DOCUMENT RESUME

ED 269 236 SE 046 550

AUTHOR	Pfeiffer, Mark G.; Scott, Paul G.
TITLE	Experimental and Analytic Evaluation of the Effects of Visual and Motion Simulation in SH-3 Helicopter Training. Technical Report 85-002.
INSTITUTION	Naval Training Systems Center, Orlando, FL.
PUB DATE	Dec 85
NOTE	60p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Aerospace Technology; *Motion; Postsecondary Education; *Simulation; *Training; *Vision
IDENTIFIERS	*Helicopters

ABSTRACT

A fly-only group (N=16) of Navy replacement pilots undergoing fleet readiness training in the SH-3 helicopter was compared with groups pre-trained on Device 2F64C with: (1) visual only (N=13); (2) no visual/no motion (N=14); and (3) one visual plus motion group (N=19). Groups were compared for their SH-3 helicopter performance in the transition stage of contact flight training. All criterion measures, such as trials-to-mastery, the number of flights and number of hours in the transition stage of flight training, were generally consistent. Transfer ratios, averaged across the three criterion measures, resulted in the best transfer of training under visual plus motion conditions (TR=.31) and about equal transfer for motion only (TR=.22), visual only (TR=.23), and no visual/no motion groups (TR=.24). While all device features reduced the effort for pilot transition to the SH-3 helicopter, the particular task was more important in determining transfer than the device feature used in training the task. (Author/JN)

Naval Training Systems Center

TECHNICAL REPORT 85-002

EXPERIMENTAL AND ANALYTIC EVALUATION OF THE EFFECTS OF VISUAL AND MOTION SIMULATION IN SH-3 HELICOPTER TRAINING

NAVAL TRAINING SYSTEMS CENTER
ORLANDO, FLORIDA

TECHNICAL REPORT 85-002

EXPERIMENTAL AND ANALYTIC EVALUATION OF THE EFFECTS OF VISUAL AND MOTION SIMULATION IN SH-3 HELICOPTER TRAINING

Mark G. Pfeiffer
Paul G. Scott

Training Systems Evaluation Division
Training Analysis and Evaluation Department
Naval Training Systems Center
Orlando, Florida 32813-7100

December 1985

T.W. McNaney

T.W. McNANEY, Director
Training Analysis and
Evaluation Department

W. H. Lindahl

W. H. LINDAHL, Technical Director
Naval Training Systems Center

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Limited to U.S. Govt agencies & their contractors: Admin/Op use (December 1985). Refer other requests to: Commanding Officer, Naval Training Systems Center (Code 1) Orlando, Fl 32813-7100
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report 85-002	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Naval Training Systems Center 6b OFFICE SYMBOL (If applicable) 122	7a. NAME OF MONITORING ORGANIZATION
6c ADDRESS (City, State, and ZIP Code) Orlando, Florida 32813-7100	7b ADDRESS (City, State, and ZIP Code)
8a NAME OF FUNDING/SPONSORING ORGANIZATION 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO / TASK NO / WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification) Experimental and Analytic Evaluation of the Effects of Visual and Motion Simulation in SH-3 Helicopter Training (Unclassified)

12 PERSONAL AUTHOR(S) Mark G. Pfeiffer, Paul G. Scott

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM ____ TO ____		62

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Flight Simulator, SH-3 Helicopter Training, Performance Measures, Trials to Mastery, Transfer of Training, Fleet Readiness Squadron.
05	08		

19 ABSTRACT (Continue on reverse if necessary and identify by block number) A fly-only group (N=16) of Navy replacement pilots undergoing fleet readiness training in the SH-3 Helicopter was compared with groups pre-trained on Device 2F64C with visual-only (N=13), no-visual/no-motion (N=14), and one visual+motion group (N=19). Groups were compared for their SH-3 helicopter performance in the transition stage of contact flight training. All criterion measures, such as trials-to-mastery, the number of flights and number of hours in the transition stage of flight training, were generally consistent. Transfer ratios, averaged across the three criterion measures, resulted in the best transfer of training under visual+motion conditions (TR= 31) and about equal transfer for motion only (TR=.22), visual-only (TR-.23) and no-visual/no-motion groups (TR=.24). While all device features reduced the effort for pilot transition to the SH-3 helicopter, the particular task was more important in determining transfer than the device feature used in training the task.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL LEE SHEPPE	22b TELEPHONE (Include Area Code) 305--646-4498 22c OFFICE SYMBOL NAVTRASYSCEN (Code 10)

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

ACKNOWLEDGEMENTS

Appreciation is expressed to the following persons for their contributions to this study:

- LCDR J. T. Donnelly, LT F. S. Pierce, and Mrs. S. A. Ward of the HS-1 Training Department, NAS Jacksonville, for their cooperation and for report data collection.

- Dr. A. F. Smode and Mr. Robert F. Browning of the Naval Training Systems Center for their study plan and leadership in developing curricula, flight training scenarios, and data collection instruments necessary for prior phases of this study.

- Dr. Richard M. Evans and Mr. William C. McDaniel of the Naval Training Systems Center for technical consultation.

- Dr. Louise G. Yates and Mr. Douglas MacPherson of the Army Research Institute for their contributions to the analysis of analytic data.

- Dr. Arthur S. Blaiwes and Dr. Robert A. Evans of the Naval Training Systems Center for their critical review of this report.

EXECUTIVE SUMMARY

This study examines the separate and joint effects of visual and motion simulation on pilot flight performance in the transition phase, "A" stage, of the SH-3 helicopter flight syllabus to resolve that issue for contact flight in helicopters. Empirical data from the study are also used to validate two models designed to forecast the effectiveness of training devices.

Since 1979, the Training Analysis and Evaluation Department, Naval Training Systems Center, has been studying the use of Device 2F64C in training the SH-3 fleet readiness squadron pilots of Helicopter Antisubmarine Squadron ONE (HS-1).

Technical Report 108 (Browning, McDaniel, & Scott, 1981) presented an evaluation plan for the simulator under conditions with motion, no motion, motion-visual, and visual with no motion. Technical Reports 127 (Browning, McDaniel, Scott, & Smode, 1982), and 153 (McDaniel, Scott, & Browning, 1983) considered the training effectiveness of the device with and without motion. The joint contribution of visual+motion simulation to pilot training was reported by Evans, Scott, & Pfeiffer (1984). The current report concerns the separate and joint effects of visual and motion simulation on piloting performance, and is the final report in this series. It differs methodologically from previous reports in that both experimental and analytic methods are employed. These studies reveal that regardless of the particular device features employed, Device 2F64C significantly reduced the number of flights, flight time, and trials-to-mastery for training replacement pilots to fly the SH-3 helicopter. Most transfer of training occurred with the difficult training tasks (those maneuvers requiring visual and motion cues). Device 2F64C provided least benefit for training non-flying tasks, e.g., normal start, systems checks, etc.

All measures of training effort (trials, aircraft flights, and aircraft flight time) consistently pointed toward visual and motion (VISMOT) as the condition for achieving best transfer of training averaged across tasks. However, because of interactions, device features motion only (MOTNLY), and VISMOT were best for training motion-based tasks and device features visual only (VISNLY), and VISMOT were best for training visual-based tasks.

While transfer ratios varied by task grouping, device feature, and performance measure, an average of ratios yielded 31% savings in effort to fly the aircraft after pretraining pilots in the VISMOT mode. However, it should be noted here that the experimental design lacked a no-practice baseline for the FLYNLY control group. Accordingly, the transfer ratios from the experiment are attenuated by an amount proportional to the extra flying proficiency that had been given for the fly-only control group trained in the Cockpit Procedures Trainer (CPT). Actual transfer probably exceeds 31% for transition tasks in the 2F64C.

Device Effectiveness Forecasting Technique (DEFT) and Forecasting Training Effectiveness (FORTE) showed promise as analytic techniques for modeling transfer coefficients: satisfactory reliability was achieved for most scales with only two raters; DEFT's concurrent validity for estimating transfer efficiency was $r = .55$. FORTE's concurrent validity for transfer was $r = .78$; convergent validity for the transfer coefficients of DEFT and FORTE was estimated at $r = .92$.

Based on these findings the following recommendations are made:

That Device 2F64C be used for training contact flight tasks even when motion and visual systems are not operating. Significant transfer of training occurs without these systems.

That motion cues be emphasized by instructors when training motion-based tasks, e.g., servo malfunctions and ASE malfunctions.

That visual cues be emphasized by instructors when training visual-based tasks, e.g., normal approach and run-on landing.

That visual and motion cues be emphasized by instructors when cue redundancy is important for training, e.g., running takeoff, normal landing and autorotation.

That DEFT scaling be modified to capture the true range of such scales as trials-to-mastery, hours-to-mastery, the transfer ratio and/or the transfer effectiveness ratio. This scaling requires a data base.

That validation of DEFT and FORTE continue with a variety of fielded devices and operational equipments in order to build a data base for forecasting the effectiveness of training devices not yet fielded.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

INTRODUCTION

PROBLEM

The greater importance of visual over motion simulation in the training of military aviators has long been recognized in the fixed wing flight community. However, there is a need for more scientific data concerning the separate and joint effects of visual and motion simulation within the helicopter community (Puig, Harris & Ricard, 1978; Semple, Hennessy, Sanders, Cross, Beith & McCauley, 1981). Design requirements are not yet well defined.

OBJECTIVE

An earlier study by the Training Analysis and Evaluation Group (TAEG Technical Note 6-83, 1983) suggested that visual systems were generally more important than motion systems for the transfer of weapons delivery skills to the aircraft. This study examines the separate and joint effects of visual and motion simulation on pilot flight performance in the transition phase, "A" stage, of the SH-3 helicopter flight syllabus in an effort to resolve that issue for contact flight in helicopters.

