented much more complex flight control tasks. These experiments often used open- or closed-loop simulators of airplanes or space vehicles. All retention interval axes here were 0-6 months. Dependent measures usually consisted of a large collection of parameters, of which a subset was chosen, such as altitude error from a preset flight path.

Table 1. Skill Retention Studies

| Investigator(s) | Retention | Task | Indp.Var | Task Description | Cmplxty | Training | #Ss | Exper | Dependent Var. |
|---|---|---|---|---|---|---|---|---|---|
| Neumann and Ammons (1957) | 1.75, 12 months | P | R | Switch setting for abstract task; 16 steps | Low | Actual task/ to criterion: 2 error free trials | 40 | I | Number of correct switch settings |
| Ammons et al (1958); Exp 1 | 1,6,12,24 months | P | D,R | Switch and control setting for an abstract task; 17 steps | Low | Actual task/ 5 vs. 30 trials | 538 | I | performance time |
| Ammons et al (1958); Exp 2 | 1,6,12,24 months | T | D,R | Pedal and stick control of model airplane | Medium | Actual task/ 1 vs. 8 hours | 465 | I | Mean time on target |
| Mengelkoch et al (1960; 1971) | 4 months | T,P | D | Partial closed-loop control of airplane simulator; 50 min. flight | High | Classroom; Simulator/ 9.0 to 13.5 hours | 26 | I | Procedural errors; Flight control error measures |
| Fleishman and Parker (1962) | 1,5,9,14 months | T | S,R | Partial closed-loop control of aircraft simulator; tracking dot on oscilloscope | Medium | Self-study; simulator/ 6 hours | 70 | I | Integrated tracking error |
| Naylor et al (1962; 1968) | 1 month | T,P | D,O | Procedural switch setting (9 steps) & 3-D tracking task | Medium | Simulation; part- and full-task/ 2 vs. 3 weeks | 128 | I | Integrated tracking error; Procedural response time and errors |
| Hammerton (1963) | 6 months | T | D | Second-order (acceleration) tracking of dot on CRT | Medium | Actual task/ to criterion | 18 | I | Target acquisition time |

Table 1 (cont). Skill Retention Studies

| Investigator(s) | Retention | Task | Indp.Var | Task Description | Cmplxty | Training | •Ss | Exper | Dependent Var. |
|---|---|---|---|---|---|---|---|---|---|
| Trumbo et al (1965a) | 1, 5 months | T | D,O,R | Tracking line on oscilloscope | Low | Actual task/ 50 vs. 100 trials | 250 | I | Integrated tracking error; temporal and spatial accuracy |
| Trumbo et al (1965b) | 1 month | T | D | Tracking line on oscilloscope | Low | Actual task/ 115 vs. 130 trials | 120 | I | Integrated tracking error; other tracking indeces |
| Grodsky et al (1964; 1966) | 1, 2 months | T,P | R | Full-scale simulation of 7 day lunar landing mission | Very high | Simulator | 3 | E | Procedural errors; flight control error |
| Cotterman and Wood (1967) | 4,8,9,13 weeks | T,P | R | Full closed-loop control and landing of lunar excursion module | Very high | Simulator/ 6 weeks | 12 | E | Probability of successful performance with respect to criterion |
| Swink et al (1967) | 3, 5 months | T | S,O,D,R | Tracking line on oscilloscope | Low | Actual task/ to varying criterions | 120 | I | Integrated tracking error; temporal index |
| Youngling et al (1968) | 1, 3, 6.7 months | T | D,R,S | Closed-loop control of space vehicle flight simulator | Very high | Simulator/ 60 vs. 120 trials | 96 | I | Mean time on tracked target |
| Grimsley (1969) | 1 month | P | F | Setting switches and controls on panel; 92 steps | Medium | Actual task, varying fidelity/ to criterion | 60 | E | Number correct control settings; training time; time to retrain |

Table 1 (cont). Skill Retention Studies

| Investigator(s) | Retention | Task | Indp.Var | Task Description | Cmplxty | Training | #Ss | Exper | Dependent Var. |
|---|---|---|---|---|---|---|---|---|---|
| Sitterley and Berge (1972) | 1,2,3,4,6 months | T,P | R,RR | Closed-loop control of space vehicle simulator; lift-off to orbit insertion | Very high | Simulator; classroom/ ave. 34 flights | 45 | I | Flight control error measures |
| Sitterley et al (1972) | 4 months | T,P | RR | Closed-loop approach and landing of Shuttle simulator | Very high | Simulator; classroom/ to criterion (ave 10 hrs) | 15 | E | 32 Flight control error measures; procedural completion time |
| Sitterley (1974) | 4 months | T,P | F | Closed-loop approach, landing; Shuttle simulator; 7 min. flight | Very high | Self-study; Classroom; Simulator/ 8.1 hours | 5 | E | Integrated flight performance |
| Shields et al (1979) | 4, ,12 months | P | R | 20 common soldier tasks; e.g., donning gas mask | Very high | Army basic training | 523 | E | % go/no go task performance |
| Schendel and Hagman (1980) | 2 months | P | D,S | Disassembly and assembly of M60 machine gun | Medium | Actual task/ to criterion | 26 | E | Procedural errors |
| Johnson (1981) | 8-11 weeks | P | F,X | Setting controls for a simulated industrial process | Medium | Simulator; Photos/ to criterion | 60 | I | Procedural errors; performance time; device transfer |
| Goldberg et al (1981) | 5 weeks | P | D | Boresight and zero main gun of M60A1 tank, 27 steps | Medium | Actual task/ to criterion | 42 | E | Procedural errors |

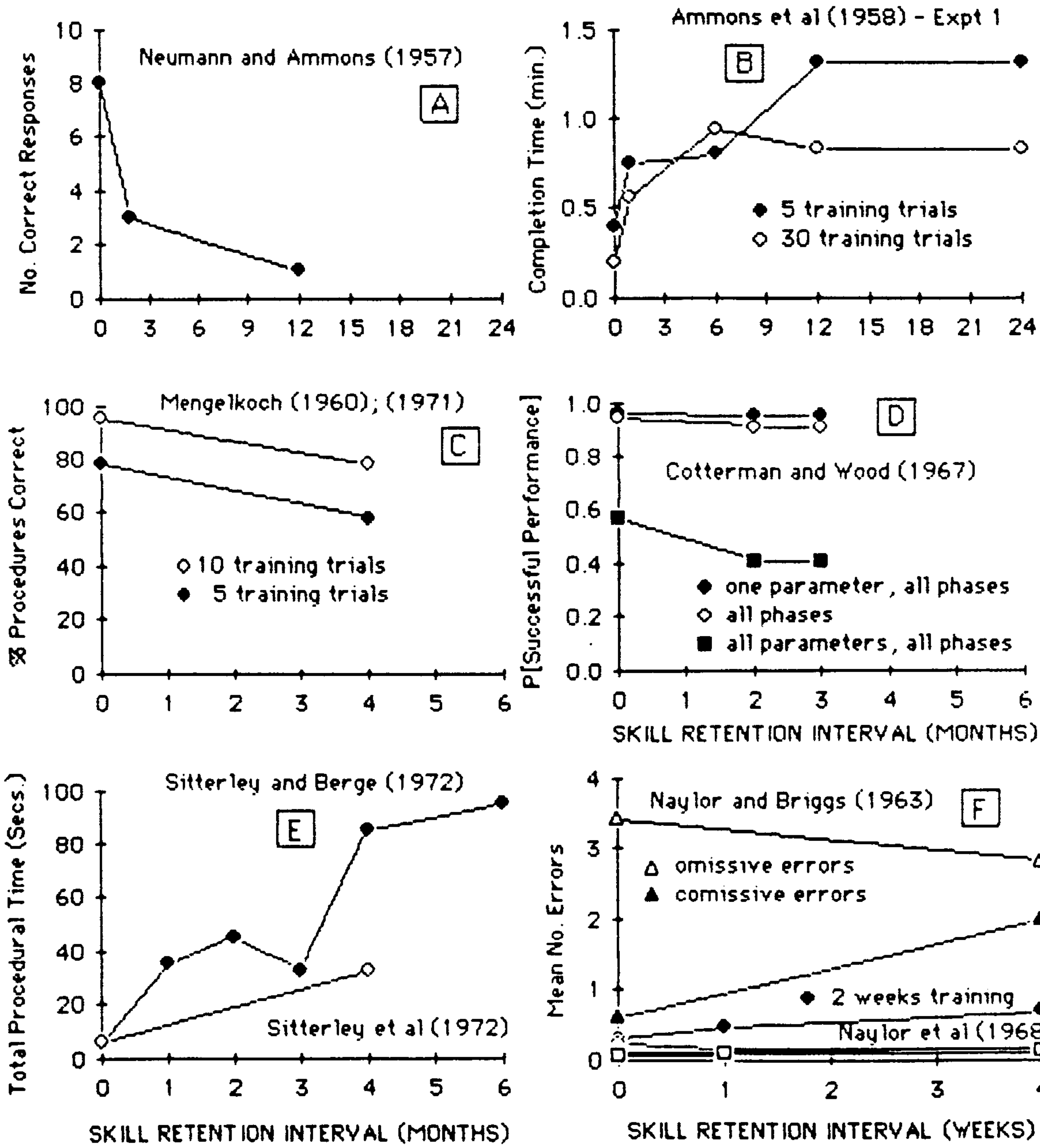


Figure 1. Procedural Skill Retention as a Function of Time Interval.
Note Varying Retention Interval Axes.

