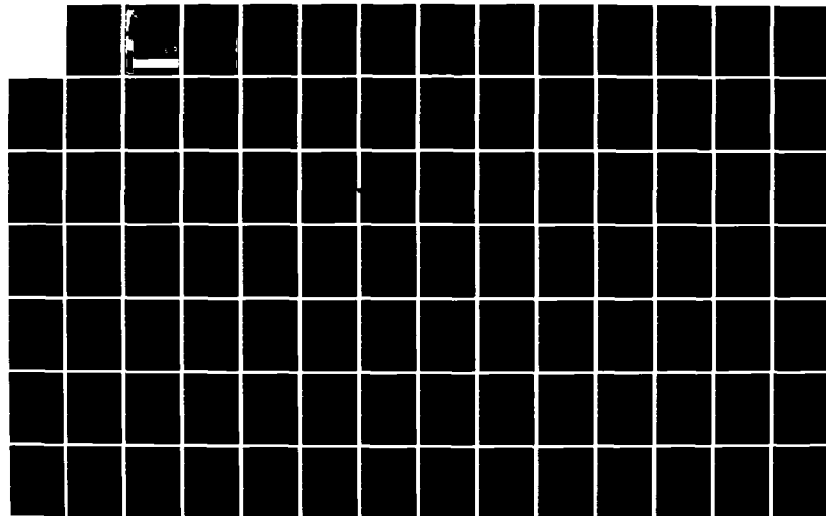
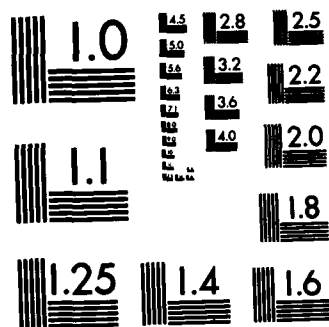


AD-A123 289 TRAINING EVALUATION OF AN AUTOMATED TRAINING SYSTEM FOR 1/2  
AIR INTERCEPT CON. (U) CANYON RESEARCH GROUP INC  
WESTLAKE VILLAGE CA M E MCCAULEY ET AL. DEC 82  
UNCLASSIFIED CRG-TR-82-015 NAVTRAQUIPC-81-C-0055-1 F/G 5/9 NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

(12)



Technical Report NAVTRAEQUIPCEN 81-C-0055-1

AD A 123289

TRAINING EVALUATION OF AN AUTOMATED  
TRAINING SYSTEM FOR AIR INTERCEPT  
CONTROLLERS

Michael E. McCauley  
Robert W. Root  
Frederick A. Muckler

Canyon Research Group, Inc.  
741 Lakefield Rd., Suite B  
Westlake Village, CA 91361

December 1982

FINAL REPORT APRIL 1981 - JULY 1982

DoD Distribution Statement

Approved for public release;  
distribution unlimited.