BACKGROUND

The Navy antisubmarine warfare helicopter community has been at the forefront in using flight simulation technology. The SH-3 fleet readiness squadron (HS-1) has used a motion-based flight simulator since 1979. During this time, the Training Analysis and Evaluation Department, Naval Training Systems Center, has been studying the SH-3 simulator (Device 2F64C). Technical Report 108 (Browning, McDaniel, & Scott, 1981) presented an evaluation plan for the simulator under conditions with motion, no motion, motion+visual, and visual with no motion. Technical Reports 127 (Browning, McDaniel, Scott, & Smode, 1982), and 153 (McDaniel, Scott, & Browning, 1983) considered the training effectiveness of the device with and without motion. The joint contribution of visual+motion simulation to pilot training was reported in Technical Report 161 by Evans, Scott, & Pfeiffer (1984). The current report concerns the separate and joint effects of visual and motion simulation on piloting performance, and is the final report in this series. It differs methodologically from previous reports in that both experimental and analytic methods are employed.

ORGANIZATION OF THE REPORT

The body of this report is divided into three sections. The first provides a full description of the "Approach" used in conducting the study and defines the salient variables. "Results", the "Conclusions, Discussion, and Recommendations" sections complete the body of the report. Rating scales used as part of an analytic evaluation of Device 2F64C are included as appendices.

APPROACH

EXPERIMENTAL DESIGN

The flight performance of a control group of helicopter replacement pilots who had undergone fleet readiness training without using a flight simulator (FLYNLY) was examined and compared with matched experimental groups of pilots who had (1) simulator training with visual and motion capabilities (VISMOT), (2) simulator training with only visual capabilities (VISNLY), (3) simulator training with only motion cues (MOTNLY) and (4) simulator training with no visual or motion cues (NVSMOT). Data from previous studies (Browning, et al., 1982; McDaniel, et al., 1983; Evans, et al., 1984) were used to form groups matched on undergraduate pilot training (UPT) standard scores. The results are shown in Table 1 and Table 2.

Table 1

UPT Composite Statistics for Five Treatment Groups

STATISTIC	VISMOT	VISNLY	MOTNLY	NVSMOT	FLYNLY
M	47.84	50.66	47.92	50.29	46 31
SD	4.96	4.87	5.97	3.65	2
N	19	13	26	14	

Note: - Based on an ANOVA, treatment means did not differ significantly. ($P > .05$).

Each group of experimental subjects was exposed to a variety of tasks in Device 2C44, a Cockpit Procedures Trainer (CPT) and in Device 2F64C, an Operational Flight Trainer (OFT). All experimental subjects were exposed to all tasks in the OFT and in the SH-3 aircraft. FLYNLY control subjects were not trained in the OFT. However they were trained in the CPT. For analysis purposes, tasks were grouped into four categories according to the cueing needed to accomplish them. The four categories are non-flying (NONFLY), motion-based (MOTION), visual-based (VISUAL) and visual + motion (VMOTION). In the taxonomy of simulator evaluation designs presented in Technical Report 157 (Pfeiffer & Browning, 1984), the present investigation is a variant of Comparison Design 2A. Table 2 illustrates the experimental design with data taken from the experiment. Mastery was defined by the Computer Aided Training Evaluation & Scheduling (CATES) criterion (Rankin & McDaniel, 1980). CATES is a decision model designed to improve efficiency in reaching training decisions. Where insufficient trials existed to implement the CATES criterion, satisfactory completion of the "A" stage check flight was used to define mastery. Additional measures of performance included the number of flights-to-mastery (AFLTS) in "A" stage and hours-to-mastery (ATIME) in "A" stage.

In order to assess the ability of the four simulator conditions in reducing training in the helicopter, transfer ratios (TR) were calculated, using the formula (Roscoe, 1980). When TR is multiplied by 100 it expresses the percent effort saved in the aircraft through pretraining in the simulator.

$$TR = \frac{\text{Mean Effort of FLYNLY} - \text{Mean Effort of Specific Group}}{\text{Mean Effort of FLYNLY}}$$

Table 2

Mean Effort in Aircraft after Training

Under Various Conditions of Simulated Flight for

Four Experimental (E) Groups and One FLYNLY Control (C) Group

CPT Training	Training Condition	Mean Hours in 2F64C	Mean Aircraft Effort*		
			TRIALS	AFLTS	ATIME
YES	VISMOT	11.4	4.3	4.7	12.3
YES	VISNLY	12.0	4.3	5.2	15.1
YES	MOTNLY	12.9	4.9	5.4	13.5
YES	NVSMOT	11.8	4.5	5.6	13.2
YES	FLYNLY	00.0	6.2	7.3	17.2

* TRIALS = Trials-to-Mastery in "A" Stage
AFLTS = Fights-to-Mastery in "A" Stage
ATIME = Hours-to-Mastery in "A" Stage

ANALYTIC DESIGN

The effectiveness of Device 2F64C was also evaluated using the Device Effectivenesss Forecasting Technique (DE..., developed by Rose, Wheaton & Yates

(1985) and Forecasting Training Effectiveness (FORTE) developed by Pfeiffer, Evans and Ford (1985). Both systems are based on expert opinion. Both convert information about various facets of the training system into a forecast of transfer effectiveness. FORTE is presented as Appendix A and DEFT is presented as Appendix B.

Device Effectiveness Forecasting Technique (DEFT)

While a DEFT analysis can be conducted at three levels, only two levels (DEFT I and DEFT II) were deemed necessary for this study. The analyst conducts four major analyses for DEFT I and DEFT II. The training problem and efficiency analyses comprise the first two major analyses. First is an analysis of the training problem to define the deficiency in skills and knowledge that replacement pilots have relative to performance on the training device. As part of the first analysis, the analyst makes magnitude estimates of the difficulty replacement pilots would have in overcoming identified deficiencies. Second, the analyst makes magnitude estimates of the quality of training provided by the training device. During the acquisition efficiency analysis, the analyst determines which instructional features and training principles have been incorporated in the device to help trainees overcome their deficiencies. The transfer problem and transfer efficiency analyses comprise the third and fourth major analyses. In the third, the analyst assesses the transfer problem by scaling the deficiency in flying the aircraft that remains after trainees have practiced on the device and satisfied the device proficiency criterion. The analyst also scales the difficulty in overcoming these residual deficits. Fourth, and finally, the analyst conducts a transfer efficiency analysis. Here the analyst scales how the training device will promote transfer of the learning to the aircraft equipment. All scaling is done in the range of zero to 100.

In order to apply DEFT to the four simulator conditions used in reducing training in the helicopter, transfer (T) and transfer efficiencies (TT) are calculated using the formulas for DEFT I and DEFT II by Rose, Wheaton & Yates (1985). This permits a comparison of modeled and actual transfer coefficients produced by Device 2F64C.

Forecasting Training Effectiveness (FORTE)

FORTE models a variety of aviation training device evaluation outcomes. This simulation model is designed to explore sources of error threatening the sensitivity of device evaluations done by the performance method. Selection of evaluation designs is guided by a model that elicits information from experts. This practical knowledge is transformed into data that are used in simulating a training effectiveness evaluation. Effects of variables such as instructor leniency, task difficulty, and student ability are estimated by two different methods. First, an analyst estimates the joint effect of instructor leniency, task difficulty, and student ability on trials-to-mastery in the aircraft. Second, the analyst estimates the separate effect of the three variables on trials-to-mastery. These estimates are made for the aircraft with and without pretraining. Available in the output is an estimate of transfer ratios based on trials-to-mastery, a diagnosis of deficiencies, an exploration of possible sources of variance, and an estimate of statistical power and required sample size for a transfer experiment.

In order to apply FORTE to the four simulator conditions used in reducing training in the helicopter, transfer ratios are calculated using the formula developed by Roscoe (1980). This permits a comparison of modeled and actual transfer produced by Device 2F64C. Transfer coefficients developed by FORTE are scaled the same as values developed by the performance methods in transfer experiments, i.e., the ratio utilizes trials-to-mastery in the aircraft with and without pretraining in the simulator.

EXPERIMENTAL TEST SUBJECTS

Nineteen newly designated Naval aviators undergoing replacement pilot training at Helicopter Antisubmarine Squadron 1 (HS-1) comprised the visual+motion (VISMOT) simulation group. Data for this group was provided from the previous study by Evans, et al., (1984). Thirteen subjects comprised the visual-only group (VISNLY). The motion-only (MOTNLY) group (N=26) was comprised of the 12 subjects used in the McDaniel, et al., (1983) study, and 14 similarly-trained students used in the Browning, et al., (1982) study. The no-visual/no-motion (N=14) group (NVSMOT) was also drawn from the McDaniel, et al., (1983) data set. Data of the 16 fly-only (FLYNLY) subjects were made available from the data collected during the Browning, et al., (1981) study.

ANALYTIC TEST SUBJECTS

Two evaluators from the Naval Training Systems Center served as test subjects in the DEFT and FORTE evaluation of Device 2F64C. The rating scales used are included as an appendix to this report. These raters had differing degrees of familiarity with Device 2F64C and DEFT itself: Rater 1 was somewhat familiar with DEFT and familiar with FORTE and Device 2F64C. Rater 2 was unfamiliar with DEFT and FORTE and highly familiar with Device 2F64C. Four task categories (NONFLY, MOTION, VISUAL AND VMOTION) representing varying levels of difficulty were scaled with the DEFT Model for each of four device features: VISMOT, VISNLY, MOTNLY, NVSMOT. The FORTE model scaled two levels of instructor leniency, two levels of student ability, and two levels of task difficulty. These three variables were scaled separately and in combination.