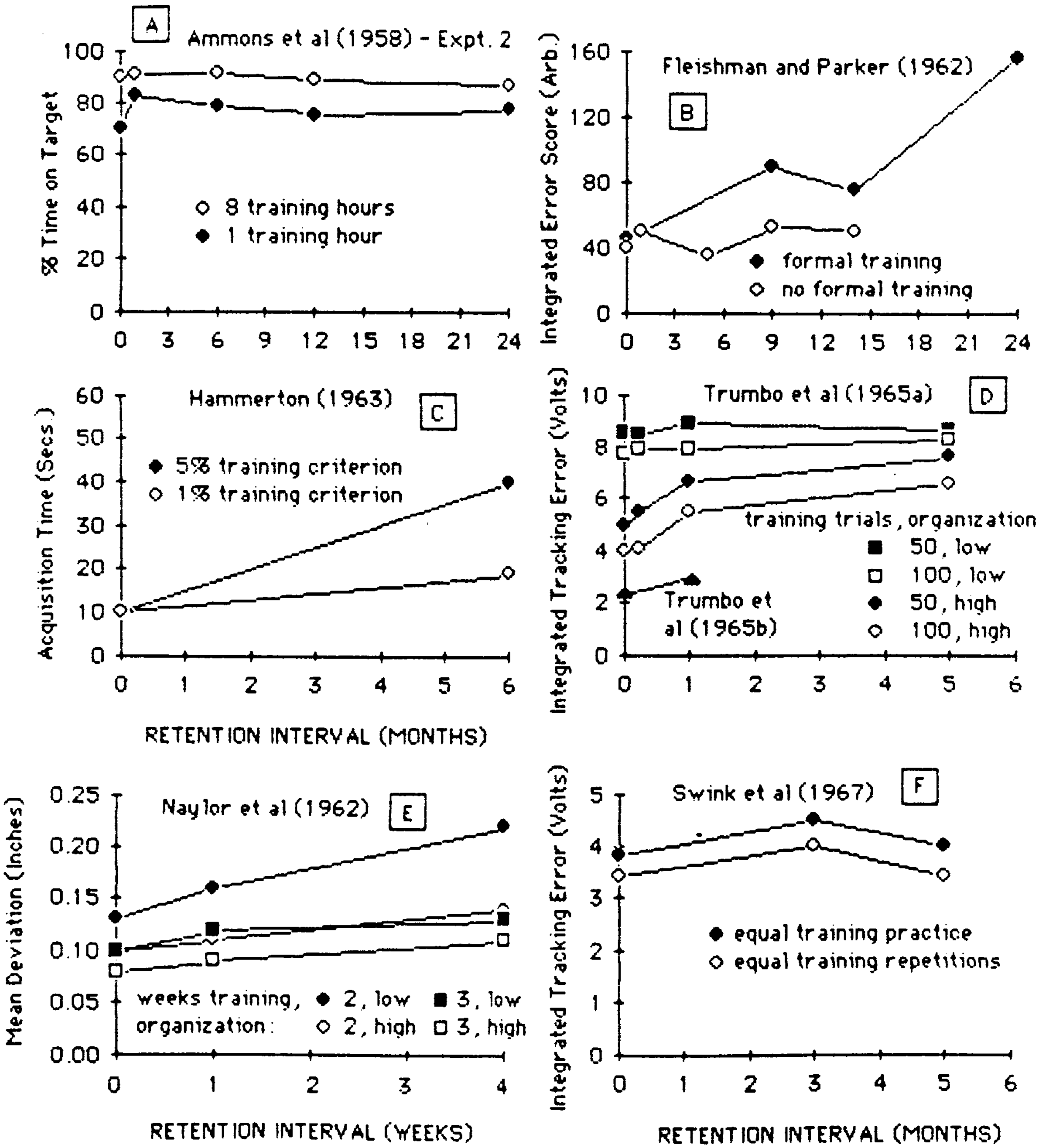


Figure 2. Simple Tracking Skill as a Function of Retention Interval. Note Varying Retention Time Axes.

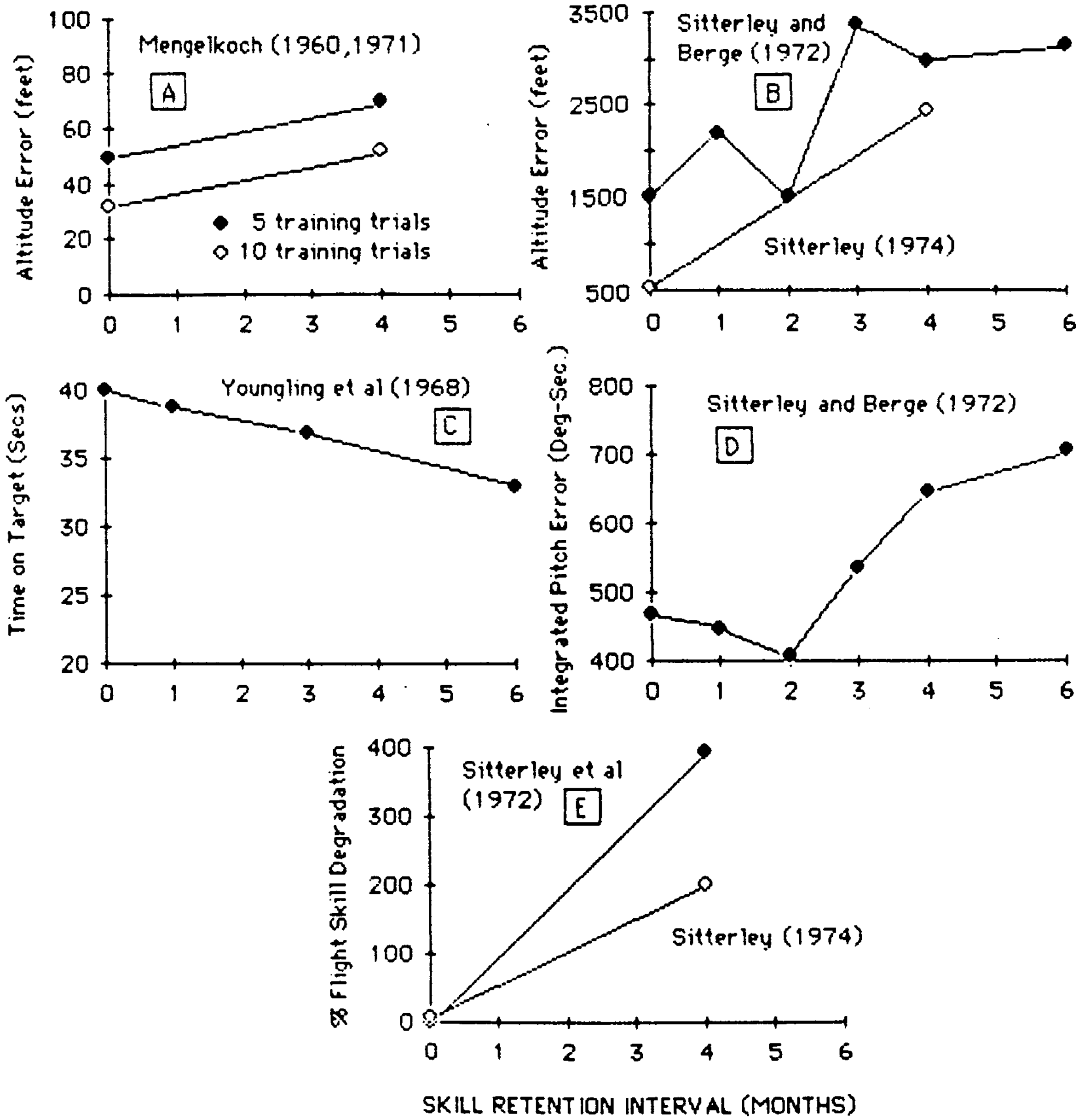


Figure 3. Complex Flight Control Skill as a Function of Retention Time Interval

## 2.1. TASK FACTORS

Those factors affecting skill retention that are direct properties of a task are considered in this section. Clear trends that appear in these factors are valuable in the design and evaluation of task training, as those factors responsible for a large amount of skill degradation are good candidates for potential elimination or control. Presence of a large number of these critical factors can point to tasks requiring frequent refresher training. Great care must be used when evaluating tasks containing more than one of these factors, as interacting factors have rarely been investigated in a controlled manner. Anything other than qualitative comparisons across studies are dangerous, due to countless numbers of uncontrolled factors. Rather, generalizations should be drawn by first noting within-study conclusions, then qualitatively comparing these across many studies.

### 2.1.1. Type of Task

Since some of the earliest skill retention research, a distinction has been made between overall types of tasks. Procedural tasks are those requiring discrete, ordered responses. Some have labelled these as cognitive tasks, referring to the large amount of non-automatic cognitive resource required when trying to recall a long sequence of task steps, while others have called them discrete tasks, for the isolated responses that are required. Examples here include checklists on aircraft or space vehicles, emergency procedures, and more abstract tasks such as setting sequences of switches in a proper order. The other class of tasks are those which require psychomotor skills for successful completion. Some have also labelled these as continuous control or tracking tasks, because discrete responses are not given. Typically, these require an individual to keep some stimulus on target, or within a specified range of conditions. Common examples here are driving a car, controlling an airplane or space vehicle, or simply manually controlling a lever so that a displayed shape remains between two points on an oscilloscope. Most real-world tasks contain an element of both of these. Early research efforts simplified these as much as possible to obtain a high degree of experimental control. These studies indicated that procedural skills degrade more quickly than operational or continuous skills. Only those studies which presented both types of tasks to subjects, measuring relative skill degradation differences, are appropriate in the present assessment.

In early studies, continuous tasks were simple tracking movements while procedural tasks consisted of sequence memorization. Ammons et al. (1958) tested over 1000 subjects on either a 17-step procedural task or a model airplane control task, over many retention intervals. Greater skill loss occurred on the procedural, cognitive task than on the motor control task over retention intervals of up to two years. Naylor et al. (1962) combined a procedural switch setting task (9 sets of 3 switches) with a three dimensional joystick-controlled tracking task and found similar skill retention for both tasks. In this instance, however, Gardlin and Sitterley (1972) noted that the procedural task was very simple, and greater skill degradation relative to the tracking task might have otherwise been expected. Using only a tracking task, Trumbo et al. (1965a) broke down overall performcance into both temporal and spatial accuracy dependent measures. The former measured anticipations and lag time, whereas the latter measured absolute positioning accuracy. Interestingly, the temporal skill performance was lost more quickly than the spatial performance over the one week to five month retention intervals. The better subjects may have emphasized the temporal aspects of the task more than the spatial aspects, suggesting that more effort should be spent on maintaining temporal task performance.

Later studies have used more complex, flight control tracking tasks with procedural checklists, measuring performance on both. Mengelkoch et al. (1971) combined a flight control task with a procedural checklist task containing 125 discrete items. Over a four month retention interval, discrete procedural responses were more susceptible to forgetting than the continuous flight control responses. Though the procedural losses were great enough to be practically significant, the investigators were careful to qualify this conclusion. It is not possible to define equal levels of learning between the two types of dependent measures; tracking is measured as a continuous variable (e.g., altitude error), while procedural lists are measured by percentage error. In this context of real aviation training, however, more emphasis should logically be placed on the learning and retention of procedural checklists than on continuous flight control (Gardlin and Sitterley, 1972).

Sitterley and others conducted a series of skill retention studies for NASA in the early 1970's. All tasks required both active flight control and the use of procedural checklists, and utilized a complex, closed-loop space shuttle cockpit simulator. Comparing manual control with emergency procedural skill retention from one to six months, Sitterley and Berge (1972) found that procedural skills degraded much more rapidly than operational skills. Flight control skills were acceptably retained for two months, whereas procedural performance degraded after only one month; flight performance degraded by a factor of ten after an interval of four months. Sitterley et al. (1972) found similar patterns of degradation, but the procedural skill loss was not as great as that found in the previous study. The studies differed in that the present study used experienced pilots (the previous did not) and allowed warm-up techniques prior to retention testing.

The studies cited above have all had in common the performance and measurement of both continuous, tracking and discrete, procedural skills. While differing retention intervals and task complexities were used, the consensus has been that procedural skills are (1) more quickly lost, and (2) lost to a lower relative skill level than tracking skills.