DTIC  
ELECTE  
JAN 11 1983  
A

DTIC FILE COPY

NAVAL TRAINING EQUIPMENT CENTER  
ORLANDO, FLORIDA 32813

83 01 11 014

NAVTRAEQUIPCEN 80-C-0055-1

GOVERNMENT RIGHTS IN DATA STATEMENT

Reproduction of this publication in whole or in part is permitted for any purpose of the United States Government.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                              | READ INSTRUCTIONS<br>BEFORE COMPLETING FORM |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------|---------------------------------------------|
| 1. REPORT NUMBER<br>NAVTRAEQUIPCEN 81-C-0055-1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 2. GOVT ACCESSION NO.<br>AD-A12328                                           | 3. RECIPIENT'S CATALOG NUMBER               |
| 4. TITLE (and Subtitle)<br>TRAINING EVALUATION OF AN AUTOMATED<br>TRAINING SYSTEM FOR AIR INTERCEPT CONTROLLERS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 5. TYPE OF REPORT & PERIOD COVERED<br>FINAL REPORT<br>April 1981 - July 1982 |                                             |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 6. PERFORMING ORG. REPORT NUMBER<br>CRG-TR-82-015 ✓                          |                                             |
| 7. AUTHOR(s)<br>Michael E. McCauley, Robert W. Root, and<br>Frederick A. Muckler                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 8. CONTRACT OR GRANT NUMBER(s)<br>N61339-81-C-0055                           |                                             |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Canyon Research Group, Inc.<br>741 Lakefield Rd., Suite B<br>Westlake Village, CA 91361                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 10. PROGRAM ELEMENT, PROJECT, TASK<br>AREA & WORK UNIT NUMBERS<br>7796-5P1   |                                             |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Naval Training Equipment Center<br>Code N-711<br>Orlando, FL 32813                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 12. REPORT DATE<br>December 1982                                             |                                             |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 13. NUMBER OF PAGES<br>128                                                   |                                             |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)<br>DCASMA<br>Los Angeles Region, Van Nuys Branch<br>6230 Van Nuys Blvd.<br>Los Angeles, CA 91408                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 15. SECURITY CLASS. (of this report)<br>UNCLASSIFIED                         |                                             |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 15a. DECLASSIFICATION/DOWNGRADING<br>SCHEDULE                                |                                             |
| 16. DISTRIBUTION STATEMENT (of this Report)<br><br>Approved for public release; Distribution unlimited                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                              |                                             |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                              |                                             |
| 18. SUPPLEMENTARY NOTES                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                              |                                             |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number)<br>VTAG<br>Air Intercept Controller<br>Automated Training System<br>Automated Speech Technology<br>Computer Speech Recognition<br>Interactive Simulation<br>Training Effectiveness Evaluation<br>Human Factors Engineering                                                                                                                                                                                                                                                                                                                          |                                                                              |                                             |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number)<br>→ This report describes the evaluation of an experimental prototype training system called the Air Controller Exerciser (ACE). The system was developed to demonstrate the use of new technologies for training, including computer speech recognition and generation, videodisc, automated instruction, automated performance measurement and syllabus control, and speech-interactive simulation. Combining these emerging technologies promises to reduce the requirements for instructor manpower and other training support personnel (cont) |                                                                              |                                             |



PREFACE

The authors are in debt to the personnel at the Navy Fleet Combat Center Pacific (FLECOMBATRACENPAC) for their active participation in this project. Special thanks to three instructors assigned to ACE, OSC Dick Meyer, OSC Jim McPhearson, and OSI Larry Nielsen. The following FLECOMBATRACENPAC personnel also contributed substantially to the project: LCDR Tana, LT Dehner, LT Eckrich, Mr. Spencer, OSSC Barney, OSSC Vincent, OSSC Larson, OSC Sabre, OSC Fife, OSMC Cross, OSSC Stites, and the remainder of the staff of Code 31. In addition, we are grateful to the AIC trainees who contributed considerable time and effort in their interaction with ACE, and the personnel of FLECOMBATRACENPAC Program Control and Evaluation who supported the transfer of training simulations.

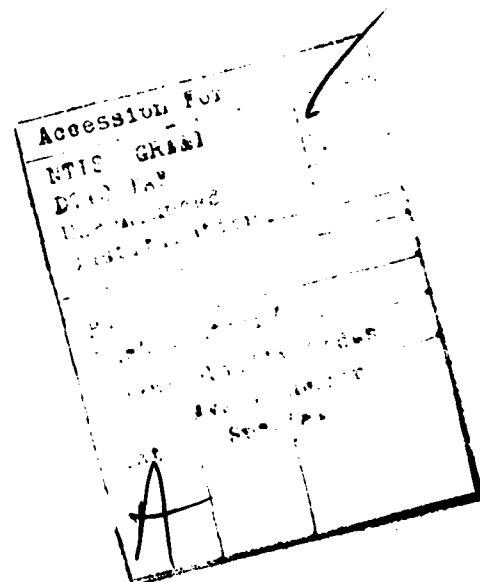
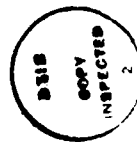
Encouragement and support were given by the Naval Training Equipment Center (NAVTRAEQUIPCEN) Scientific Officer and Project Engineer, Dr. Robert Breaux and Mr. Joe Puig.

Logicon, Inc. employees were helpful in providing information about ACE, particularly Mr. Michael Grady, Dr. Robin Halley, Mr. Ron Anders, Ms. Gail Slemon, and Ms. Mary Hicklin.

We are thankful to Dr. Sam Bowser for his participation in the early stages of the evaluation, including the development of the Exit Questionnaire.

Our consultant, Dr. Douglas Chatfield of Behavioral Evaluation and Training Systems (BETS), was invaluable for his professional approach to observational analysis and assessment of the system's temporal characteristics.