AIRCRAFT AND TRAINING DEVICES

General descriptions are provided for the aircraft (SH-3), operational flight trainer (2F64C), and cockpit procedures trainer (2C44).

Aircraft

Replacement pilots were trained in the Sikorsky SH-3 "Sea King" helicopter. The SH-3 is designed for a primary mission of antisubmarine warfare and a secondary mission of search and rescue. The replacement pilot receives flight instruction while occupying the first-pilot position (right seat). The instructor occupies the copilot position (left seat) and performs copilot and safety pilot duties in addition to providing flight instruction.

Operational Flight Trainer

Simulator training for the replacement pilots was conducted in Device 2F64C. The flight section provides training for most tasks associated with transition to the SH-3 and the maintenance of piloting skills. The cockpit area is a high fidelity replication of the SH-3. Training is normally administered simultaneously to two students in the cockpit area. The replacement pilot receiving first-pilot training occupies the right position. The second replacement pilot is positioned in the left seat and serves as copilot. The instructor is positioned at the on-cab instructor station of the flight section. The instructor station is equipped with controls for establishing environmental conditions, problem parameters, malfunction insertion, problem or parameter freeze, and record/playback.

Visual/Video System

The visual system incorporated in Device 2F64C is a VITAL IV system, manufactured by McDonnell Douglas Electronics Company (MDEC), St. Charles, Missouri. The VITAL system, which is an acronym for "Virtual Image Takeoff and Landing," creates a realistic computer generated, high resolution, multicolor image, displayed to the flight crew in virtual image form. The image is reflected from a large concave mirror, providing depth to enhance the realism of the out-of-window illusion. VITAL IV has no flight characteristics of its own but responds to simulated aircraft position, altitude, attitude, and other data provided by the host flight simulator; and to inserted commands such as problem freeze, reset, and slew. The image is generated through the real-time solution of equations in a general purpose digital computer with appropriate peripherals, special components, and specially prepared programs.

Motion System

The motion system consists of controls, Central Processing Unit (CPU) interface circuitry, and the hydraulic system. The motion system is a six-degree of freedom (roll, pitch, yaw, heave, lateral translation, and longitudinal translation) hydraulic powered system which is controlled by the CPU through the interface circuitry and activators. Various safety circuits inhibit motion if a safety violation occurs. The hydraulic system consists of pumps, filters, and valves which provide hydraulic power to both the motion system and the control loading system.

Cockpit Procedures Trainer

Cockpit procedures training for all groups was conducted in Device 2C44. This trailerized device includes a facsimile of the SH-3 cockpit, an instructor console, and a digital computer. It provides training in powerplant management, systems tests, and normal and emergency procedures. Flight is simulated by setting in fixed altitude and airspeed parameters.

TREATMENT

The simulator-trained groups all underwent a six-flight simulator syllabus in Device 2F64C, as described in Browning, et al. (1982), and McDaniel, et al., (1983). A set of flight scenarios was available to ensure that students were exposed to all maneuvers in a standard way. Instructor pilots were enjoined to use these scenarios, and one of the authors rode in the jump-seat of the simulator to ensure standardization. On completion of the simulator sequence, pilots moved to "A" stage in the SH-3 aircraft for training. The flight maneuvers used for grading are defined as "Task Variables" after the "Matching Variable" definition.

DEFINITIONS

The experimental study variables fall into four groups: matching variable, task variables, other dependent variables, and conditions of training.

Matching Variable. Undergraduate pilot training (UPT) composite flight scores were used as a basis for matching. This variable was computed from the mean primary, intermediate, and advanced UPT grades. These grades have a mean of 50 and a standard deviation of 10.

Task Variables. Seventeen task variables were scored in terms of number of trials-to-mastery in the "A" stage flight training syllabus of the SH-3 aircraft. These tasks were selected to represent a much larger set of tasks that were actually trained.

1. Normal start. Starting the No. 1 turbine engine on the SH-3.

2. Blade spread. Properly using the hydraulic blade spread mechanism to change the five rotor blades from a folded to a flying position.

3. System check. Ensuring that the blade spread, hoist, rotary wing head and servos, and basic automatic stabilization equipment (ASE) systems are operating.

4. No. 2 Engine start. Properly starting the number 2 engine.

5. Rotor engagement. The act of bringing the rotor blades up to a speed necessary to maintain control of the helicopter.

6. Taxi. Movement of the airc 't to and from the takeoff/ landing area on the surface.

7. Normal takeoff. Flight beginning from a hover.

8. Normal approach. Making an approach to a landing.

9. Normal landing. Landing from a hover.

10. Running takeoff. Takeoff on a hard surface with no hover.

11. Run-on landing. Landing to a runway with no hover.

14. AUX-off landing. Auxiliary control boost is turned off.

16. Single-engine approach to runway. Usually involves a run-on landing.

17. Single engine wave-off. Aborting a normal approach and taking the aircraft around the flight pattern for another approach.

18. Autorotation. Downward flight using the kinetic energy stored in rotor system for control and emergency landing. In this syllabus there is power recovery at 15 feet altitude.

21. Servo malfunctions. Dealing with hydraulic control failures.

22. ASE malfunctions. Dealing with failures within the electronic automatic stabilization equipment.

Other Dependent Variables

- "A" stage flight hours. Number of flight hours as a first pilot logged in attaining mastery of the syllabus.

- Number of flights in "A" stage. The flight training syllabus calls for five to six flights as a first pilot depending on student's level of mastery of tasks.

Conditions of Training

- o VISMOT. Pilots had both visual and motion aspects of simulator training.
- o VISNLY. Visual-only aspects of simulator training were available.
- o MOTNLY. Motion-only simulator experience.
- o NVSMOT. No visual or motion available during simulator training.
- o FLYNLY. A fly-only group. No simulator experiences in Device 2F64C prior to "A" stage flight training.

TASK CLASSIFICATION

To examine the effect of common task categories on contact flight performance, the authors and squadron training officers categorized the 17 tasks as follows for the purposes of this experiment:

1. NON-FLYING. (NONFLY) Comprised of tasks 1, 2, 3, and 4, involving engine start, blade spread, and system checks.
2. MOTION-BASED. (MOTION) Included tasks 5, 21, and 22 which are thought to be aided by motion cues and do not have visual cues in the simulator.
3. VISUAL-BASED. (VISUAL) Included tasks 6, 8, 11, and 17 involving approach and landing which are preferred to be trained with visual cues but can be trained without motion cues.
4. VISUAL+MOTION. (VMOTION) Included tasks 7, 9, 10, 14, 16 and 18 involving takeoffs and landings under visual flight rules for which both visual and motion cues are considered important for training.

RESULTS

The major questions to be addressed here include the following:

(1) Task difficulty - Does trainee performance in the aircraft support the assumption that the tasks under examination reflect ordered difficulty levels?

(2) Device features - Does trainee performance in the aircraft, when pretrained in the simulator, support the assumption that the groups trained with the device features are distinguishable from the FLYONLY control group?

(3) Interaction - Does trainee performance support the view that there is an interaction between device features and task difficulty?

(4) Transfer - Is there general evidence for transfer of training and do some device features cause some tasks to transfer better than others?

(5) Predicted transfer - With what accuracy is it possible to predict transfer using the DEFT and FORTE analytic models?

TASK DIFFICULTY

Tasks were examined for difficulty level. First, a representative sample of 17 tasks from "A" stage training was selected and tasks were grouped into clusters based on the required cueing of tasks as hypothesized by instructor pilots. Performance on tasks within the clusters was averaged and then compared across categories.

Table 3 presents the mean trials-to-mastery of the five groups of replacement pilots in the four categories of "A" stage flight maneuvers. This grouping was based on the opinions of instructor pilots. The column means show the four categories of tasks grouped in increasing order of difficulty. Non-flying tasks were generally learned most quickly in the aircraft; tasks requiring visual and motion cueing were found to be most slowly learned in the aircraft. The effect of task difficulty was very large, as indicated by the steep increase in the mean trials-to-mastery down each column of Table 3.

Table 3

Mean Trials-to-Mastery in "A" Stage

TASK GROUPING	TRAINING FEATURE				
	N=19 VISMOT	N=13 VISNLY	N=26 MOTNLY	N=14 NVSMOT	N=16 FLYNLY
NON-FLYING TASKS	2.263	2.404	2.231	2.036	2.766
1. Normal start	2.105	2.231	2.346	1.929	2.313
2. Blade Spread	1.684	1.846	1.462	1.143	1.750
3. System Check	2.684	3.077	2.731	2.929	4.063
4. No. 2 Eng Start	2.579	2.462	2.385	2.143	2.938
MOTION-BASED TASKS	3.895	4.205	4.013	4.500	5.271
5. Rotor Engagement	4.789	5.077	5.577	5.429	5.563
21. Servo Malfunction	3.368	3.231	4.346	5.357	4.563
22. ASE Malfunction	3.526	4.308	2.115	2.714	5.688
VISUAL-BASED TASKS	4.539	4.538	5.875	4.964	7.266
6. Taxi	3.158	2.923	2.615	2.000	2.625
8. Normal Approach	7.000	6.462	9.462	8.429	11.688
11. Run-on Landing	4.263	4.462	7.615	5.500	10.125
17. Single Eng WO	3.737	4.308	3.808	3.929	4.625
VISUAL & MOTION TASKS	6.342	6.064	7.615	6.393	9.573
7. Normal Takeoff	6.000	7.846	6.038	5.286	9.750
9. Normal Landing	4.947	4.846	6.192	4.071	10.563
10. Running Takeoff	3.632	3.538	4.038	3.143	5.250
14. Aux-off landing	6.895	7.000	5.500	6.500	7.875
16. Single Eng Appr.	6.947	6.538	7.846	6.286	8.250
18. Autorotation	9.632	6.615	16.077	13.071	15.750

DEVICE FEATURES

The performance of groups exposed to training conditions in the form of different device features was examined. Four device features were studied and compared with a FLYNLY control group: VISMOT, VISNLY, MOTNLY and NVSMOT. Significant savings in effort as a function of prior simulator training was found among all groups trained regardless of different device features. Replacement pilots trained in the simulator with both visual and motion systems (VISMOT) generally required least effort to mastery in the aircraft. Replacement pilots without pretraining in the simulator (FLYNLY) required most effort to mastery.