The underlying factors responsible for this differential retention loss are still unknown, however. Task organization differences may be responsible for this differential (see Section 2.1.2.). Alternatively, proportionally more training may be achieved by tracking than by procedural performance in a short period of time, because of its continuous nature, as discussed in Section 2.2.1. Also, there may naturally be more transfer and practice of tracking skills in a given retention interval than is allowed for procedural skills (see Sections 2.2.3. and 2.3.2.). The fact that procedural skills degrade more rapidly and fully than operational skills may thus be an emergent property of other underlying mechanisms. Different types of parameters are also measured in these two classes of tasks. Procedural skills are measured by accuracy in following the established order of a task, while operational skills are typically measured by temporal parameters. If procedural tasks are measured by such parameters as the time required to complete a sequence of switch settings, it in effect becomes an operational task, so the true measure of task type may lie in the parameters measured, not in the actual task itself.

### 2.1.2. Task Organization

The actual or perceived task organization, in addition to the procedural/operational dichotomy, influences skill degradation. Procedural tasks may have less spatial and temporal organization than tracking tasks (Gardlin and Sitterley, 1972; Trumbo et al., 1965; Swink et al., 1967; Noble et al., 1967). Unless a study manipulates an organizational variable in a highly controlled manner, the task type and organizational variable will be confounded.

Perceived organization has been most commonly manipulated by altering task predictability. Using a procedural task combined with a tracking task, Naylor et al. (1962) systematically manipulated task organization by illuminating light pairs in a predictable or an unpredictable order. The task organization had a greater influence for lesser trained conditions, in that more organization was required as less training was given. Trumbo et al. (1965a) created four differential conditions of target predictability in a tracking task by selecting tracking targets so that every one followed in spatial sequence, or every one, second, or third target was chosen randomly. At retention testing, subjects who received the most predictable target sequence retained skills better than all other levels of predictability. In fact, performance after a five month retention interval using the predictable sequence was superior when compared to performance after only one week of retention using the less predictable sequences. When compared with the predictable target organization, the less predictable tasks showed 80% to 100% more error, but little practical difference was noted between the low organizational task performances. Swink et al. (1967) had tracking targets either appear in a deterministic order on every trial, or with every fourth target randomly selected. Again, the predictable target sequence produced superior tracking performance for retention intervals of both 3 and 5 months. In another horizontal tracking task, Trumbo et al. (1967) produced three levels of target predictability, corresponding to every 4th target, 6th target, or no targets randomly assigned in a sequence of 12 target locations. All levels of predictability produced the same performance by the end of training sessions, but retention at the end of one week was greatest with the most predictable task. The lowest predictability task produced better performance than the medium predictability task by the end of the retention testing session, perhaps due to the fact that differential training had been given to these two groups of subjects; low predictability subjects received 195 training trials, while the medium predictability subjects received only 80 trials. Noble and Trumbo (1967) reviewed a number of experiments, breaking down retention loss by spatial and temporal uncertainty variables. In general, the greatest retention losses were noted in the most uncertain task conditions, and response strategies by subjects varied with the amount of task uncertainty.

Manipulations of procedural task uncertainty have also been used to alter organization. Naylor et al. (1968) manipulated the predictability of a secondary procedural task, while subjects were simultaneously performing a primary tracking task. Subjects had to depress buttons in varying orders, depending on which of several lights were illuminated at a point in time. Two levels of procedural task organization were defined: a light sequence in numerical order, and a sequence in random order. After retention intervals of 1 to 4 weeks, the well organized secondary task had lesser performance decrements than did the less organized task. In addition to decrements on the secondary task, performance on the primary tracking task was also influenced. The well organized secondary task produced superior retention on the primary tracking task after both 1 and 4 weeks, than did the more poorly organized secondary task.

These have clearly demonstrated that task organization directly influences both procedural and tracking skill retention duration. The actual task predictability was manipulated; retention may also be influenced by perceived organization within a task. While no studies have compared actual versus perceived predictability, this is precisely the function of training.

### 2.1.3. Task Workload

Workload refers to the physical and cognitive effort imposed by an operational task. Earlier work in this area studied the effects of time-sharing between several visual displays, as is required when driving or flying, while later work utilized other, joint combinations of abstract and operational tasks. For many reasons, some effort has been made to develop models and measures of workload in generic tasks (e.g., Brown, 1978). Rouse (1979) has reviewed information theory, control theory, and queueing theory models of operator workload, as well as performance, physiological, and subjective workload measures. Workload is of interest due to its effect upon skill training and retention. Tasks requiring many simultaneous control movements, as in aircraft control, or tasks containing long strings of branching point decisions, can both be considered to be of high complexity and required workload. Johnson (1981) suggested that the number of steps required in a procedural task should be a determining factor in its probability of successful skill retention.

It is interesting to ask (1) whether skills for a high workload task are retained for a different time period than those for a lower workload task, and (2) whether an individual can be trained to perform a given task under higher workload conditions. The organization of a task may be considered to directly affect workload, in that the two are inversly related (see Section 2.1.2.).

The most valid method of increasing workload has been to add a secondary task on top of a measured primary task. Garvey (1960) trained subjects for 25 days on a tracking task, then added a different secondary task on three subsequent days. Inclusion of the secondary tasks greatly increased tracking error to levels above initial, unpracticed levels. Single-task, low workload training did not transfer to dual-task, higher workload tasks. Briggs and Wiener (1966) noted that higher fidelity training is required in high workload, dual-task performance, than in lower workloads. This result was generalized to flight control simulators. Rudimentary flight control, having low time-sharing requirements, may be trained on low fidelity devices, but greater workload requires a higher-fidelity simulator. Trumbo et al. (1967) combined a tracking task with a verbal number anticipation task, of varying difficulties. Addition of the secondary task again dropped tracking performance to below that at the start of training. Performance after retraining did not increase to the level shown by those not performing secondary tasks. Further, performance loss after 8 days of retention was independent of the introduction of the secondary task. Naylor et al's (1962, 1968) subjects performed in dual-task combinations of a tracking task with a switch-setting procedural task, with predictability in the procedural task manipulated. This also influenced workload, because much more attention had to be placed on procedural task performance in conditions of low predictability. For each of the two tasks, the low predictability procedural conditions produced both poorer absolute performance and poor retention after a 4 week retention interval. The amount of training was the greatest predictor of absolute performance level. Gopher and North (1977) combined a primary tracking task with a secondary digit-processing task, and manipulated training conditions. Greater performance improvements from training occurred under dual-task than under single task conditions, as if the motivation from a harder task was beneficial to learning.

The definition and measurement of mental workload is a young science, but some investigators have clearly implicated it as a factor in training and retention. Higher workload tasks, defined by large levels of cognitive time-sharing, are harder to learn and retain. Large performance improvements, due to training, have been noted, however. Relative skill retention duration between single and dual-task conditions also remains to be modeled.

### 2.1.4. Performance Measurement

The methodology used to assess one's performance in a task has been a continuing issue for many years. In fact, an entire issue of a journal was recently devoted to this topic (Human Factors, Volume 21(2), 1979). Measurement of procedural and operational skill following a no-practice retention interval has relied on measures of accuracy or speed. These dependent variables are then plotted as a function of time, across several experimental subject groups. Bahrick (1964) critiqued the basic skill retention curve, suggesting that observed changes are due to varying sensitivities of an overserver's perceptual/cognitive system between testing sessions, and not necessarily due to forgetting or retention differences. Converting degradation scores into Z-scores may aid in stabilizing the variance between testing sessios (Bahrick, 1965).

Single-task performance measurement may not capture the concurrent demand, time-sharing requirements of real work environments, and many have artificially combined many tasks together into multiple task batteries (e.g., Alluisi, 1967). These task batteries provide high validity, precision in measurement, and sample a broad range of abilities (Akins, 1979). However, some have argued that the batteries are unnecessarily artificial, and performance scores may be defined rather arbitrarily (Chiles, 1967; Akins, 1979).

Swezey (1979) introduced a Bayesian-oriented utilities model to determine what criterion level should be achieved at the end of training sessions for gunnery trainees. This called for a 10-step decision model, identifying components of the model and calculating utility. Other, empirically-based methods of assignment to training programs have also been presented (e.g., Savage et al., 1982).

The appropriate choice of useful performance variables and methodologies is still very much at issue, particularly in light of the fact that the degree of observed retention is dependent on which perameters are used. Some investigators have measured absolute performance, whereas others used difference scores, subtracting post-training performance from retention performance. Both types of scores are required to evaluate loss during a no-practice retention interval (Gardlin and Sitterley, 1972). Also, variance measures, in addition to means, have rarely been used as performance indicators.

## 2.2. TRAINING FACTORS

This section covers those factors having their primary influence on initial task training. Many have suggested that these variables are at least as influential as task variables in determining the duration of skill retention. Factors covered here include the duration of training, the distribution of initial training sessions, transfer of training, and fidelity issues. The central questions to consider concern the required amount of training (cost) to expect adequate skill retention for a desired interval (benefit). Such issues as whether whole or part-task training are needed, and the degree to which already trained skills can be transferred, are relevant in answering this question.

2.2.1. Amount of Initial Training

A consistent finding across many skill retention studies has been that the relative amount of initial training one receives is a strong predictor of level and duration of skill retention. Overtraining on one task may be considered insufficient training on a similar, analogous task. Also, overtraining can be very expensive if it requires significant high-fidelity simulator time. Due to the importance of this issue, reference is made to the figures presented earlier.

Using only a 15-step procedural task, Ammons et al. (1958, Experiment 1) trained subjects for either 5 or 30 training trials. Initial completion time for the two subject groups at the end of training was 0.4 and 0.2 minutes for the 5 and 30 trial groups, respectively. After a 2 year no-practice retention interval, the 5 trial training group performed the task in 1.3 minutes (a 3-fold increase) and the 30 trial group performance rose to 0.5 minute (2.5-fold increase). Proportionally fewer trials were required for the 5 trial subjects to regain their initial performance than were required by the 30 trial subjects. This training difference is plotted in Figure 1(B). In their tracking task (Ammons et al., 1958, Experiment 2), subjects were trained in aircraft control (using an airplane model) for a period of either 1 or 8 hours. Results from this study, plotted in Figure 2(A), showed that skills increased somewhat over retention intervals of up to 2 years. This increment was approximately equal for both training durations, but the 8 hour group maintained about a 2%-10% superiority over the 1 hour group performance throughout all retention intervals. While the superiority of longer training remained clear in this task, the reason for the performance increment did not.

Mengelkoch et al. (1960, 1971) trained inexperienced subjects for either 5 or 10 daily 50-minute sessions, in an aircraft flight simulator. As shown in Figure 1(C), the two groups had approximately the same skill degradation, after a 4 month retention interval, on a procedural task (losses for the 5 and 10 trial groups were, respectively, 20% and 16% of training levels). The effect of greater training was in achieving a nearly 20% increase in initial training level on the procedural task. The flight control or tracking portions of Mengelkoch's study only showed significant skill degradation, from both training groups, for the airspeed error parameter. The 5 and 10 trial training groups showed altitude error increases of about 10 feet over the 4 month interval (see Figure 3(A)), or about 20%-30% increase from initially trained levels. This loss was significant for the 5 trial group, but not the 10 trial group. Like performance of the procedural task, the primary difference between training duration groups was in the performance level at the end of training, rather than the relatively long skill retention. The fact that the 5 trial group retained their skills to the same magnitude as the longer trained group is meaningful.