Thanks to Rosemary Wescott for responding to a series of requests by the authors for "final" corrections to the manuscript.



## TABLE OF CONTENTS

| <u>Section</u> |                                                                     | <u>Page</u> |
|----------------|---------------------------------------------------------------------|-------------|
| I              | INTRODUCTION.....                                                   | 7           |
|                | Background.....                                                     | 7           |
|                | New Technologies for Training.....                                  | 7           |
|                | The AIC Task.....                                                   | 10          |
|                | AIC Training.....                                                   | 11          |
|                | ACE Training Overview.....                                          | 11          |
|                | ACE Evaluation Objectives and Constraints.....                      | 17          |
| II             | ANALYSES OF ACE TRAINING FEATURES AND<br>DESIGN.....                | 19          |
|                | Overview.....                                                       | 19          |
|                | Automated Instruction.....                                          | 19          |
|                | Simulation.....                                                     | 26          |
|                | Speech System.....                                                  | 27          |
|                | Instructor's Role.....                                              | 31          |
|                | Temporal Factors in ACE Training.....                               | 33          |
|                | User Acceptance.....                                                | 36          |
| III            | SPEECH RECOGNITION ACCURACY EVALUATION.....                         | 37          |
|                | Introduction.....                                                   | 37          |
|                | Procedure.....                                                      | 39          |
|                | Digit Error.....                                                    | 41          |
|                | Sample Description.....                                             | 41          |
|                | Results.....                                                        | 42          |
| IV             | STUDIES OF PERFORMANCE MEASUREMENT AND<br>TRANSFER OF TRAINING..... | 51          |
|                | Performance Measurement Validation Study.....                       | 51          |
|                | Transfer of Training (TOT) Test.....                                | 55          |
| V              | COST-BENEFIT ANALYSIS.....                                          | 63          |
|                | The Cost Benefits of ACE.....                                       | 63          |
|                | Principal Cost Components.....                                      | 64          |
|                | Reducing Automated Instruction Features.....                        | 68          |
| VI             | CONCLUSIONS AND RECOMMENDATIONS.....                                | 71          |
|                | Conclusions.....                                                    | 71          |
|                | Recommendations.....                                                | 73          |
|                | REFERENCES.....                                                     | 75          |
|                | LIST OF ACRONYMS.....                                               | 79          |

# TABLE OF CONTENTS (Continued)

| <u>Section</u>                                                                | <u>Page</u> |
|-------------------------------------------------------------------------------|-------------|
| APPENDIX A: The Exit Questionnaire.....                                       | 81          |
| APPENDIX B: Performance Measurement<br>Validation study:<br>Grading Form..... | 85          |
| APPENDIX C: Sample Grading Sheet for<br>TOT Test.....                         | 89          |
| APPENDIX D: TOT Test Protocol.....                                            | 91          |

## LIST OF ILLUSTRATIONS

| <u>Figure</u> |                                                                                        |    |
|---------------|----------------------------------------------------------------------------------------|----|
| 1             | ACE Computer Hierarchy.....                                                            | 12 |
| 2             | Error Amplification Pathway to Syllabus<br>Control.....                                | 24 |
| 3             | Three Categories of Training Errors can be<br>Caused by Speech Recognition Errors..... | 50 |
| 4             | TOT Study: Mean Performance Scores by Type<br>of Intercept and Training Group.....     | 60 |
| 5             | Current AIC and ACE Alternative System<br>Training Configuration.....                  | 65 |

## LIST OF TABLES

| <u>Table</u> |                                                 |    |
|--------------|-------------------------------------------------|----|
| 1            | ACE Training Vocabulary.....                    | 14 |
| 2            | Time Allocation in a Sample of ACE Training.... | 33 |
| 3            | A Typical Speech-Interaction Scenario.....      | 37 |