Table 4 presents the mean effort-to-mastery of the five groups of replacement pilots trained in the "A" stage flight maneuvers. The column means on the left show that the VISMOT group required the least effort and the FLYNLY group (shown on the right) required the most effort regardless of the measure employed.

Table 4

Mean Effort-to-Mastery in Number of Flights (AFLTS),

Time (ATIME) and Trials-To-Mastery (TRIALS) as Measures

under a Variety of Training Conditions

MEASURE	VISMOT	VISNLY	MOTNLY	NVSMOT	FLYNLY
AFLTS	4.75	5.15	5.42	5.64	7.31
ATIME	12.35	15.15	13.53	13.21	17.18
TRIALS	4.26	4.30	4.93	4.47	6.22

Table 5 presents point-biserial correlations describing the accuracy of classifying replacement pilots from measures of groups trained under different conditions. When one of the two variables in a correlation problem is a genuine dichotomy (e.g., experimental vs. control groups), the appropriate type of coefficient to use is the point-biserial correlation. While VISMOT was best overall, clearly all device features produced a reduction in effort needed to fly the aircraft. All performance measures (TRIALS, AFLTS and ATIME) were generally consistent in identifying VISMOT as the most valuable device feature for training. Tabled coefficients are r = .74, r = .74 and r = .64 respectively.

Table 5

Point-Biserial Correlations

to Discriminate those Pilots Pretrained from those in the

FLYNLY Group using TRIALS, AFLTS and ATIME as Measures.

TRAINING CONDITION	N	TRIALS	AFLTS	ATIME
FLYNLY, VISMOT	16,19	.74**	.74**	.64**
FLYNLY, VISNLY	16,13	.68**	.66**	.35*
FLYNLY, MOTNLY	16,26	.72**	.56**	.51**
FLYNLY,NVSMOT	16,24	.48*	.53**	.55**

**$P<.01$, *$P<.05$

INTERACTION

An interaction effect is an effect attributable to the combination of variables above and beyond that which can be predicted from the variables considered singly. Device features differentially influenced task performance. The interaction between task configuration and training conditions presented as device features is displayed as Figure 1. Groups trained in the simulator showed some differential effects across motion-based and visually-based tasks as evidenced by the uneven heights of the bar graphs. Also contributing to the interaction is the gradual diverging of heights for different training conditions. Greater differences occured across training conditions for the complex tasks (VISUAL and VMOTION) than for the simple tasks (NONFLY and MOTION). This divergence from the FLYNLY group illustrates how Device 2F64C had greater training value for tasks which made use of device features than for tasks which made little use of device features. Statistical support for the above is provided in Table 6, the summary of a split plot ANOVA.

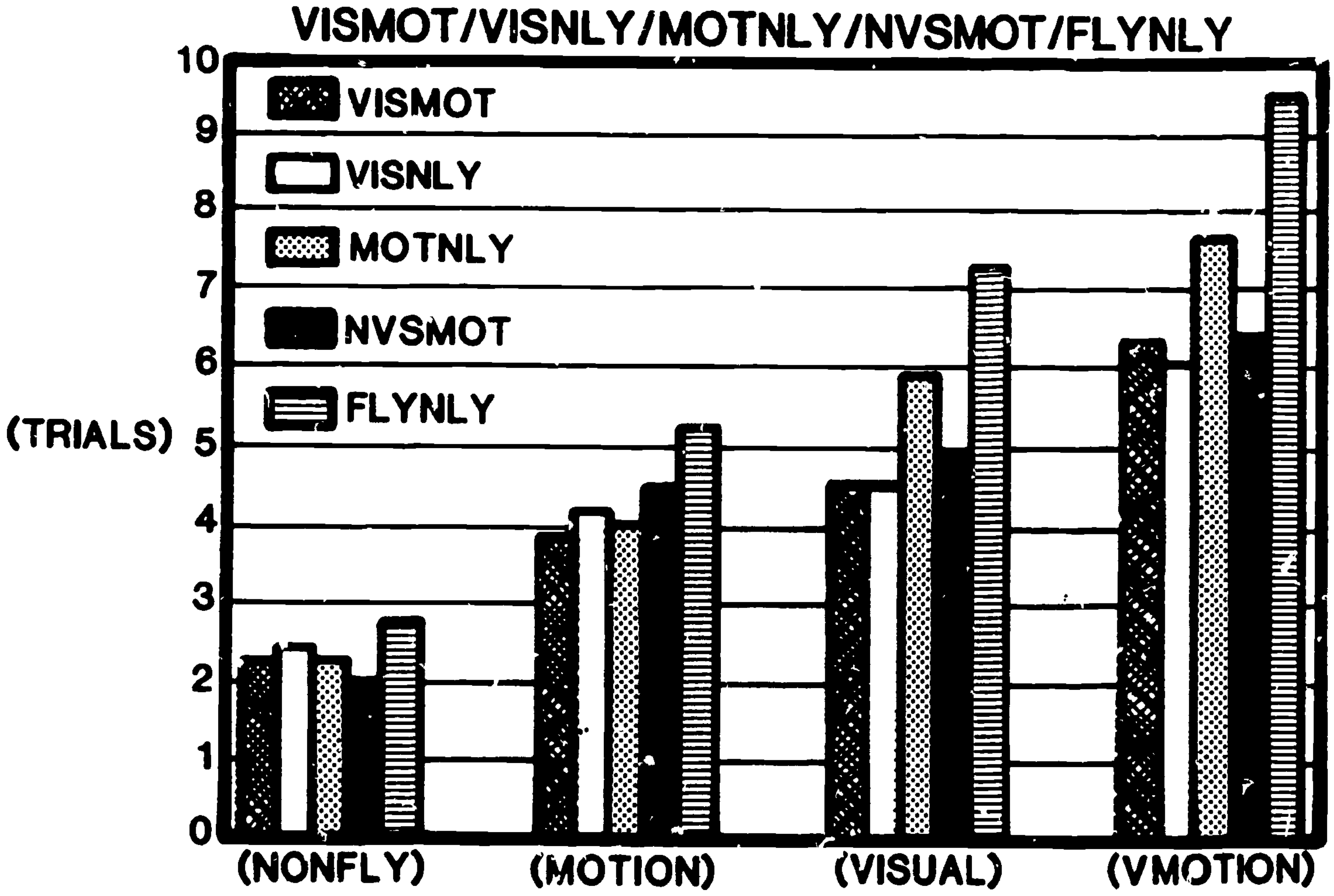

Figure 1. Interaction between training device feature and task configuration.

Table 6

Split Plot Analysis of Variance (ANOVA)

SOURCE	df	SS	MS	F
Between subjects	87	604.847	6.952	
Training Condition(A)	4	170.677	42.669	8.2**
Subjects w/Groups	83	434.170	5.231	
Within subjects	264	1734.608	6.570	
Task category(B)	3	1136.867	378.956	180.9**
A x B	12	76.148	6.346	3.0**
B x subject w/groups	249	521.594	2.095	
TOTAL	351	2339.[illegible]55	6.665	

NOTE: df = Degrees of Freedom; SS = Sum of Squares
MS = Mean Square; F = Fisher's Ratio

**P<.01

Figure 2 presents learning curves for groups trained on tasks requiring both visual and motion cues (VMOTION). The vertical axis shows the cumulative percentage of tasks at mastery level on a trial by trial basis. Replacement pilots trained under FLYNLY conditions are clearly distinguishable from pilots trained in Device 2F64C (VISMOT, VISNLY, MOTNLY and NVSMOT). The learning rate for the FLYNLY control group is slower than the rate for the experimental groups. Ninety percent of all tasks were learned to mastery by the VISMOT, VISNLY and NVSMOT groups in about 10 trials or less, by the MOTNLY group in 13 trials or less, and by the FLYNLY group in 17 trials or less.

Table 7 presents point-biserial correlations among training conditions and tasks. Device feature VISMOT had the greatest value for training tasks that required visual cues (VISUAL r = .67) and the combination of visual and motion cueing (VMOTION r = .56). VISNLY had greatest value for training tasks that required visual cueing (VISUAL r = .63). MOTNLY was most important for training tasks requiring motion cueing (MOTION r = .41). NVSMOT was equally effective for all categories of tasks except motion-based (MOTION) tasks. Tasks requiring motion cueing were not effectively trained without the visual and motion systems (NVSMOT). Similarly, tasks requiring visual cueing were not effectively trained with platform motion on and the visual system off (MOTNLY). Non flying tasks (NONFLY) were not effectively trained with the visual system on (VISNLY).