Naylor et al. (1962, 1968) trained subjects on a dual tracking and procedural task for either 2 or 3 weeks of daily sessions. The longer training produced relatively superior performance at the end of both 1 and 4 week retention intervals when compared with end of training scores, but only in omissive errors. The comissive errrors here did not significantly differ. Figure 1(F) shows the omissive errors from these subjects. Naylor et al. (1963) used the same task and trained subjects for either 5 or 10 daily sessions (one or two weeks). Only comissive errors here differed as a function of training duration. The tracking performance from these dual-task studies showed similar trends. Three weeks of training produced less integrated tracking error and greater skill retention than two weeks of training. Integrated error was also significantly lower for 2 weeks of training, compared with 1 week of training. The 1 week trained group did, however, display increases in skill during the retention interval, not shown by the 2 week group.

Using a simple abstract line tracking task, Hammerton (1963) varied desired initial training criterions, as opposed to varying initial training durations. A 5% criterion group required 3 successive daily elapsed target acquisition times that did not differ at the 5% level of significance. Likewise, those in the 1% criterion group had 3 successive scores not differing at the 1% level. Retention of tracking skill after 6 months is shown in Figure 2(C). While two groups did not differ in mean time after training, the 5% group required more than 10 additional seconds than the 1% group after 6 months. This difference was both statistically and practically significant. Even the 1% group exhibited significant skill degradation, in spite of their extensive training. In this study, the additional training to achieve the 1% criterion was 9 to 17 days beyond the 8 to 22 days required for the 5% criterion. This degree of overlearning significantly decreased, but did not entirely alleviate, skill degradation.

Trumbo et al. (1965a) presented a similar line tracking task to 250 subjects. Half were trained for 50 trials and half for 100, over a 3 day interval. As shown in Figure 2(D), both training groups showed significant retention losses (increased tracking error) over a 5 month interval. The task organization was a stronger predictor of retention loss than was the absolute training duration. The high training level group did exhibit less skill loss than the low training group at all tested retention intervals. A subsequent analysis of separate skill retention components, from only the 100 training trial group, demonstrated that the best tracking subjects retained temporal accuracy (as measured by lead or lag time) better than spatial accuracy (as measured by percentage of over or undershoot errors). Thus, temporal as opposed to spatial training may be more important in retaining tracking skills over a long duration.

In a complex simulation of an Apollo mission, Youngling et al. (1968) trained their subjects for either 60 or 120 days. The overall skill retention, measured by time on target, was twice as great for the 120 trial group (5.5 seconds) than the 60 trial group (2.4 seconds on target).

Hagman (1983) summarized several skill retention studies performed in military contexts. Hagman (1980) varied the number of times Army personnel repeated a procedural electrical alternator output test during training. Increased task repetition, from 1 to 4 times, reduced performance time and errors by approximately constant amounts during training and after a two week retention. Increasing repetitions linearly increased performance until the 4 repetition duration training. Schendel and Hagman (1980) trained Army groups to either one correct performance or two correct performances in the disassembly and assembly of an M60 machine gun. After an 8 week retention interval, the greater trained group committed fewer errors than the lesser trained group. Goldberg et al. (1981) trained Army personnel to either 1 or 3 successive correct performances of boresighting and zeroing the main gun of an M60A1 tank. Again, higher performance after a 5 week retention was achieved by the more highly trained personnel. Schendel and Hagman also varied the time at which extra training was actually given. One group of subjects received extra task repetitions during the intitial training, while a second group received theirs half-way during the retention interval, at 4 weeks. They found no significant difference between the two modes, implying that it is more cost effective to supply all training at one time.

Most of the above studies have been in agreement in that skill retention is a function of training duration. Many questions still remain, such as whether training duration is more or less important than task organization or the retention interval in determining the magnitude of skill retention.

2.2.2. Training Distribution

The way in which a given amount of training is distributed over a time interval is also predictive of skill retention success. Fleishman and Parker (1962) manipulated the method of retraining following a no-practice retention interval. A massed practice group received 4 practice sessions within a 2 hour period, while a distributed practice group received the same level of training, spread across 4 subsequent days. The distributed training group outperformed the massed practice group by the end of retraining, but both groups performed equally well after another 1 week retention interval. Thus, the distributed practice may have had its effect on temporary performance factors. Hagman and Rose (1983) reported that insertion of time between repetitions of a task increases skill retention, but the problems associated with the disruptive training may overshadow their benefits in actual tasks. Hagman (1980) compared massed versus spaced training for Army electrical alternator testers and repairers. The massed training group took 51% longer and made 40% more errors than the spaced group. Schendel and Hagman (1980) either gave task repetitions as part of initial training or one month later, and found no difference in ability after a two month retention interval. Spacing of task trials and/or sessions may be helpful, but there is some question as to whether its effectiveness varies with task proficiency level (Hagman and Rose, 1983). A model is needed here, and must consider the initial level of post-training skill proficiency, which also determines the required frequency of task repetitions.

2.2.3. Transfer of Training

Training transfer refers to the ability of a trained skill to generalize to a new setting. From cost considerations, positive skill transfer means that performance on a task can utilize already trained skills, saving time and money. Also, highly generalizable skills can easily be used in new settings or situations, for which no training has been constructed. The term validity refers to the degree to which training readies one for performance on a task, and is a measure of training transfer.

Briggs and Wiener (1966) trained subjects to perform an abstract two dimensional tracking task, and transferred this training to an easier task requiring the setting of a control knob. High fidelity training (achieved through proprioceptive control feedback) was only required when the transfer task reqiuired a high level of time sharing, by forcing constant positioning. Thus, when proprioceptive cues and high levels of time sharing are required in a task, the training program should be of high fidelity.

Reid (1975) assessed training transfer from a formation flight simulator to actual formation flying. Untrained, formation simulator trained, and aircraft trained pilots were compared in actual flight formation flying. Evidence of positive simulator skill transfer was obtained, as these pilots did not fly significantly differently from conventionally trained pilots. The simulator provided the same degree of training as the flight sorites, indicating a high level of skill transfer. Carter and Trollip (1980) showed that training transfer between skills may be compared by plotting iso-transfer curves between pairs of skills and noting maximum transfer pairings. An operations research technique, the Lagrange Multiplier, was useful for determining costs and benefits of training.

Validity of training can also be evaluated by the method proposed by Goldstein (1978), who used a four level approach to evaluation: (1) Training validity, determined by trainee performance relative to standard training criteria, (2) Performance validity, measured by transfer of job performance, using criteria from the actual job, (3) Intra-organizational validity, measured by the performance of a group of new trainees based on the performmance of a previous group, and (4) Inter-organizational validity, measured by the degree to which a training program validated in one organization can be used in another organization. All of these levels must be evaluated to determine the effectiveness or validity of a given training program. Moving from the first level to the fourth, an increasing number of variables influence the success of training. Also, the necessary level of complexity in a training needs analysis must depend on the final goal of training. If one's goals do not reach beyond the second level, for instance, there is no need to consider levels 3 or 4. Such a structural assessment of validity is required to transform training needs assessment from art to engineering. The validity of training apparatus, according to Crawford and Crawford (1978), lies more in the manner in which it is used, rather than in the degree of its similarity to actual equipment. These investigators substituted conventional hands-on practice for part-task computer-based training on the use of an integrated control panel in an anti-submarine airplane. Control subjects performed on a high fidelity simulation of the control panel, while experimental subjects were trained using a graphic simulation on a touch screen display. The experimental subjects completed more tests, in less time than the control subjects. The computer-based training was found to provide at least as good skill acquisition, in less time and at lesser cost, than the full simulator training. A cost analysis indicated a substantial two-thirds cost savings over the conventional training method, much of which was due to a smaller number of instructor man-hours.

Adams (1979) contrasted the shortcomings of two methods of rating flight simulators for aircrew training. A transfer of training study measures the relationship between achieved task competence and proficiency on the flight simulator, while the rating method requires an engineering and experienced pilot assessment of hardware and flight similarity between the simulator and actual aircraft. Adams reviewed many studies with the thesis that both techniques are flawed. This does not mean that simulators are not useful, though. Humans require the perceptual-learning, stimulus-response learning, and feedback provided by simulator sessions. In addition, simulators successfully motivate trainees better than lower fidelity learning environments. Because simulators are based on these well-founded principles, simulators need not be evaluated for their effectiveness; this may be taken on faith (Adams, 1979).

As part of his procedural control setting training study, Johnson (1981) measured skill transfer by manipulating the sequential steps of the original task. In two experiments, low-fidelity paper-and-pencil training transferred very well to the new operational task. Although the two tasks likely utilized similar skills, this was further evidence of the utility of analogous tasks for training purposes. Validity determination and training needs assessment are still very much debated topics; this section has only briefly introduced these issues.

### 2.2.4. Training Fidelity

A large and detailed lieterature has developed, concerning required fidelity in task training. At issue here is whether full fidelity is required to achieve adequate and retained skill performance. Considerations such as whether open or closed-loop simulation control is required, whether simulators must necessarily move, or whether adequate whole task performance can be achieved from part-task training have been addressed. In this paper, fidelity refers to the degree to which a training device can mimic an actual task of interest, such as flying an aircraft.

Naylor et al. (1963) manipulated the type of rehearsal subjects received in a dual tracking and procedural task. The procedural task consisted of 9 pairs of lights to which responses had to be made, and the tracking task was a three dimensional meter nulling task, in which roll, pitch, and yaw were simulated. In whole-task training, subjects practiced with both tasks simultaneously, as required for the measured performance task. Part-task training required separate practice on each task. Retention differences between the training conditions were significant on tracking performance, with the part-task rehearsal group displaying inferior performance. Whole-task rehearsal was also superior for procedural task performance, but this superiority lessened over the retention testing. Whole-task rehearsal was superior with a small amount of training (up to 5 days), but after 10 days of training, the two types of rehearsal were not signifianctly different. Naylor and Briggs (1963) manipulated rehearsal conditions on this procedural switch setting task. Whole-task rehearsal consisted of repeating the original task half-way through the 2 month retention interval. Part-task rehearsal conditions consisted of either (1) spatial rehearsal, with stimulus events occurring at equal temporal intervals, or (2) temporal rehearsal, with stimulus events occurring at varying times as in the original task, and stimuli appearing in a regular spatial order. The whole-task group produced far fewer omissive procedural errors than the part-task groups upon initial retention testing. The whole-task and part, temporal-task rehearsal were superior to spatial-task rehearsal when considering comissive errors. Thus, whole-task rehearsal here was best, closely followed by part-task temporal rehearsal in upholding skill retention over a 1 month retention interval.