LIST OF TABLES  
(Continued)

| <u>Table</u> |                                                                                                                 | <u>Page</u> |
|--------------|-----------------------------------------------------------------------------------------------------------------|-------------|
| 4            | Sample Size (Number of Transmissions) by Syllabus Level and Transmission Type (With/Without Digit Strings)..... | 42          |
| 5            | Recognition Accuracy of the DP-100 and the Speech Understanding Software (SUS).....                             | 43          |
| 6            | Speech Recognition Errors by Type.....                                                                          | 44          |
| 7            | Recognition Accuracy for "Accepted" and "Rejected" Transmissions.....                                           | 44          |
| 8            | Error Types for 1440 Transmissions Containing Digit Strings.....                                                | 45          |
| 9            | Error Types for 325 Transmissions Containing No Digit Strings.....                                              | 46          |
| 10           | Number of Cases Correct, Rejected, and Misrecognized for Single Digits and Compound Numbers.....                | 47          |
| 11           | Single Digit Errors in Heading Calls, by Position.....                                                          | 48          |
| 12           | Transmission Recognition Accuracy by Syllabus Level for Eight Trainees.....                                     | 51          |
| 13           | "Pass" Percentage Figures for Specific PMVS Measured During PMV Validations Study.....                          | 54          |
| 14           | Correlation Matrix of Instructor Ratings and School Performance.....                                            | 55          |
| 15           | TOT Test Scores for Two Training Groups by Type of Intercept.....                                               | 59          |
| 16           | Analysis of Variance Summary Table: TOT Results.....                                                            | 61          |
| 17           | Personnel Loading for Current and ACE Systems...                                                                | 66          |
| 18           | Comparative 10-Year Operating Cost.....                                                                         | 67          |
| 19           | Comparative 10-Year Operating Costs: Limited ACE System.....                                                    | 69          |

## SECTION I

## INTRODUCTION

## BACKGROUND

For nearly ten years the Naval Training Equipment Center (NAVTRAEQUIPCEN) has conducted a series of research and development projects on the use of automated speech technology (AST) in training systems. Two prototype training systems have been developed under this program, one for precision approach radar controllers and the other for air intercept controllers. Both of these systems were developed by Logicon, Inc., Tactical and Training Systems Division. The precision approach radar training system (PARTS; also known as Ground Controlled Approach-Controller Training System (GCA-CTS)) was completed in 1979 and subsequently evaluated at the Air Traffic Control School, Naval Technical Training Center (McCauley and Semple, 1980). The air intercept controller (AIC) training system was completed approximately two years later. It has been known by two names also, Air Controller Exerciser (ACE) and the Prototype Air Controller Training System for AICs (PACTS AIC). The present study, conducted at the Fleet Combat Training Center Pacific (FLECOMBATRACENPAC), dealt with the cost and training effectiveness of the use of computer speech recognition in training, using ACE as the vehicle for the evaluation.

Automated speech recognition and synthesis provide the potential to reduce the instructor's workload in the training of verbal tasks by automating many of his functions. At the extreme, a system employing these technologies would be capable of providing "instructorless" training. When AST is combined with other emerging technologies, it has the potential to provide consistent, effective training while reducing: (1) instructor manpower requirements; (2) requirements for other training support personnel; and (3) training time. Clearly, these reductions in resource requirements also imply cost reduction. This report will deal with both training effectiveness and cost effectiveness because of their combined relevance for decisions about procurement of operational training systems.

## NEW TECHNOLOGIES FOR TRAINING

ACE included several new technologies which may be applicable to automated training systems. A brief review of these technologies follows.

**AUTOMATED SPEECH TECHNOLOGY.** AST (also called voice technology) refers to both computer speech recognition and speech generation.

Speech Recognition. This technology promises to usher in a new era for the interaction between humans and computers. Reviews of the current state-of-the-art in AST can be found in Dickson and Martin (1979) and Lea (1980). Some observers have predicted that AST will have a large impact on society in the near future because it will provide a natural form of

communication between humans and computers (Evans, 1979; Toffler, 1980). The NAVTRAEQUIPCEN has been actively pursuing the practical applications of AST for automated training systems since 1972 (Breux, Curran, and Huff, 1978; Breux and Goldstein, 1975). An overview of the potential applications of AST to training systems is given by Cotton and McCauley (1982).

One application of AST is in the training of tasks which have a significant verbal-interaction component, such as air traffic control. Both the PARTS and the ACE systems were designed on this premise, but there are other applications of AST for training. For example, a maintenance task that involves little or no verbal performance from the operator could make use of AST in a training system as a communication medium between the trainee and the computer. While the NAVTRAEQUIPCEN projects have emphasized the use of AST for training verbal tasks, there are many other types of tasks for which AST could be applicable, such as hands/eyes busy or tasks involving mobile human-computer interaction without the constraint of being tied to a terminal/keyboard.