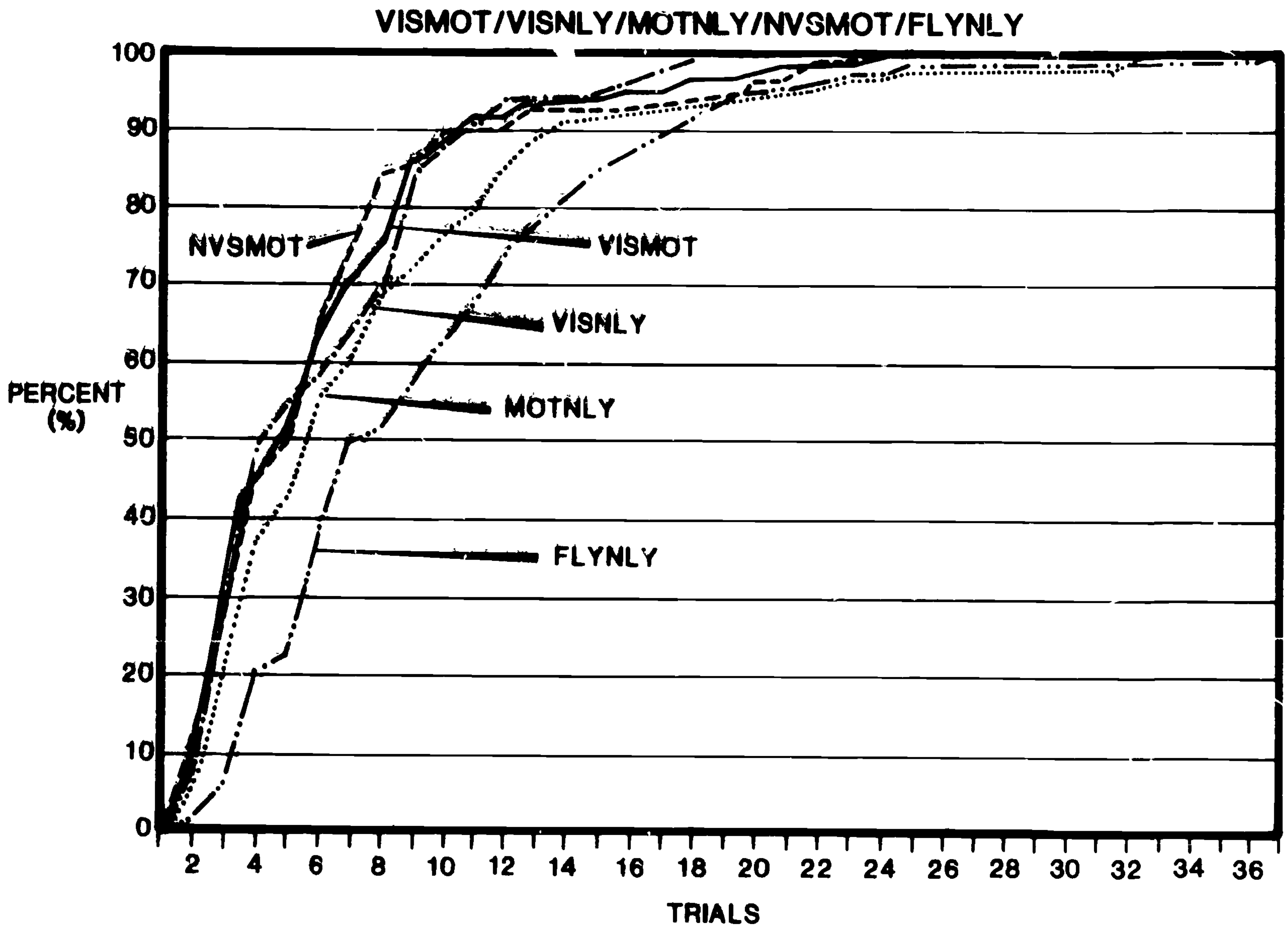

Figure 2. Learning curves for groups trained on VMOTION tasks under different simulator conditions.

Table 7

Point-Biserial Correlations
among Training Conditions and Tasks

TRAINING CONDITIONS	N	----TASK ----			
		NONFLY	MOTION	VISUAL	VMOTION
FLYNLY, VISMOT	16,19	.39**	.49**	.67**	.56**
FLYNLY, VISNLY	16,13	.27	.37*	.63**	.53**
FLYNLY, MOTNLY	16,26	.35*	.41**	.25	.34*
FLYNLY, NVSMOT	16,14	.53**	.26	.52**	.54**

$*P<.05$, $**P<.01$

TRANSFER OF TRAINING

Transfer ratios (TR) for three performance measures are shown in Table 8. TR is an expression of the ability of simulator training to reduce effort needed to master the aircraft. The ratios are presented by device feature and performance measure. Generally, the transfer ratios are quite consistent across all performance measures. Based on a variance components analysis of the data in Table 8, the variance among transfer ratios due to device features was only superficially greater than the variance due to different performance measures. However, the VISMOT group showed greatest transfer regardless of the measure employed. The average transfer for tasks trained under VISMOT is TR = .31. Transfer ratios for the remaining groups trained with device features VISNLY, MOTNLY and NVSMOT ranged between TR = .22 and TR = .24. These mean values are shown at the bottom of Table 8.

Table 8

Transfer Ratios (TR) of Four Simulator

Device Features by Three Performance Measures

PERFORMANCE MEASURE	DEVICE FEATURE				
	VISMOT	VISNLY	MOTNLY	NVSMOT	
FLIGHTS	.35	.29	.25	.23	
FLIGHT TIME	.28	.12	.21	.23	
TRIALS	.29	.27	.20	.26	
MEAN	.31	.23	.22	.24	

Table 9 presents transfer ratios of four simulator device features across four categories of tasks with means across the bottom of the table. Highest transfer resulted from training with both the visual and motion systems on (VISMOT TR = .29). Motion simulation, when used alone, contributed least to the transfer of all tasks (MOTNLY TR = .20). Furthermore, when motion was combined with the visual system (VISMOT) it contributed little over the visual system (VISNLY) alone for training complex tasks. Generally, simple tasks (NONFLY and MOTION) showed less transfer than complex tasks (VMOTION and VISUAL) and were less influenced by the device feature used in training. Based on a variance components analysis of the data in Table 9, the variance among transfer ratios due to tasks was more than 16 times as large as the variance due to device features. Accordingly, the amount of transfer from simulator to aircraft was determined much more by the task than by the device feature used in training the task.

It seems clear from an examination of Table 9 that device features VISMOT, VISNLY, and MOTNLY provided a degree of distraction from tasks not requiring visual and motion cues. For example, the best device feature for nonflying tasks (NONFLY) was NVSMOT (TR = .26). Device features VISMOT and MOTNLY were best for training tasks requiring motion (MOTION) cueing (TR = .26, .24). Device features VISMOT and VISNLY were best for training tasks requiring visual cueing (VISUAL) and the combination of visual and motion cueing (VMOTION).

Table 9

Transfer Ratios of Four Simulator Device

Features Across Four Categories of Tasks

TASK CATEGORY	DEVICE FEATURE				
	VISMOT	VISNLY	MOTNLY	NVSMOT	
NONFLY	.18	.13	.19	.26	
MOTION	.26	.20	.24	.15	
VISUAL	.37	.37	.19	.32	
VMOTION	.34	.37	.20	.33	
MEAN	.29	.27	.20	.26	

MODELING TRANSFER USING DEFT AND FORTE

Device effectiveness forecasting was attempted using the DEFT and FORTE models. DEFT and FORTE employ a series of interactive, menu-driven computer programs that guide an analyst through the evaluation of a training device. Questionnaire versions of DEFT and FORTE are presented as Appendix A and Appendix B. DEFT quantifies opinions about the acquisition and transfer parts of the training system into a variety of estimates of acquisition and transfer effectiveness. FORTE quantifies opinions about the transfer parts of the training system into a transfer ratio based on trials to mastery. The results presented here are based on two raters who applied DEFT I, DEFT II, and FORTE to representative "A" stage task categories of varying difficulty to evaluate four device features: VISMOT, VISNLY, MOTNLY, and NVSMOT. Estimates of rater reliability, concurrent validity and convergent validity are provided in Tables 10, 11 and 13.

Table 10 presents the inter-rater reliability of DEFT scales for the average of two raters using tasks from "A" stage training. DEFT I demonstrated much lower reliability than DEFT II. Scales in DEFT II measuring acquisition demonstrated higher reliability than scales measuring transfer.

Table 10

Reliability of DEFT Scales for

The Average of Two Raters Using Tasks

From "A" Stage Training

SCALE	N ITEMS	RELIABILITY
DEFT I		
VISMOT	8	.72
VISNLY	8	.55
MOTNLY	8	.39
NVSMOT	8	.60
DEFT II ACQUISITION		
Performance Deficit	16	.97
Learning Difficulty	16	.97
Quality of Training Acquisition	16	.96
DEFT II TRANSFER		
Residual Learning Difficulty	16	.96
Physical Similarity	16	.85
Functional Similarity	16	.81
Quality of Training Transfer	12	.92

Table 11 presents the inter-rater reliability of FORTE scales for the average of the same two raters. The scales based on additive methodology (FORTE II) showed higher reliability than the scales based on interactive methodology (FORTE I).

Table 11

Reliability of FORTE Scales for the Average of Two Raters Using Tasks from "A" Stage Training

SCALE	N ITEMS	RELIABILITY
FORTE I		
VISMOT	8	.73
VISNLY	8	.73
MOTNLY	8	.74
NVSMOT	8	.69
FLYNLY	8	.77
FORTE II		
Student Ability	10	.98
Instructor Leniency	10	.99
Task Difficulty	10	.99

Table 12 compares modeled and actual transfer ratios by device feature. Modeled transfer coefficients where TT = $(\text{rating}/100)^{0.5}$ using DEFT I show a greater range than those modeled by DEFT II. However, the ordinal relation of coefficients across DEFT I and DEFT II is identical. Clearly, the modeled transfer coefficients overestimated the actual transfer ratios as shown in table 12. This directional error was large and more consistent for DEFT than for FORTE. FORTE was more accurate than DEFT. Finally, DEFT failed to show that NVSMOT (a lower fidelity condition) was superior to MOTNLY (a higher fidelity condition). FORTE failed in this same way for the interactive method (FORTE I) but succeeded for the additive method (FORTE II).