Fleishman (1965) presented a multidimensional tracking task to inexperienced Air Force trainees, with the objective of predicting whole-task performance from various combinations of part-task training. The performance measurement device contained display dials for heading, altitude, bank, and airspeed, which all had to be simultaneously centered. Subjects were first proficiency trained on one dial, then two dials, then the entire task. Correlations between one-dial, part-task performance and whole-task performance ranged from .46 to .54 across the subjects. Between two-dial, part-task performance and whole-task performance, the range was .63 to .70. Multiple component practice was a better predictor of whole-task performance than single task performance in this multidimensional task. In addition, the multiple component performance was at least as predictive as linear combinations of the single task performances. The greatest correlations (.74) between part-task and whole task performance were found with linear combinations of two, two-dial practice. In this work, the actual task components that were used was less important than the fact that simultaneous practice had occurred. All predictive tasks here were part-task practice, but this investigation suggested that a continuum exists in training effectiveness, between various integrative or combinatory levels of sub-task performances.

In his space vehicle approach and landing simulator, Sitterley (1974) varied the fidelity of pilot retraining methods, following a 4 month no-practice interval. The number of visual cues present in the training session strongly predicted the level of performance achieved, and the level of visual cueing was independent of the fidelity of the simulator session. Static photographic training was superior to open-loop, dynamic training. All cues, however, present in the static pictorial method were present in the dynamic display training, so the total number of cues was not soley predictive of training effectiveness. As stated by Sitterley, the most important element in these training alternatives was the presentation of efficient cues which assisted pilots in recalling their basic flight experiences. Thus, open-loop, static training methods may actually be superior to more costly methods, given careful training program design. Trollip (1979) compared a computer-based with a simulator-based training program for aircraft flight control. Control subjects were trained in a flight simulator, while experimental subjects were trained on a plasma touch screen terminal with an attached hand controller. The computer-based training produced significantly fewer critical errors and better flight control than the simulator. This trend was identical in both no wind and crosswind flight patterns. When generalizaing flight control to a new procedure, no difference was found between the two methods of training. The computer-trained subjects performed better, learned quicker, and made fewer mistakes than their control counterparts. It allowed students to develop better mental images of the ideal flight characteristics. Johnson (1981) also found that training requiring large use of mental imagery cues can produce the highest level of skill retention. Even in high complexity flight simulation and control environments, the highest level of fidelity is not required. Sitterley and Berge (1972) and Sitterley (1974) concluded that static rehearsal or training may be superior to the dynamic, higher fidelity rehearsal because of the artificially increased importance of visual cues.

One variable significantly influencing fidelity in aircraft simulations is flight motion. This has been a controversial topic over the past decade, with many insisting that motion cues are unnecessary for general aviator training. Caro (1979) discussed this issue with reference to two different motion cues, maneuver motion and disturbance motion. The former motion cue refers to those motion changes initiated by the pilot, whereas the latter refers to those cues initiated outside the immediate control loop, such as turbulance or engine effects. While maneuver motion moves the aircraft platform, it does not cause important alerting cues provided by the disturbance cues, which lead to quicker and more accurate pilot control of the simulator. No motion, on the other hand, is required if the simulated aircraft is easy to control and relatively stable (Caro, 1979). Thus, required fidelity here was based on a logical analysis of task training requirements.

In the monitoring and control of a procedural industrial operation, Johnson (1981) utilitzed three different training strategies: (1) conventional, full-fidelity practice on the actual task, (2) medium-fidelity reproduction study of photographs, where the subject was allowed to draw on the photos, indicating his procedural responses, and (3) low-fidelity, blind practice, where the subject was allowed to study, but not write on photos of the control equipment. Although the conventional strategy provided the quickest learning time (the blind practice required 1.5 times as long to reach criterion), the conventional and reproduction training did not produce different control setting errors after a 3 month retention interval. This illustrated that the highest fidelity training is not required in procedural tasks. Johnson and Rouse (1982) also found that low and medium fidelity training in aircraft power plant troubleshooting is very competitive to high fidelity simulation. The highest-fidelity method in this study included training on the actual task, medium fidelity required a special power plant simulation, while the lowest fidelity condition utilized videotaped lectures and live quizzes. Video training produced the greatest performance with all simulated failures, and the original task and simulation were similar in their effectiveness. From a cost consideration, the low and moderate fidelity devices provided sufficient problem solving experience to effectively compete with the conventional training methods.

According to guidelines posed by Cream et al. (1978), the specification of required training fidelity appears to be art, rather than engineering. They stated that (1) essential and nonessential aspects of controls and displays must be differentiated, in that many of these elements are not required for proper training. (2) The choice of fidelity is more complicated when dealing with displays rather than simple indicators. Also, no rigorous decision-making procedures have been developed in the area of cost/benefit fidelity analyses. Though experiential-based fidelity definintion has been used for many years, no useful guidelines exist for the development of training for new tasks.

Perhaps the degree of required fidelity is a function of how little is understood of the processes required to carry out a given task. Seemingly complex tasks may only be combinations of a limited number of combined operations. On these, perhaps part-task training would be sufficient, if the actual components could be identified. Full-fidelity, whole-task training would then only be required in very complex tasks.

### 2.2.5. Adaptive Training

Both ground-based and on-orbit training systems should be adaptive to trainee performance, for maximum efficiency. This research area has recently shown substantial growth, as a result of the development of specific adaptive systems. Machine-controlled adaptive training simply automates a skilled instructor, by modifying the training stimuli as a function of trainee performance. Training efficiency is maximized, because effective learning only takes place when training is at an appropriate level of difficulty (Kelly, 1969). Adaptive learning curves typically show a linear relationship between ability and time, as opposed to conventional training curves.

The marker variable for training adaptation may vary, depending on the nature of a task. Johnson and Haygood (1984) utilized performance on a secondary light recognition task to adapt the difficulty of a primary tracking task. Williges and Williges (1978) concluded that the most effective adaptive parameter should be a multivariate combination of several performance skills. Matheny (1969) argued that the time lag between a system response and an operator's subsequent performance should serve as the adaptive parameter in general man-machine systems. While many parameters have been used, they have all served the function of varying the difficulty of a primary task.

## 2.3. RETENTION INTERVAL FACTORS

This section discusses those skill degradation issues directly related to the retention interval, between the initial training and actual performance. Two factors here have important implications for skill retention. (1) The duration of the retention interval has been extensively studied, using intervals from a few minutes to well over two years. (2) The nature of activities performed during the retention interval influences skill degradation, just as practice of any task should do. Most investigators in this area have concluded that learned skills regularly degrade with increasing no-practice retention intervals. Bahrick (1964) has, however, disputed this concept, claiming that retention curves based on anticipation, recognition, or free recall reflect changes in one's sensitivity from session to session. This complaint, however, was only drawn against measures with only right/wrong responses. Many procedural studies, using other continuous measures such as completion time, have indeed found evidence for regular skill degradation over time.

### 2.3.1. Length of Interval

Retention intervals are an important consideration when designing refresher training on lengthy crew missions. The many skill retention/ degradation studies have all shown large performance decrements upon post-retention testing, when no retention practice is allowed. Estimates of the percentage skill loss at various time intervals allow an empirically determined estimate for the frequency of refresher training, given that retraining is required when performance falls below a set criterion. Longer intervals are generally accompanied by greater loss in skills, but this is very task specific. Ideally, this review should provide an overall skill retention function, mapping percentage of skills retained versus retention duration. However, reality dictates that between-study variations make such generalizations and models very hard to achieve. The most important question that can be answered here is whether a constant degradation across many tasks is found, with all other factors being equal. A cursory analysis of the results in Figures 1 through 3 indicates that skills degrade with time when not subjected to interim practice, and that the level of degradation reaches as asymptote in some studies.

Procedural Tasks. With few exceptions, procedural performance is marked by consistently increasing decrements with progressively longer retention intervals. Neumann and Ammons (1957) found that a one year, 90% loss in post-training performance was about the same as initial performance at the start of training, but proficiency was quickly regained upon retraining (Figure 1A). On a task of nearly equal complexity, Ammons et al. (1958, Expt. 1) found a 2 to 3-fold increase in task completion time after a one year interval, which did not appreciably increase after 2 years of retention (Figure 1B). The magnitude of relative skill degradation was the same here, regardless of the original number of training trials. Mengelkoch (1960; 1971) also found that relative mognitude of skill loss was independent of the amount of training (Figure 1C), where subjects showed a 20% decrease in correct procedures after a 4 month retention. In an extremely complicated 169 hour mission simulation, Cotterman and Wood (1967) found relatively small degradation over a 3 month retention when only a single parameter was considered (Figure 1D). The probability of successful performance over the interval fell by about .03. However, when all parameters in all phases of the simulation were considered, the probabilities dropped significantly over the interval; initially at an average of about 0.6, it fell to about 0.4 after the retention interval, suggesting that a failure was highly likely in some mission phases. This study was flawed, however, due to uncontrolled retention interval activities and small sample sizes (Gardlin and Sitterley, 1972). The performance of complex control and emergency procedures clearly degrade in required procedural time after 6 months, and Sitterley et al. (1972) noted a 4.5 fold increase after 4 months of retention (Figure 1E). Johnson (1981) measured the time required to set controls in an 87-step procedural task, and found a mean time of 8 minutes after training had increased by 50% to 12.8 monutes after about 2.5 months. As gross estimates of the magnitude of skill degradation in procedural tasks, 20% to 50% degradation in 3 to 6 months, and 50% to 100% or more in more than 6 months may be made, based on the above data.

Investigations over shorter retention intervals, up to 1 month, have not found consistent skill degradation patterns. For example, using their switch setting task, Naylor and Briggs (1963) found a 20% decrease in omissive errors, but a 233% increase in comissive errors after 1 month (Figure 1F). Likewise, using the same task paired with a tracking task, Naylor et al. (1968) found the only retention degradation in comissive errors from the subject group with lesser training and low task organization (Figure 1F). Thus, skill retention of less than one month is harder to predict than longer durations, and may be dependent on many other task factors.

Simple Tracking Tasks. Performance on tracking tasks have not as a rule shown the predicatble and regular retention decrements shown by procedural task performance. In controlling the flight characteristics of a model airplane, Ammons et al. (1958, Expt. 2) found only a small 5% decrease in tracking time on target, between 1 and 24 months (Figure 2A). These slight skill decrements followed slight but significant increments between the end of training and 1 month retention. Beyond the absolute performance difference at training, the duration of training did not alter the relative decay rate of skills. Hammerton (1963) used a measure of target acquisition time in a tracking task (Figure 2C). By varying the allowable amount of session-to-session variability at the end of training, differential skill degradation was observed at a retention interval of six months. The looser criterion group showed a 3-fold increase in target acquisition time, whereas the tighter criterion showed only about a 1-fold increase. Thus, in tracking, regularity of performance as well as absolute magnitude appears to predict the degree of skill degradation. Over a short, 1 month retention, Naylor et al. (1962; 1968) demonstrated statistically significant skill loss at 2 levels of training duration and two levels of task organization (Figure 2E). Relative losses averaged about 16% at one week, and 44% at one month.