AST in training systems enables automated interactive simulation based on the trainee's verbal output, as well as automated measurement of his verbal performance. Previously, verbal performance, such as in air traffic control, was amenable only to instructor grading. Furthermore, in training that requires an interaction between the trainee and the controller of the simulated vehicle (a pseudo pilot), AST enables such pseudo pilots to be replaced by computer models of the pilot/aircraft (McCauley and Cotton, 1982). Similarly, instructor models can aid the human instructor and reduce his workload by enabling verbal interchange between the trainee and the automated training system.

These potential resource-reduction benefits which may be derived from the use of AST in training systems must be weighed against their potential risks. Most of the risks stem from the current technological limitations of AST, such as speech recognition accuracy, speech stylization requirements, limited vocabulary size, and the need to sample each individual's speech characteristics. The latter requirement characterizes speaker-dependent systems, which represent the current state-of-the-art in AST. A few speaker-independent systems have been developed which can recognize General American without pretraining. However, these systems tend to have extremely limited vocabulary sizes.

ACE used the most sophisticated recognition system on the market at the time, the Nippon Electric Company NEC DP-100. This device is a connected speech recognizer (CSR), which eliminates the requirement of isolated word recognizers (IWR) to pause between each specified vocabulary item. The DP-100 allows up to five vocabulary items to be spoken without pausing.

Speech Generation. Speech generation is advancing rapidly and is now well within the state-of-the-art. The two major techniques for producing speech by machine are: (1) to record, digitize, and playback an actual human

voice (digitized speech) or; (2) to synthesize and concatenate speech from a set of machine-generated phonemes or words (synthesized speech). Digitized speech sounds very normal but it requires that each word or phrase be prerecorded. Currently, synthesized speech tends to sound somewhat unnatural but it requires no prerecording. Summaries of speech generation technology have been given by Kaplan (1980) and by Michaelis and Wiggins (1981) and Brightman and Crook (1982). Applications of this technology in commercial aviation have been studied by Simpson and Williams (1980).

Speech generation in ACE was accomplished by two different methods. Synthesized speech was produced by a Votrax VS-6. Digitized speech was produced by a Logicon VoxBox. These two methods generated perceptibly different voices which helped the ACE trainee to determine "who" was talking to him during a simulated intercept.

The combination of speech recognition and speech generation allows communication to flow both directions across the man-computer interface. ACE demonstrates such an application for training systems, but much more complex interchanges undoubtedly will be seen in the future when the trainee and computer are able to dialogue.

**OTHER NEW TECHNOLOGIES.** New technologies other than AST were included in ACE and represent a secondary set of training system features to be evaluated. Some of these emerging technologies include: automated performance measurement, videodisc, and interactive simulation of the AIC's task.

Automated Performance Measurement. ACE had a set of over 80 Performance Measurement Variables (PMVs) which defined the scoring criteria for various classes of performance to be exhibited during AIC training. The behaviors that were monitored were both manual and verbal actions, which were checked for accuracy and time of occurrence. A subset of the PMVs were active for any specific portion of training. On a particular problem, each PMV was worth a total of 100 points, and a criterion score for passing was designated.

The major objectives of automated performance measurement in ACE were to provide automatic performance feedback to the trainee and to provide performance information to the automated syllabus control and record keeping subsystems, as well as to the human instructor.

Videodisc. A videodisc system (MCA Discovision) was included in ACE as a demonstration of this new medium for presenting audiovisual information in training. The videodisc employed both still-frame picture and moving audiovisual presentations to support instruction of the AIC task. It was integrated with more traditional computer assisted instruction (CAI) presented on a CRT.

Speech-Interactive Simulation. The AIC trainee was able to practice intercepts tasks on a simulated Naval Tactical Data System (NTDS) UYA-4 console, designed specifically for ACE. This practice console was called the Training Enhancement Console (TEC), and included simulations of the radar display and the NTDS program for the console operation. The simulated air intercept scenarios were based on the interaction between the AIC trainee and various computer models of his task environment, such as pilot/aircraft models and simulated bogeys. The trainee gave both verbal and manual inputs and the system responded with simulated speech outputs and changes in the visual display.