34

Table 12

Comparison of Modeled and Actual

Transfer by Device Feature

Using Tasks from "A" Stage Flight Training

DEVICE FEATURE	MODELED TRANSFER COEFFICIENT				ACTUAL TRANSFER RATIO
	(DEFT I)	(DEFT II)	(FORTE I)	(FORTE II)	(TR)
VISMOT	.92	.84	.37	.34	.29
VISNLY	.88	.82	.33	.31	.27
MOTNLY	.82	.80	.26	.16	.20
NVSMOT	.79	.77	.24	.20	.26

NOTE: Actual transfer ratio (TR) is based on trials-to-mastery measure taken from Table 9; modeled transfer is based on average of coefficients by two raters.

The columns of transfer coefficients shown in Table 12 and developed by DEFT, FORTE and actual field methods were correlated two at a time to provide estimates of concurrent validity and convergent validity. The validity coefficients are presented in Table 13. Convergent validity for the DEFT and FORTE transfer coefficients is estimated at r = .92; concurrent validity for the DEFT estimate of transfer is r = .55; concurrent validity for the FORTE estimate of transfer is r = .78, a value close (r = .80) to that found by Pfeiffer, Evans and Ford (1985) who used flight instructors as subjects. Table 13 presents the range of validity coefficients and the mean for each type of validity.

Table 13

Validity of DEFT and FORTE

for Estimating Transfer of Training

MODEL	TYPE VALIDITY	RANGE	MEAN
DEFT AND FORTE	Convergent	.81 - .99	.92
DEFT	Concurrent	.45 - .63	.55
FORTE	Concurrent	.68 - .87	.78

NOTE: Validity coefficients were squared prior to averaging.

Another DEFT index is called transfer (T) and is obtained by combining the transfer problem (TRP) and transfer efficiency (TT) indexes where T = TRP/TT. The best device feature (VISMOT) had the smallest score on this index. Except for negative direction, the concurrent validity coefficients had about the same range for T as for TT. Accordingly, the correlations for T are not presented in Table 13.

CONCLUSIONS, DISCUSSION, AND RECOMMENDATIONS

CONCLUSIONS

Regardless of the particular device features employed, Device 2F64C significantly reduced the number of flights, flight time, and trials-to-mastery for training replacement pilots to fly the SH-3 helicopter.

1. Device 2F64C provided the most transfer value for training difficult flying maneuvers (those tasks requiring visual and motion cues).

2. Device 2F64C provided the least benefit for training non-flying tasks (NONFLY). Non-flying tasks are mostly procedural tasks.

3. All three measures of training effort (TRIALS, AFLTS and ATIME) consistently pointed toward device feature VISMOT as the condition for achieving best transfer of training averaged across tasks. However, because of interactions, device features MOTNLY and VISMOT were best for training motion-based tasks (MOTION) and device features VISNLY and VISMOT were best for training visual-based tasks (VISUAL and VMOTION).

4. The amount of transfer from Device 2F64C to the SH-3 aircraft was determined much more by the tasks than by the device features used in training the tasks.

5. While transfer ratios varied by task grouping and device feature, an average of ratios yielded 31% savings in effort to fly the aircraft after pretraining pilots in the VISMOT mode. However, it should be noted here that the experimental design lacked a no-practice baseline for the FLYNLY control group. Accordingly, the transfer ratios from the experiment are attenuated by an amount proportional to the extra flying proficiency that had been given for the FLYNLY control group trained in the CPT. Actual transfer probably exceeds 31% for transition tasks in the 2F64C.

6. DEFT and FORTE showed promise as analytic techniques for modeling transfer coefficients; satisfactory reliability was achieved for most transfer scales with only two raters; DEFT's concurrent validity for estimating transfer efficiency was $r = .55$. FORTE's concurrent validity for transfer was $r = .78$; convergent validity for the transfer coefficients of DEFT and FORTE was estimated at $r = .92$.

DISCUSSION

While the pilot is a visually oriented person, and tends to seek visual information first, he/she is prepared to take whatever information is available to him/her. The need for maneuver motion cues and external visual cues for contact flight depends in part on what is available. When visual and motion cues are taken away, then the pilot still has instruments. Evidence from the present study indicated that training without visual and motion cues (NVSMOT) provided significant transfer of training for contact flight tasks which seemingly required simulation of these cues. This evidence provides some support for a positive transfer from instrument to contact flight, a finding, reported by other investigators many years ago (Muckler, Nygaard, O'Kelly & Williams, 1959; Ritchie & Michael, 1955; Lee, 1935).

Unfortunately, the rather subtle finding of the experiment that a lower fidelity feature such as NVSMOT provided equal or better transfer of training for contact flight than a higher fidelity feature such as MOTNLY was not duplicated with application of the DEFT model The DEFT model does not properly combine physical and functional fidelity scales to yield a transfer coefficient. One possibility for adjusting the DEFT model to account for this problem has been described by Adams (1979). Adams proposed a family of relationships between transfer of training and similarity of the training and transfer tasks that permit proper coupling of similarity and transfer data.

RECOMMENDATIONS

It is recommended that:

1. Device 2F64C be used for training contact flight tasks even when motion and visual systems are not operating. Significant transfer of training occurs without these systems.

2. Motion cues be emphasized by instructors when training motion-based tasks, e.g., servo malfunctions and ASE malfunctions.

3. Visual cues be emphasized by instructors when training visual-based tasks, e.g., normal approach and run-on landing.

4. Visual and motion cues be emphasized by instructors when cue redundancy is important for training, e.g., running takeoff, normal landing and autorotation.

5. DEFT scaling be modified to capture the true range of such scales as trials-to-mastery, hours-to-mastery, the transfer ratio and/or transfer effectiveness ratio. This scaling requires a data base.

6. Validation of DEFT and FORTE continue with a variety of fielded devices and operational equipments in order to build a data base for forecasting the effectiveness of training devices not yet fielded.

POST NOTE

During conduct of the visual only (VISNLY) study it was noted that several pilots experienced a phenomenon known as "Simulator Sickness." This sickness included symptoms such as general discomfort, headache and eyestrain.

The problem of "Simulator Sickness" has been studied through a special focus 6.2 research program by NAVTRASYSCEN Codes 711 and 732. Field studies and laboratory research will produce guidelines for students and instructors in operational settings. Specific research of the NTSC Visual Technology Research Simulator will provide guidelines for engineers concerning throughput delay specifications that have a bearing on "Simulator Sickness." Future NTSC technical reports will elaborate on this research area.

39

REFERENCES

Adams, J. A. On the evaluation of training devices. Human Factors, 1979, 21, 711-720.

Browning, R. F., McDaniel, W. C., & Scott, P. G. Preparation and design for a training effectiveness evaluation of device 2F64C for rep' cement pilot training (TAEG Technical Report 108). Orlando: Training Analysis and Evaluation Group, August 1981.

Browning, R. F., McDaniel, W. C. Scott, P. G., & Smode, A. F. An assessment of the training effectiveness of device 2F64C for training helicopter replacement pilots (TAEG Technical Report 127). Orlando: Training Analysis and Evaluation Group, July 1982. (DTIC No. AD A118[illegible])

Evans, R. M., Scott, P. G., & Pfeiffer, M. G. SH-3 Helicopter flight training: an evaluation of visual and motion simulation in device 2F64C (TAEG Technical Report 161). Orlando: Training Analysis and Evaluation Group, Naval Training Equipment Center, December 1984. (AOB 090 114)

Lee, T., Jr. Instrument flying from scratch. Aviation, 1935, 34, (12).

McDaniel, W. C., Scott, P. G., & Browning, R. F. Contribution of platform motion simulation in SH-3 helicopter pilot training (TAEG Technical Report 153). Orlando: Training Analysis and Evaluation Group, Naval Training Equipment Center, October 1983. (DTIC No. AD A134905)

Muckler, F. A., Nygaard, J. E., O'Kelly, L. I. & Williams, A. C., Jr., Psychological variables in the design of flight simulators for training (WADC Technical Report 56-369) Dayton, OH: AERO Medical Laboratory, Wright Air Development Center, Air Research and Development Command, January 1959. (ASTIA NO. AD97130)

Pfeiffer, M. G., & Browning, R. F. Field evaluations of aviation trainers (TAEG Technical Report 157). Orlando: Training Analysis and Evaluation Group, Naval Training Equipment Center, June 1984. (AD B083584)

Pfeiffer, M. G., Evans, R. M., & Ford, L. H. Modeling field evaluations of aviation trainers (TAEG Technical Note 1-85). Orlando: Training Analysis and Evaluation Group, Naval Training Equipment Center, January 1985.

Puig, J. A., Harris, W. T., & Ricard, G. L. Motion in flight simulation; an annotated bibliography (NAVTRAEQUIPCEN Technical Note IH-298). Orlando: Naval Training Equipment Center, July 1978.

Rankin, W. C., & McDaniel, W. C. Computer aided training evaluation and scheduling (CATES) system: assessing flight task proficiency (TAEG Technical Report 94). Orlando: Training Analysis and Evaluation Group, December 1980. (DTIC No. AD A095007)

Ritchie, M. L., & Michael, A. L. Transfer between instrument and contact flight training. Journal of Applied Psychology, 1955, 39, 145-149.