Four studies cited here used integrated tracking error as a dependent measure. Fleishman and Parker (1962) had two groups of subjects perform tracking tasks. A group receiving no formal training showed no performance decrement at up to 14 months of retention (Figure 2B). A second group who received formal training on the task showed a 1-fold increase in error after one year, but then showed a 4-fold increase after 2 years. Trumbo et al's (1965a) subjects showed virtually no performance decrement with intervals up to 5 months, when the task was unpredictable, with random targets located on every trial (Figure 2D). However, when the target position was more predictable, post-training tracking error was about 50% less than in the predictable condition. Retention intervals of 1 and 5 months produced large decrements in performance, upwards of 50%-60% from training levels. Equal degradation rates were found for both 50 and 100 training trial conditions, with the latter condition always producing better performance. Trumbo et al. (1965b) also demonstrated a 24% skill loss after a 1 month retention interval (Figure 2D). Swink et al. (1967) also manipulated task predictability and training duration in a tracking task (Figure 2F). The retention interval in this study, however, was unrelated to tracking error, as tracking ability did not degrade over a 5 month no-practice interval. Roehrig (1964) had several subjects stand on a small balancing platform, and measured the time duration that they could balance to within ± 1.5° of horizontal. After a 50 week hiatus from the task, all subjects demonstrated performance at least as great as shown at the end of training. Much like the well-known fact that "one never forgets how to ride a bicycle," this task, once trained, seemed to trigger the same skill retention. Perhaps there is much ability transfer from balancing in ordinary walking (once the body has been trained to use the muscle groups required by the task), and subjects were unknowingly practicing the task. This conversly suggests that we can forget how to ride a bicycle, if balancing is not normally practiced, as in bed-ridden or spacefaring individuals. Research needs to first be conducted to determine which tasks are dependent on balancing practice in a gravity environment, as was implied by an earlier report (National Academy of Sciences, 1972, p. 245). A taxonomy of tasks, organized by gravity dependency, should be developed.

Simple tracking performance, not requiring a large number of simultaneous decisions and elements of concious cognitive control, does not appear to degrade as requlariy and predictably as procedural skill performance. While some studies did find large decrements after a few months (e.g., Fleishman and Parker, 1962; Hammerton, 1963), others have found no evidence of skill degradation (e.g., Ammons et al., 1958; Swink et al., 1967). Clearly, in simple tracking, other factors are important in determining the degree of retention loss. From those studies cited here, those factors must include duration of initial training and task organization or predictability.

Complex Tracking Tasks. In those few studies using tracking tasks in higher complexity flight control contexts, performance on at least one parameter has shown strong interval-related degradation. Mengelkoch (1971) found significant increases in altitude error after a 4 month retention interval (Figure 3A). The skill degradation rate was equal between the 5 and 10 training trial groups, but the 10 trial group consistently made about 20% less error than the 5 trial group. Sitterley and colleagues also used altitude error, among many other parameters, in their space vehicle simulations. Sitterley and Berge (1972) measured a 2-fold increase in error over a 6 month interval, whereas Sitterley (1974) found a 5-fold increase over a 4 month retention interval (Figure 3B). In an alternate parameter, Sitterley and Berge measured a 55% increase in integrated pitch error after a no-practice retention of 6 months duration (Figure 3D). When measuring ability to null complex movements in the display within a simulator, Youngling et al. (1968) found a nearly linear relationship between the length of retention and performance loss (Figure 3C). Here, total tracking time on target decreased from approximately 40 seconds at training to about 33 seconds after 6 months, or a 20% loss. Percent flight skill degradation, a composite of many flight parameters, is perhaps the best overall measure of flight performance. Sitterley et al. (1972) noted a 400% decrease while Sitterley (1974) noted a 200% decrease in skills over a 4 month interval (Figure 3E). Clearly, flight skills are very sensitive to no-practice retention intervals, and may degrade by 4 or 5-fold over a few months.

2.3.2. Interpolated Activities

Practicing critical skills during the retention interval does aid retention performance. The relevant issues here are (1) for which types of tasks does practice aid, (2) what are the practice task transfer characteristics to the job performance task, and (3) are these dangers of negative task transfer; i.e., practice that can accelerate performance degradation.

Brown et al. (1963) required subjects to perform Naylor's switch setting task as well as a three dimensioal tracking task. Rehearsal on these tasks was manipulated on 4 days of a 15 day retention interval. For the tracking task, rehearsal greatly aided retention performance, but the fidelity of the rehearsal did not alter this result. Performance decrements were effectively erased with task rehearsal. On the procedural task, rehearsal had influence on both commissive and omissive errors. For both types of tasks, sufficiently long original training attenuated the positive effects of rehearsal. When training was more limitied in scope, practice during the retention interval lead to large increases in skill retention. Naylor and Briggs (1963) tested variations in type of rehearsal, on retention performance of their switch setting task. The four rehearsal sessions occurred mid-way in their 25-day retention interval. One group received actual task rehearsal, one received no rehearsal training, and two groups received either part-task temporal or part-task spatial rehearsal. On omissive errors, the actual task group committed about half the errors of the other three rehearsal groups. On commissive errors, the actual task and temporal part-task groups were superior. Whole-task rehearsal was superior to part-task rehearsal conditions. Spatial rehearsal was barely any better than no rehearsal at all, but the time dimension may have been more difficult than the spatial dimension in this task. Trumbo et al. (1965b) compared verbal rehearsal with no rehearsal in a tracking task, over a one month retention interval. Part-task rehearsal required subjects to verbally repeat the tracking target location, referencing to its presented numerical location. On this task, no tracking mean performance retention difference was due to rehearsal, but a greater variability in tracking in the rehearsal group than non-rehearsal group was found.

Sitterley and colleagues investigated the type and distribution of rehearsal for their complex spacecraft simulation tasks. Sitterley and Berge (1972) presented both emergency procedural and flight control tasks to thier inexperienced subjects. After four months of inactivity, both task performances were greatly degraded, beyond the minimal proficiency level. As part of their experimental design, two subject groups recieved static rehearsal training during the retention interval period, where a session consisted of a review of the flight training manual, photographs of the cockpit environment, and a written evaluation test. The static rehearsal greatly countered performance degradation for the procedural task, at both 3 and 6 month intervals. The interim rehearsal aided performance as much as allowing dynamic warmup immediately prior to the start of retention testing. The continuous task, on the other hand, responded differently to rehearsal and warmup training. At a 6 month retention interval, static rehearsal was insufficient to maintain performance in all control skills; dynamic warmup was required to insure reliability. The regular rehearsal sessions were, however, adequate for skill maintenance over the 3 month retention interval. Thus, long retention intervals require both rehearsal and warmup for flight control, but only require static rehearsal for procedural tasks. Using an even more complex space vehicle approach and landing under both visual and instrument flight conditions, Sitterley et al. (1972) added a condition of dynamic rehearsal training, in addition to improving the static rehearsal training method. The improved static method utilized photographs of flight instruments and scenes at critical times, and allowed the subject to sit in the simulator cockpit for refamiliarization prior to testing. The dynamic rehearsal condition included the above static rehearsal, then the pilots were allowed to view three dynamic flights from the cockpit, in an open-loop fashion. The pilot still did not have direct interaction as he would have during warmup practice. Results showed that, like Sitterley and Berge (1972), static rehearsal improved skill retention, but required dynamic warmup practice for adequate proficiency. The dynamic rehearsal prevented skill degradation for all procedural and flight control tasks, with the visual flight control portions receiving the greatest training benefit. The static method was only slightly worse than the dynamic method in retention of flight control or continuous skills.

From a cost/benefit viewpoint, Sitterley (1974) claimed that the static rehearsal method had the greatest development potential. To test an advanced static rehearsal version, Sitterley presented more pictorial information along both normal and sub-nominal flight paths, and enhanced pilot involvement to reinforce critical perceptual cues in the visual environment. The static rehearsal was presented in a booklet format for self-study by the pilot. All retention testing was preceded by a 40 minute slide show of real time cockpit views of the approach and landing. After a 4 month retention interval, the advanced static retraining countered all skill degredation, more so than even the dynamic rehearsal of the previous study. Sitterley suggested that the carefully structured visual cues at critical moments were sufficient to key appropriate pilot responses.

The above studies have nearly all confirmed the utility of retention interval practice in countering both continuous and procedural skill degradation. When training is of insufficient duration, rehearsal methods can be substituted to some degree. The rehearsal training should be carefully designed to provide minimal cues required to successfully perform the task of interest. Experiments by Sitterley have demonstrated that rehearsal for complex flight control does not have to be closed loop and high fidelity. So long as the important visual cues have been provided, open-loop, pictorial reviews may adequately be substituted for the real task. Of those studies reported here, none have concluded that rehearsal degrades retention when compared with no rehearsal. However, none have systematically varied rehearsal tasks so as to provide negative skill transfer.

## 3. SPACE MISSION TRAINING

It is frustrating to study empirical research on task training factors, then consider the techniques that are actually used to define training requirements. Cream et al. (1978) outlined "systematic methods" usually used to specify training objectives in a specific task. First, the behavioral skills and knowledge required of graduated trainees are identified. Next, these are matched against the actual ability of new students. The identified differences then define training requirements of a program. As recognized by Cream et al., a lack of task analysis data for defining training requirements exists, especially with new systems. They recommended seeking out analogous tasks, again avoiding the issue of task analysis. Such experiential-based development can be a costly error in new systems development, where many competing task factors can eclipse unforseen interactions. This report defends the need for a quantitative model of training requirements as a function of task factors. Of course, much research will be required to specify this model.

This section presents a brief outline of operational space mission training at NASA, for the purpose of establishing a foundation from which to define Mars mission training requirments.

### 3.1. CURRENT MISSION TRAINING

Training programs for Space Shuttle missions proceed from part-systems teaching and practice to more complex, fully integrated simulations. A typical astronaut training program currently requires about 5 years from start to flight. Training starts with stand-alone equipment, then proceeds to joint, integrated mission simulation. Initially, workbooks and self-paced computer aided trainers are used to gain knowledge and proficiency on specific systems. Single and multiple part-task trainers are then used to gain required psychomotor proficiency. Examples of these trainers are specific shuttle control panels and the RMS. An underwater weightless training facility, and airborne parabolic flights may be used for specific procedure training. Shuttle systems simultators are now used for many tasks. The multiple-task shuttle simulator may be tied with one or more flight centers in partial-mission simulations. The full mission may then be simulated by tying in all payload customers and control centers. For very complicated maneuvers, this joint simulation may even be repeated on-orbit just prior to the actual performance.

Very little active astronaut performance measurement is currently conducted, once selected (Akins, 1979; Nicogossian, 1984; Christensen and Talbot, 1985). Throughout the literature, a prerequisite for the evaluation and developpment of training procedures is unbiased performance data (see Goldstein, 1978; Cream et al., 1978; Swezey, 1979). As an illustration of this, consider a part-mission simulator session. Trainers prepare scripts of system failures that occur at regular intervals, every few minutes. The task of the trainees is to make educated diagnoses and decisions, while controlling the space vehicle. After completing each simulation run, the training scripts are reviewed with the trainees, pointing out mistakes that were made. A new run then begins, with the hope that lessons have been learned. While trainees clearly learn from this training scenario, limitations in vehicle design or human capabilities are not collected. A series of failures made by all trainees would not be noted; such failures are valuable data to use in the redesign of systems. A separate performance monitoring system, invisible to trainers and trainees, would be useful here.