#### THE AIC TASK

The AIC's combat task may be summarized as providing information about the tactical situation in general, and threat aircraft in particular, to friendly fighter aircraft and surface combat coordinators via communication links, including voice transmission. The AIC also may provide information about tanker join-ups for air refueling, or he may provide control services for two friendly aircraft practicing air intercepts. Examples of typical subtasks performed by the AIC include: determine nearest collision intercept (NCI), conduct a nearest collision intercept conversion (NCIC) in appropriate circumstances, execute a tanker join-up (also called rendezvous), and perform these tasks using the Navy Tactical Data System (NTDS). In addition, the AIC must communicate with a number of tactical personnel using proper Navy phraseology. These personnel include the ships weapons coordinator (SWC), and the combat air patrol (CAP). The AIC's information comes from his NTDS console displays and from radio communications. His actions include verbal communications (radio transmissions) and manual operation of the NTDS console controls. These actions must be taken promptly with respect to the dynamic spatial and temporal relationship of the tactical elements shown on the NTDS display.

Air intercept controller training at FLECOMBATRACENPAC is divided into two basic groups: conventional (i.e., Radar-only, non-NTDS) and Naval Tactical Data System (NTDS). The ACE system was directed only at NTDS qualification.

The AIC's task is described in the mission of the school at FLECOMBATRACENPAC as follows: "...to analyze and transmit to the interceptor aircrew the necessary tactical information required to perform their mission in both the combat and training environments in accordance with current fleet doctrine and directives; advise command during the planning stages of anti-air warfare on the capabilities, limitations, and employment of current U.S. and threat aircraft and associated weapons systems in accordance with current tactical manuals and threat intelligence; and, inform command of mission progress and aircraft status on a continuing basis using standard Navy phraseology and appropriate relativity code words" (U.S. Navy, 1980).

## AIC TRAINING

At the time of this evaluation the length of the NTDS AIC course at FLECOMBATRACENPAC was 40 days. Normally there were four trainees per class, and a new class began every two weeks. The first week was spent in classroom lecture. The next two weeks of training were spent in a simulator using pseudo-pilots and an NTDS radar console. This simulation is called "Synthetics."

In Synthetics, the trainees practiced intercepts on the NTDS console with an instructor looking over their shoulders, giving instruction and providing feedback on their performance. A trainee's verbal commands are transmitted to the Problem Control and Evaluation (PC&E) room where the pseudo-pilot hears the command and responds on a keyboard to simulate the action of the pilot/aircraft. The need for human pseudo-pilots to support AIC training is costly and it introduces potential uncontrolled variability in the responsiveness of the pilot/aircraft. This two week Synthetics phase of AIC training was the test bed for ACE.

Following completion of Synthetics, the trainees proceeded to "lives," in which they controlled actual aircraft (both contracted Lear jets and actual Navy Fighter aircraft on training missions) in a special sector over the Pacific Ocean off of San Diego. Successful performance in "lives" marked completion of the trainee's training as an AIC.

Prerequisites for AIC training were as follows: all candidates for AIC training must have a minimum of one year recent operational Combat Information Center (CIC) experience in a combatant ship, the last six months of which must include performance in the anti-air warfare (AAW) environment; candidates must be ordered to, or serving in, a unit requiring AICs; and personnel and eligible ratings include: operations specialists (OS), E-5 and above.

## ACE TRAINING OVERVIEW

A schematic representation of the ACE system, is given in Figure 1. The entire ACE system was housed in a single room, including the following: Student Station (videodisc monitor and keyboard/visual display unit (VDU)), Training Enhancement Console (TEC; a simulated AN/UUA-4(V10) NTDS console), three Data General Eclipse computers, Instructor's Station (three keyboards/VDUs), an NEC DP-100 speech recognition unit, and a Printronix printer.

Trainees selected for ACE spent nine days training on the system while their fellow classmates continued normal training in Synthetics. One of the three ACE instructors was assigned exclusively to ACE each day, and spent nearly the entire time in the same room with the trainee. The instructors provided direct instruction or intervention only when requested by the trainee or when ACE failed to perform adequately.