Roscoe, S. N. Aviation psychology. Ames, Iowa: Iowa State University Press, 1980.

Rose, A. M., Wheaton, G. R, & Yates, L. G. Forecasting device effectiveness: Volume I. Issues (Technical Report 680). Alexandria, VA: U. S. Army Research Institute, December 1985.

Rose, A. M., Wheaton, G. R., & Yates, L. G. Forecasting device effectiveness: Volume II. Procedures (Research Product 85-25). Alexandria, VA: U. S. Army Research Institute, June 1985.

Rose, A. M., Martin A. W., & Yates, L. G. Forecasting device effectiveness: Volume III. Analytic assessment of device effectiveness forecasting technique (Technical Report 681). Alexandria, VA: U. S. Army Research Institute, June 1985.

Semple, C. A., Hennessy, R. T., Sanders, M. S. Cross, B. K., Beith, B. H.,& McCauley, M. E. Aircrew training devices: Fidelity features (AFHRL-TR-80-36). Brooks AFB, TX: Air Force Human Resources Laboratory, [illegible]ary 1981.

Training Analysis and Evaluation Group, Effectiveness of simulation for air-to-air and air-to-surface weapons delivery training (TAEG Technical Note 6-83). Orlando: Author, September 1983.

APPENDIX A

QUESTIONNAIRE

FORECASTING TRAINING EFFECTIVENESS (FORTE)

FORECASTING TRAINING EFFECTIVENESS (FORTE)

OVERVIEW: This questionnaire is designed for senior officers, flight instructors, and experienced squadron pilots in Navy Fleet replacement squadrons.

It elicits information that will enable evaluators to guide the design and execution of transfer of training studies involving flight simulators. We are particularly interested in your estimates of the number of trials a student pilot needs to demonstrate NATOPS-level mastery of a variety of training tasks taught by a variety of instructors both with and without the aid of a flight simulator.

I. First, think of a group of student pilots in your squadron who have completed simulator training prior to checking out in the aircraft. Please make estimates of the number of trials needed for mastery under each of the following eight sets of conditions.

INSTRUCTOR	STUDENT	TASK	NUMBER TRIALS IN AIRCRAFT (FORTE I)
1. Easy	Fast	Easy	
2. Easy	Fast	Tough	
3. Easy	Slow	Easy	
4. Tough	Fast	Easy	
5. Easy	Slow	Tough	
6. Tough	Fast	Tough	
7. Tough	Slow	Easy	
8. Tough	Slow	Tough	

9. Now, please rank the following variables for their importance to the estimations you just made:

Rank	Variable
	Instructors
	Students
	Tasks

Administrator: Sum the eight sets of trials recorded above and divide by 8. Insert this mean value (rounded to a whole number) following the symbol "*N*" in questions 10-12. (FORTE II)

10. If an average student requires *N* trials to learn to mastery, how many trials will

... a fast learner require?
... a slow learner require?

11. If an average instructor requires *N* trials to train students, how many trials will

... an easy instructor need?
... a tough instructor need?

12. If *N* trials are needed for average tasks, how many trials would

... an easy task require?
... a tough task require?

II. Now we will answer similar questions for a group of students who <u>have not had simulator experience</u>.

INSTRUCTOR	STUDENT	TASK	NUMBER TRIALS IN AIRCRAFT (FORTE I)
13. Easy	Fast	Easy	
14. Easy	Fast	Tough	
15. Easy	Slow	Easy	
16. Tough	Fast	Easy	
17. Easy	Slow	Tough	
18. Tough	Fast	Tough	
19. Tough	Slow	Easy	
20. Tough	Slow	Tough	

21. Now, again rank these variables for their order of importance in determining trials to mastery:

Rank	Variable
	Instructors
	Students
	Tasks

Administrator: Sum the trials listed in response to questions 13-20 and divide by 8. Enter this rounded value appropriately following the symbol "*M*" in the three questions that follow. (FORTE II)

---

22. If an average student requires *M* trials-to-mastery, how many trials will

    ... a fast learner need?
    ... a slow learner need?

23. If an average instructor requires *M* trials to train students, how many will

    ... an easy instructor need?
    ... a tough instructor need?

24. If *M* trials are needed for average tasks, how many trials would

    ... an easy task require?
    ... a tough task require?

APPENDIX 3

QUESTIONNAIRE

DEVICE EFFECTIVENESS FORECASTING TECHNIQUE (DEFT)

DEVICE EFFECTIVE FORECASTING TECHNIQUE (DEFT)

Training Problem Analysis: DEFT I

PERFORMANCE DEFICIT

I. Examine the statement of the training objective(s). Considering what you know about the typical trainee's background, work experience, and prior training, what proportion of the skills and knowledges required in order to meet the training objective(s) will the trainee still have to learn in order to reach criterion proficiency in the training device?

0 = None; the trainee can already meet the training objective(s).

0

100

100 = All; the trainee has to learn all of the skills and knowledges needed to meet the training objective(s).

LEARNING DIFFICULTY

II. Consider the enabling skills and knowledges required to meet the training objective(s) that the typical trainee does not currently possess. Rate the difficulty of acquiring the remaining skills and knowledges.

0 = Very easy to learn; it will take practically no training or practice on the device to learn the skills and knowldeges needed to meet the training objective(s).

0

100

100 = Very difficult to learn; it will take a lot of training or practice on the device to learn the skills and knowledges needed to meet the training objective(s).

Acquisition Efficiency Analysis: DEFT I

QUALITY OF TRAINING ACQUISITION

I. Examine information about the instructional features of the training device, the training principles it incorporates, the program for its implementation, and the larger training context in which the device is embedded. Consider the performance deficits you have identified and how utilization of the device will overcome these deficits.

To provide "excellent" training, the training system should:

- o make the performance requirements of the training objective(s) explicit to the trainees:
- o provide meaningful and understandable feedback to the trainee regarding the results of his performance as soon as possible following his performance:
- o provide sufficient practice where specific and hard-to-learn physical skills are involved; and
- o provide a record of trainee performance.

Rate the quality of the training provided by this training system, considering only the training problems you have identified.

0 = Poor training; the system embodies few if any sound training principles and instructional features.

0

100

100 = Excellent training; the system makes maximum use of sound training principles and instructional features.

Transfer Problem Analysis DEFT I

RESIDUAL DEFICIT

I. Assume that the trainee has achieved the training objective(s) (i.e., has reached criterion proficiency on the training device). What proportion of the skills and knowledges required in order to reach criterion proficiency on the operational equipment will the trainee still have to learn?

0 = None; the trainee can already meet the operational performance objectives.

0

100

100 = All; the trainee has to learn all of the skills and knowledges needed to meet the operational performance objective(s).

RESIDUAL LEARNING DIFFICULTY

II. Consider the skills and knowledges that a graduate of the training device must still acquire in order to perform at criterion level(s) on the operational equipment. Rate the difficulty of acquiring the remaining skills and knowledges.

0 = Very easy to learn; it will take practically no training or practice on the operational equipment to learn the skills and knowedges needed to meet the operational performance objectives(s).

0

100

100 = Very difficult to learn; it will take a lot of training or practice on the operational equipment to learn the skills and knowledges needed to meet the operational performance objective(s).

PHYSICAL SIMILARITY

Physical similarity is based on the similarity between physical characteristics of the training device and those of the operational situation. The assessment is based on the physical similarity (e.g., location, appearance, and feel) of displays, controls, and ambient conditions in the training and operational setting. Determine the physical similarity between the training device and the operational equipment.

0 = Totally dissimilar; there would be a large noticeable difference, quite apparent to the trainee at transfer and a large performance decrement, given that the trainee could perform at all; specific instruction and practice would be required on the operational equipment after transfer to overcome the deficit.

0

100

100 = Identical; the trainee would not notice a difference between the training device and the operational equipment at the time of transfer.

FUNCTIONAL SIMILARITY

Functional similarity is based on the operator's behavior in terms of the information flow from each display to the operator, and from the operator to each control. The assessment is made in terms of the amount of information transmitted from each display to each control and the type of information-processing activity performed by the operator. Determine how functionally similar the training device and operational equipment are.

0 = Totally dissimilar; the trainee acts on completely different types and amount of information in the training device and the operational equipment; the trainee carries out different information-processing activities.

0

100

100 = Identical; the trainee acts on the same types and amounts of information in the training device and the operational equipment;' the trainee carries out the same information-processing activities.

Transfer Efficiency Analysis DEFT I

QUALITY OF TRAINING TRANSFER

1. Consider the statement of the operational performance objective(s), as given in the Training Device Requirement Document, the statement of the training objective(s), and descriptions of the operational equipment and the training device.

Consider the instructional features and training principles that are included in the device to increase the probablility that the skills and knowledges acquired on the device will be used effectively in the operational situation. Rate how well the training device will promote transfer to the operational situation.

0 = Poor transfer; the device embodies few if any sound training principles and instructional features to promote transfer to the operational equipment.

0

100

100 = Excellent transfer; the device makes maximum use of sound training principles and instructional features to promote transfer to the operational equipment.

DEVICE EFFECTIVENESS FORECASTING TECHNIQUE (DEFT)

Training Problem Analysis: DEFT II

PERFORMANCE DEFICIT

I. Examine the statement of the training objective(s). Considering what you know about the typical trainee's background, work experience, and prior training, what proportion of the skills and knowledges required in order to meet the training objective(s) will the trainee still have to learn in order to reach criterion proficiency in the training device?

0 = None; the trainee can already meet the training objective(s).