Soviet cosmonauts also utilize simulators and part-task trainers, but their training philosophy differs in a basic way from U.S. philosophy. Rather than rely on basic documentation in training programs, they listen to a lecture from a specialist several times, taking notes (Lenorovitz, 1982). The Soviet program also places more emphasis on psychological status than the U.S. program, with tests given at training to assure psychological compatibility with crew members, and regular psychological monitoring during and after flight (Borrowman, 1982; Bluth, 1982; Oberg and Oberg, 1986). Future U.S. missions must concentrate more on psychological status of crew members during training (Conners et al., 1985; Collins, 1985).

## 3.2. SPACE STATION TRAINING

A recently published document (NASA, 1986) detailed training requirements for the near-operatioal space station, to be launched in the early 1990's. This section will summarize important aspects of this paper.

### 3.2.1. On-Orbit Versus Ground-Based Training

Specific criteria have been imposed to assign training to on-orbit or ground-based trainers. On-orbit training is preferred for complex psychomotor skills, or time-critical procedures, safety drills, and maintenance of group behavioral dynamics. This training is preferred for microgravity, low-cost, and low probability of occurrance tasks. On-orbit refresher training will also be carried out prior to unscheduled maintenance tasks. Ground-based training will be preferred for fundamental, safety-critical tasks, such as space station activation or medical procedures. Basic training in group dynamics and habitability will be carried out on the ground.

Ground-based training will, in general, be preferred for prerequisite skill acquisition and day-to-day space station operations. On-orbit training is a supplament to basic ground-trained skills, so few skills will only be trained on-orbit.

Initially, most training will be performed on the ground, but more and more training will be performed on-orbit as the space station program matures. Ground-based training will include (1) ingress and egress to and from space station, (2) activation and deactivation of space station, (3) systems training, with emphasis on understanding, (4) spacecraft docking and tethering, (5) RMS and robotics, (6) orbital management and communications, (7) habitability systems, (8) safety, emergency, medical and maintenance procedures, (9) integrated simulations, stressing team approach to problems. On-orbit training will include (1) spacecraft docking and tethering, (2) refresher training on RMS, (3) crew rescue EVA, (4) handling of fuels and other hazardous materials, (5) use of avionics equipment, (6) emergency and malfunction procedures, (7) manned systems refresher training. Ground-based training will initially be performed for all phases and activities. Eventually, primary training for some skills or activities will be shifted to on-orbit. This training will be available both to NASA astronauts and to contractors or customers of NASA.

### 3.2.2. Training Breadth

Training programs for the space station will be developed for the crew, ground-based flight controllers, and training instructors. The crew members will require more training at greater frequency for procedural skills than for psychomotor skills. Launch schedules will impose training duration limits. There are still questions as to the relative amount of self-paced training, amount of on-board training, and relative training differences between crew members. Flight controllers will initially take part in full integrated mission simulations, however later in the space station program, fewer formal simulations will be conducted. Eventually, no joint controller-crew simulation will take place, due to their inherent complexity and time consumption. Instructors must also be trained in procedures. Questions exist as to the number of required instructors per crew member, and the complexity of their simulation scripts.

The general direction for breadth of training is one of initial full scale, integrated simulations involving all parties tapering to later separate simulations of mission components. This change will be required to shorten the training time of space station crews, and to decrease the cost of rotating crews. Some amount of procedural or psychomotor skill practice, such as one-half hour per day, will be mandatory on-orbit.

### 3.2.3. Training Technologies and Facilities

Space station training will make extensive use of computer aided and adaptive instruction. Computer aided instruction systems permit a consistent presentation of material in a given sequential order. Intelligent or adaptive computer aided instruction allows material to be resequenced or altered according to the needs of a trainee (Morgan and Erb, 1986). Both of these types of systems will be utilized in ground-based and on-orbit training. Intelligent systems will be designed to serve as a coach, rather than as a tutor or manager, in that advice is provided to the trainee trying to meet an educational objective.

These instructional programs will be implemented on interactive laser videodisk storage systems. The trainee will respond via keyboard and voice. The system will output via television monitors, wide angle visuals, helmet mounted video, and voice synthesis. Appropriate videodisks for every required repair or maintenance will be onboard; the capacity for transmitting the equivalent content of a videodisk directly into the training hardware may also be present. These systems offer many advantages over conventional computer-based trainers. They are potentially very small and portable, still or moving scenes are of higher fidelity than computer graphics, and it is cheaper to film sequences of movements for videodisk interpretation than developing reliable graphics via computer. Astronauts on EVA will have the capability of viewing procedures in a helmet mounted display as they are performed. These videodisk trainers can also contain other controller attachments to allow realistic practice with complex psychomotor skills.

Space station training will also be embedded within operational controls and displays. By monitoring performance while a trainee attempts to complete a given task, better involvement and motivation are achieved. The monitor will act as a coach, much like a master-apprentice scenario.

Ground-based NASA training facilites for the space station missions will include (1) manned test facility, for engineering design and testing, (2) mockup and integration laboratory, for RMS training, (3) weightless environment training facility, for crew training, (4) systems engineering simulator, for flight training, (5) space station training facility, for high-fidelity crew training, (6) shuttle rendezvous simulator, and (7) integrated EVA simulator. As in the shuttle training program, training will start with part-task, single-system trainers, and end with full-mission, multiple system simulation. Much of the shuttle training facilities will be utilized for space station training.

### 3.4. MARS MISSION TRAINING

Until an empirically determined model of training requirements has been developed, all conceptual designs are merely "straw-man" estimates. However, based on previous sections of this paper, some recommendations may be made.

On-Board Versus Ground-Based Training. Whether training should be given on-board or on the ground should not be an issue on a Mars mission. The crew members must have the necessary resources to rehearse procedural lists and psychomotor skills whenever required. Ground-based training should consist of academic systems overviews and those skills required for complete systems understanding. It is imperitive that complete understanding be achieved prior to flight, as effective refresher training can only be assurred with well organized tasks. Ground-based training might include other knowledge acquisition, beyond the immediate scope of the mission, to guard against unexpected events. As an example, this training might include psychological or social models of small groups. In general, ground-based training should be academic and broad, while on-board required training shoud be specific and skill oriented. Of course, pilots will require all training prior to the mission.

Scope of Training. Ground-based and on-board training and refresher programs must be designed to counteract the negative aspects of the space environment, as discussed in Section 1, in addition to maintaining skill and knowledge retention. Skills must be regularly refreshed, according to a yet-to-be developed model of retention time. Using such a model, a computer program could list, on a daily basis, those skills or procedures that need refresher training. Ideally, a crew member's required refresher training should be determined automatically and adaptively. Periodic performance measurement on a testing battery could indicate level of retention and pinpoint areas for needed training. Training for crew autonomy and confinement will be harder to define, until more is known. Drills may be required to measure the cohesiveness of the crew. Lists of critical procedures must be regularly reviewed and trained, as should the daily workload level. As measured by a model, mental and physical workload must be constantly reviewed and reallocated among the crew members.

All crew members should be encouraged to develop expertise, while in-transit to Mars, in academic fields other than their own. The on-board teaching expertise clearly will exist. Establishing a formal instructional regimen will aid in maintaining cognitive abilites of both teachers and students. Healthy interaction between the crew members will also be maintained. The outcome of such concentrated training could even consist of additional academic degrees.

Periodic Drills. Emergency and disaster drills should be conducted, as called for by either ground control or by the on-board commander. Many controls could be placed in an alternate, embedded training mode for conducting these drills. Images of impending meteorites, etc., could even be projected onto displays or windows. Crew performance should be reviewed by the commander, and necessary refresher training conducted.

Recreation. Off-duty periods also present a good oppurtunity for procedural and operational skill maintenance. Video games, music, etc., all present unique practice oppurtunities for different skills.

Hardware. A small, portable, videodisk based computer system with voice input and output may serve as a generic trainer and recreation device. Such a device will allow practice of skills anywhere and anytime on a mission. Different videodisks could be loaded for different procedures, and others could be loaded for entertainment.

A Mars mission presents many challenges far beyond those that have already been approached. This oppurtunity should be seized for pushing the state-of-the-art in knowledge of human training and skill retention. This paper has stressed the development of empirical models, the only unbiased approach to defining training needs. As the research required to achieve these models will take many years, now is the time to start. A long duration space mission will require an understanding of psychological limitations in all mission phases. This report has stressed the need for modeling these limitations in light of training requirements, whether initial or refresher training. The proper, scientific method of training definition will require a model of skill retetention, as argued here.

## 4. REFERENCES

Adams, J.A., "On the Evaluation of Training Devices," Human Factors, 1979, 21(6): 711-720.

Akins, F.R., "Performance Considerations in Long-Term Space Flight," NASA Technical Report N81-18651, Contract No. NASA-CR-152384, 1979.

Alluisi, E.A., "Methodology in the Use of Synthetic Tasks to Assess Complex Performance," Human Factors, 1967, 9(4): 375-384.

Ammons, R.B., Farr, R.B., Bloch, E., Neumann, E., Day, M., Marion, R., and Ammons, C.H., "Full-Term Retention of Perceptual Motor Skills," Journal of Experimental Psychology, 1958, 55:318-328.

Bahrick, H.P, "Retention Curves: Facts or Artifacts?," Psychological Bulletin, 1964, 61(3): 188-194.

Bahrick, H.P., "The Ebb of Retention," Psychological Review, 1965, 72: 69-73.

Bluth, B.J., "Staying Sane in Space," Mechanical Engineering, 1982, 1: 24-29.

Borrowman, G.L., "Training for Space," Spaceflight, 1982, 24: 129-132.

Brown, I.D., "Dual Task Methods of Assessing Workload," Ergonomics, 1978, 21(3): 221-224.

Brown, D.R., Briggs, G.E., and Naylor, J.C., "The Retention of Discrete and Continuous Tasks as a Function of Interim Practice with Modified Task Requirements," Wright-Patterson AFB Technical Report AMRL-TDR-63-35, Ohio, 1963.

Briggs, G.E., and Wiener, E.L., "Influence of Time Sharing and Control Loading on Transfer of Training," Journal of Applied Psychology, 1966, 50(3): 201-203.

Caro, P.W., "The Relationship between Flight Simulator Motion and Training Requirements," Human Factors, 1979, 21(4): 493-501.

Carter, G., and Trollip, S., "A Constrained Maximization Extension to Incremental Transfer Effectiveness, or, How to Mix your Training Technologies," Human Factors, 1980, 22(2): 141-152.

Chiles, W.D., "Methodology in the Assessment of Complex Performance: Discussion and Conclusions," Human Factors, 1967, 9(4): 385-392.

Christensen, J.M., and Talbot, J.M., "Research Opportunities in Human Behavior and Performance," NASA Contractor Report 3886, Contract NASW-3924, 1985.

Collins, D.L., "Psychological Issues Relevant to Astronaut Selection for Long-Duration Spaceflight," Brooks AFB Technical Report AD-A154 051, Texas, 1985.

Conners, M.M., Harrison, A.A., and Akins, F.R., Living Aloft, Human Requirements for Extended Spaceflight, Washington, D.C.: NASA, 1985.

Cotterman, T.E., and Wood, M.E., "Retention of Simulated Lunar Landing Mission Skill: A Test of Pilot Reliability," USAF Technical Report AMRL-TDR-66-222, Wright-Patterson AFB, Ohio, 1967.

Crawford, A.M., and Crawford, K.S., "Simulation of Operational Equipment with a Computer-Based Instructional System: A Low Cost Training Technology," Human Factors, 1978, 20(2): 215-224.

Cream, B.W., Eggemeier, F.T., and Klein, G.A., "A Strategy for the Development of Training Devices," Human Factors, 1978, 20(2): 145-158.

Fleishman, E.A., "The Prediction of Total Task Performance from Prior Practice on Task Components," Human Factors, 1965, 7(2): 18-27.

Fleischman, E.A., and J.F. Parker, "Factors in the Retention and Relearning of Perceptual-Motor Skills (b)," Journal of Experimental Psychology, 1962, 64: 215-226.

Fraser, T.M., "Leisure and Recreation in Long Duration Space Missions," Human Factors, 1968, 10(5): 483-488.

Gardlin, G.R., and Sitterley, T.E., "Degradation of Learned Skills, A Review and Annotated Bibliography," Report No. D180-15080-1, The Boeing Co., (NASA acquisition No. N73-10152, contract No. NAS9-10962), Seattle, WA, 1972.

Garvey, W.D., "A Comparison of the Effects of Training and Secondary Tasks on Tracking Behavior," Journal of Applied Psychology, 1960, 44(6): 370-375.

Goldberg, S.L., Drillings, M., and Dressel, J.D., "Mastery Training: Effects on Skill Retention," U.S. Army Research Inst. Behav. Soc. Sci. Technical Report 513, 1981, Alexandria, VA.

Goldstein, I.L., "The Pursuit of Validity in the Evaluation of Training Programs," Human Factors, 1978, 20(2): 131-144.

Gopher, D., and North, R.A., "Manipulating the Conditions of Training in Time-Sharing Performance," Human Factors, 1977, 19(6): 583-593.

Grimsley, D.L., "Acquisition, Retention and Retraining: Effects of High and Low Fidelity in Training Devices," Human Resources Research Office, Technical Report 69-1, 1969.

Grodsky, M.A., and Lutman, C.C., "Pilot Reliability and Skill Retention for Space Flight Missions," Air University Review, 1964, 16: 22-32.

Hagman, J.D., "Effects of Training Task Repetition on Retention and Transfer of Maintenance Skill," U.S. Army Research Inst. Behav. Soc. Sci. Technical Report 1271, 1980, Alexandria, VA.

Hagman, J.D., and Rose, A.M., "Retention of Military Tasks: A Review," Human Factors, 1983, 25(2): 199-213.

Hammerton, M., "Retention of Learning in a Difficult Tracking Task," Journal of Experimental Psychology, 1963, 66:108-110.

Joels, K.M., The Mars One Crew Manual, New York: Ballantine Books, 1985.

Johnson, D.F., "Research Note: The Use of Secondary Tasks in Adaptive Training," Human Factors, 1984, 26(1): 105-108.

Johnson, S.L., "Effect of Training Device on Retention and Transfer of a Procedural Task," Human Factors, 1981, 23(3): 257-272.

Johnson, W.B., and Rouse, W.B., "Training Maintenance Techniques for Troubleshooting: Two Experiments with Computer Simulations," Human Factors, 1982, 24(3): 271-276.

Kelly, C.R., "What is Adaptive Training?," Human Factors, 1969, 11(6): 547-556.

Lenorovitz, J.M., "Soviet Cosmonauts Training at Star City Complex," Aviation Week & Space Technology, August, 1982, 44-46.

Lincoln, R.S., "Learning and Retraining a Rate of Movement with the Aid of Kinesthetic and Verbal Cues," Journal of Experimental Psychology, 1956, 51(3): 199-204.

Matheny, W.G., "The Effective Time Constant - A New Technique for Adaptive Training," Human Factors, 1969, 11(6): 557-560.

Mengelkoch, R.F, Adams, J.A., and Gainer, C.A., "The Forgetting of Instrument Flying Skills as a Functio of the Level of Initial Proficiency," NAVTRADEVCEN Technical Report 71-16-18, New York, 1960.

Mengelkoch, R.F, Adams, J.A., and Gainer, C.A., "The Forgetting of Instrument Flying Skills," Human Factors, 1971, 13(5): 397-405.

Moray, N., "Models and Measures of Mental Workload," in Moray, N. (Ed.), Mental Workload, Its Theory and Measurement, New York: Plenum Press, 1979.

Morgan. L.W., and Erb, D., "Survey of Current Training Technology," NASA Working Paper, Crew Training Division, Johnson Space Center, April, 1986.

NASA, Space Station Training Document, Crew Training Divison, Johnson Space Center, 1986.

National Academy of Sciences, Human Factors in Long-Duration Spaceflight, Washington, D.C.: National Academy of Sciences, 1972.

National Commission on Space, Pioneering the Space Frontier, Toronto: Bantam Books, 1986.

Naylor, J.C., and Briggs, G.E., "Long-Term Retention of Learned Skills: A Review of the Literature," USAF Technical Report ASD 61-390, Behavioral Sciences Laboratory, Wright-Patterson AFB, Ohio, 1961.

Naylor, J.C., and Briggs, G.E., "Effect of Rehearsal of Temporal and Spatial Aspects on the Long-Term Retention of a Procedural Skill," Journal of Applied Psychology, 1963, 47(2): 120-126.

Naylor, J.C., Briggs, G.E., and Buckout, "Long-term Skill Transfer and Feedback Conditions during Training and Rehearsal," USAF Technical Report AMRL-TDR-63-136, Wright-Patterson AFB, Ohio, 1963.

Naylor, J.C., Briggs, G.E., Brown, E.R., and Reed, W.G., "The effect of Rehearsal on the Retention of a Time-Shared Task," USAF Technical Report, AMRL-TDR-63-33, Wright-Patterson AFB, Ohio, 1963.

Naylor, J.C., Briggs, G.E., and Reed, W.G., "Task Coherence, Training Time, and Retention Interval Effects on Skill Retention," Journal of Applied Psychology, 1968, 52: 386-393.

Neumann, E., and Ammons, R.B., "Acquisition of Long-Term Retention of a Simple, Serial, Perceptual-Motor Skill," Journal of Experimental Psychology, 1957, 53: 159-161.

Nicogossian, A.E., "Human Capabilities in Space," NASA Technical Memorandum 87360, Washington, D.C., 1984.

Noble, M., and Trumbo, D., "The Organization of Skilled Response," Organizational Behavior and Human Performance, 1967, 2: 1-25.

Oberg, J.E., Mission to Mars, Plans and Concepts for the First Manned Landing, New York: Meridian, 1982.

Oberg, J.E., and Oberg, A.R., Pioneering Space, Living on the Next Frontier, New York: McGraw-Hill Co., 1986.

Parker, J.F., and Fleishman, E.A., "Use of Analytical Information Concerning Task Requirements to Increase the Effectiveness of Skill Training," Journal of Applied Psychology, 1961, 45(5): 295-302.

Reid, G.B., "Training Transfer of a Formation Flight Trainer," Human Factors, 1975, 17(5): 470-476.

Roehrig, W.C., "Psychomotor Task with Perfect Recall After Fifty Weeks of No Practice," Perceptual and Motor Skills, 1964, 19: 547-550.

Savage, R.E., Williges, B.H., and Williges, R.C., "Empirical Prediction Models for Training-Group Assignment," Human Factors, 1982, 24(4): 417-426.

Schendel, J.D., and Hagman, J.D., "On Sustaining Procedural Skills over Prolonged Retention Intervals," U.S. Army Research Institute for the Behavioral and Social Sciences Technical Report 1234, 1980, Alexandria, VA.

Shields, J.L., Goldberg, S.L., and Dressel, J.D., "Retention of Basic Soldiering Skills," U.S. Army Research Institute for the Behavioral and Social Sciences Technical Report 1225, 1979, Alexandria, VA.

Sitterley, T.E., "Degradation of Learned Skills, Static Practice Effectiveness for Visual Approach and Landing Skill Retention," Report No. D180-17876-1, The Boeing Co., (NASA contract No. NAS9-13550), Seattle, WA, 1974.

Sitterley, T.E., and Berge, W.A., "Degradation of Learned Skills, Effectiveness of Practice Methods on Simulated Space Flight Skill Retention," Report No. D180-15081-1, The Boeing Co., (NASA acquisition No. N73-10159, contract No. NAS9-10962), Seattle, WA, 1972.

Sitterley, T.E., Zaitzeff, L.P., and Berge, W.A., "Degradation of Learned Skills, Effectiveness of Practice Methods on Visual Approach and Landing Skill Retention," Report No. D180-15082-1, The Boeing Co., (NASA acquisition No. N73-23086, contract No. NAS9-10962), Seattle, WA, 1972.

Swezey, R.W., "An Application of a Multi-Attribute Utilities Model to Training Analysis," Human Factors, 1979, 21(2): 183-189.

Swink, J., Trumbo, D., and Noble, M., "On the Length-Difficulty Relation in Skill Performance," Journal of Experimental Psychology, 1967, 74: 356-362.

Trollip, S.R., "The Evaluation of a Complex Computer-Based Flight Performance Trainer," Human Factors, 1979, 21(1): 47-54.

Trumbo, D., Noble, M., Cross, K., and Ulrich, L., "Task Predictability in the Organization, Acquisition and Retention of Tracking Skills," Journal of Experimental Psychology, 1965a, 70: 252-263.

Trumbo, D., Ulrich, L., and Noble, M., "Verbal Coding and Display Coding in the Acquisition and Retention of Tracking Skill," Journal of Applied Psychology, 1965b, 49: 368-375.

Trumbo, D., Noble, M., and Swink, J., "Secondary Task Interference in the Performance of Tracking Tasks," Journal of Experimental Psychology, 1967, 73: 232-240.

von Tiesenhausen, G., "An Approach Toward Function Allocation Between Humans and Machines in Space Station Activities," NASA Technical Memorandum 82510, Marshall Space Flight Center, Alabama, 1982.

Williges, R.C., and Williges, B.H., "Critical Variables in Adaptive Motor Skills Training," Human Factors, 1978, 20(2): 201-214.

Youngling, E.W., Sharpe, E.N., Ricketson, B.S., and McGee, D.W., "Crew Skill Retention for Space Missions up to 200 Days," Report F7666, McDonnell-Douglas Astronautics Co., Eastern Div., 1968.