0

100

100 = All; the trainee has to learn all of the skills and knowledges needed to meet the training objective(s).

LEARNING DIFFICULTY

II. Consider each task (subtask) that you indicated a trainee won't be able to perform initially on the training device. Rate the difficulty the typical trainee will have in learning to perform each task (subtask).

0 = Very easy to learn; it will take practically no training or practice on the device to reach criterion proficiency on this task (subtask).

0

100

100 = Very difficult to learn; it will take a lot of training or practice on the device to reach criterion proficiency on this task (subtask).

Acquisition Efficiency Analysis: DEFT II

QUALITY OF TRAINING ACQUISITION

I. Examine information about the instructional features of the training device, the training principles it incorporates, the program for its implementation, and the larger training context in which the device is embedded. Consider the performance deficits you have identified. Rate how well utilization of the device will overcome these deficits.

For what percentage of the tasks (subtasks) that must be learned does the training system make the criterion performance requirements explicit to the trainee?

0 = None; performance requirements are not made explicit to trainees.

0

100

100 = All; performance requirements are made explicit to trainees on all tasks (subtasks) they must learn.

QUALITY OF TRAINING ACQUISITION

II. For what percentage of the tasks (subtasks) that must be learned does the training system provide practice?

0 = None; practice is not provided for on any of the tasks (subtasks) which must be learned.

0

100

100 = All; practice is provided for on all of the tasks (subtasks) which must be learned.

QUALITY OF TRAINING ACQUISITION

III. For what percentage of the tasks (subtasks) that must be learned does the training system provide qualitative feedback to the trainees about their performance?

0 = None; feedback about performance is not provided on any of the tasks (subtasks) which must be learned.

0

100

100 = All; feedback about performance is provided on all of the tasks (subtasks) which must be learned.

QUALITY OF TRAINING ACQUISITION

IV. For what percentage of the tasks (subtasks) that must be learned does the training system provide a record of trainee performance?

0 = None; records of trainee performance are not provided for any of the tasks (subtasks) which must be learned.

0

100

100 = All; records of performance are provided for all of the tasks (subtasks) which must be learned.

Transfer Problem Analysis DEFT II

RESIDUAL DEFICIT

I. Assume that the trainee can perform all of the tasks (subtasks) comprising the training objective(s) (i.e., has reached criterion proficiency on each task (subtask) in the training device.

For each task (subtask) associated with the operational performance objective(s), enter a value from 1 to 4 as indicated below:

1. Operational task (subtask) was represented in the training objective; most trainees will be able to perform this task (subtask) with minimal exposure to or practice on the operational equipment.

2. Operational task (subtask) was not represented in the training objective; but most trainees will be able to perform this task (subtask) with minimal exposure to or practice on the operational equipment.

3. Operational task (subtask) was represented in the training objective; but most trainees will not be able to perform this task (subtask) with minimal exposure to or practice on the operational equipment.

4. Operational task (subtask) was not represented in the training objective; most trainees will not be able to perform this task (subtask) with minimal exposure to or practice on the operational equipment.

RESIDUAL LEARNING DIFFICULTY

II. Consider each operational task (subtask) that you indicted the typical trainee won't be able to perform initially on the operational equipment. Rate the difficulty the typical trainee will have in learning to perform each task (subtask).

0 = Very easy to learn; it will take practically no training or practice on the operational equipment to reach criterion proficiency on this task (subtask).

0

100

100 = Very difficult to learn; it will take a lot of training or practice on the operational equipment to reach criterion proficiency on this task (subtask).

PHYSICAL SIMILARITY

III. Physical similarity is based on the similarity between physical characteristics of the training device and those of the operational situation. The assessment is based on the physical similarity (e.g., location, appearance, and feel) of displays, controls, and ambient conditions in the operational and training tasks (subtasks). Rate the physical similarity between each operational task (subtask) and its counterpart (if any) in the training device.

0 = Totally dissimilar; although the task is represented in the training device, there would be a large noticeable difference quite apparent to the trainee at transfer and a large performance decrement, given that the trainee could perform the task at all; specific instruction and practice would be required for this task (subtask) on the operational equipment after transfer to overcome the deficit.

0

100

100 = Identical; the trainee would not notice a difference between the training device and the operational equipment for this task (subtask) at the time of transfer.

FUNCTIONAL SIMILARITY

IV. Functional similarity is based on the operator's behavior in terms of the information flow from each display to the operator and from the operator to each control. The assessment is made in terms of the amount of information transmitted from each display to each control and the type of information-processing activity performed by the operator. Rate the functional similarity between each operational task (subtask) and its counterpart (if any) in the training device.

0 = Totally dissimilar; for this task, the trainee acts on completely different amounts and types of information in the training device and the operational equipment; the trainee carries out different information-processing activities in the two versions of the task.

0

100

100 = Identical; for this task the trainee acts on the same types and amounts of information in the training device and the operational equipment; the trainee carries out the same information-processing activities in the two versions of the task.

Transfer Efficiency Analysis: DEFT II

QUALITY OF TRAINING TRANSFER

I. Consider the statement of the operational performance objective(s), as given in the Training Device Requirement Document, the statement of the training objective(s), and descriptions of the operational equipment and the training device.

Rate how well the training device will promote transfer to the operational situation. Consider the instructional features and training principles that are included in the device to increase the probability that the skills and knowledges acquired will be used effectively in the operational situation.

What percentages of the tasks (subtasks) that must be learned in the device are realistic and relevant in the sense that they are similar to the tasks that are performed in the real world?

0 = None; the learning tasks are not realistic, relevant or similar to those in the real world.

0

100

100 = All; the learning tasks are realistic, relevant or similar to those in the real world.

QUALITY OF TRAINING TRANSFER

II. For what percentage of the tasks (subtasks) that must be learned in the device are the conditions of practice late in training made to approximate those in the real world?

0 = None; late in training the conditions of practice do not approximate those likely to be encountered in the real world.

0

100

100 = All; late in training the conditions of practice are made to approximate those in the real world on all of the tasks trainees must learn in the device.

QUALITY OF TRAINING TRANSFER

III. For what percentage of the tasks (subtasks) that must be learned in the device is an extensive amount of practice given?

0 = None; not even a single task is practiced extensively.

0

100

100 = All; every task that trainees must learn in the device is practiced extensively.

DISTRIBUTION LIST

Navy

OASN (MRA)
CNO (OP-115, OP-987H, OP-987, OP-12, OP-11, OP-594, OP-596, OP-01B7)
ONR (442 (3 copies), 270)
CNET (01, 00A, 00A2, L02)
CNAVRES (02)
COMNAVAIRSYSCOM (APC 205-0, APC 205-1)
CNATRA (Library (2 copies))
CO NAVPERSRANDCEN (Library (4 copies))
NAVPGSCOL (2124, 32)
USNA (Chairman, Behavioral Science Dept.)
NAVEDTRASUPPCENLANT (N-3 (2 copies))
NAVEDTRASUPPCENPAC (2 copies)
NAVAEROMEDRSCHLAB (Chief Aviation Psych. Div.)
NAVTRASYSCEN (TIC, 10 (3 copies), 12 (10 copies), 713, 001, 002, 09, 04, 7_)
Center for Naval Analyses
OIC NODAC (2)
CO HELTRARON EIGHT
CO HELTRARON EIGHTEEN
NAVAEROSPMEDINST (Code 11)
COMSEABASEDWINGSLANT
COMNAVAIRLANT (312, Helo Trng Officer; 316)
COMHSWING ONE (3 copies)
CO HELANTISUBRON ONE (3 copies)
CO HELANTISUBRON TEN (3 copies)
COMNAVAIRPAC
CO FASOTRAGRULANT
OIC FASOTRAGRULANT DET JACKSONVILLE
CNET Liaison Officer, Williams AFB
HSL 31 (Training Dept.)
COMNAVAIRTESTCEN (RW-50)
NAVSAFCEN (Code 11)

Air Force

Headquarters, Air Training Command (XPTD, XPTIA) Randolph Air Force Base
Air Force Human Resources Laboratory, Brooks Air Force Base
Air Force Human Resources Laboratory (Flight Training Division), Williams Air Force Base
Air Force Office of Scientific Research/NL
Headquarters Tactical Air Command (DOOS), Langley Air Force Base
AFMTC/XR, Lackland Air Force Base
Headquarters MAC/DOT, Scott Air Force Base
4235 Strategic Training Squadron, Carswell Air Force Base
1550th Aircrew Training and Test Wing, Aerospace Rescue and Recovery Service, Kirtland Air Force Base

DISTRIBUTION LIST (continued)

Army

Commandant, TRADOC (Technical Library)
ARI (Technical Director, PERI-SM, PERI-IC, Library (2 copies))
ARI (Reference Service)
Director of Evaluation and Standards, USAAVNC, Fort Rucker
Director of Training Development, USAAVNC, Fort Rucker
Director of Academic Training. USAAVNC, Fort Rucker
Commanding Officer, USASC, Fort Rucker
Commanding Officer, USAARL, Fort Rucker

Coast Guard

Commandant, Coast Guard Headquarters (G-P-1/2/42, GRT/54)
U.S. Coast Guard Aviation Training Center, Mobile, Alabama (Kingsley Povenmire)

Marine Corps

CMC (OT, TDA)
CGMCDEC
Director, Marine Corps Institute
COMCCES

Other

Military Assistant for Human Resources, OUSDR&E, Pentagon
Institute for Defense Analyses

Information Exchanges

DTIC (12 copies)
DLSIE
Executive Editor, Psychological Abstracts, American Psychological Association
ERIC Processing and Reference Facility, Bethesda, MD (2 copies)