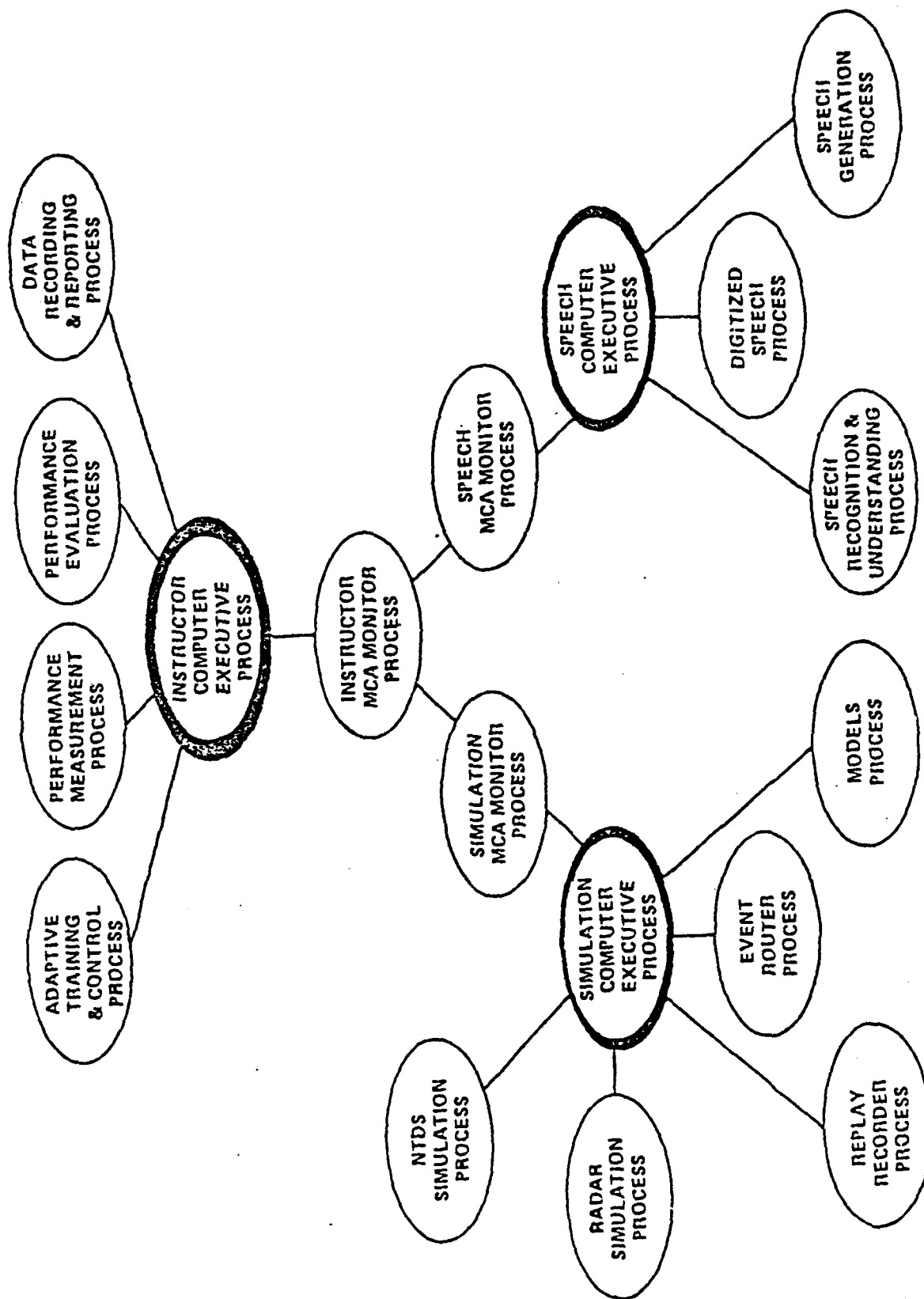


Figure 1. ACE Computer Hierarchy  
(from Clark, Halley, Regelson, Slemmon, and VerSteeg, 1979, p:28)

The ACE trainee was given interactive instruction on the Student Station VDU, supplemented by demonstrations on the videodisc monitor. After an instructional segment, the trainee was referred to the adjacent TEC for initialization of a set of graded practice problems. The automated instruction guided the trainee through the eight levels of the ACE syllabus using sequences of these instruction or practice modes. The instructional segments were termed "Interactive Teaching" (IAT). Two types of practice modes were used - commented practice (CP) and free practice (FP). Both types of practice were graded, but the problem froze in the CP mode if too many errors were made. The FP mode was a simulation of an intercept problem, followed by the presentation of performance feedback to the trainee.

The eight syllabus levels in ACE were as follows: (1) System Use and Pretest; (2) Basic Skills; (3) Basic Intercept; (4) Hostile aircraft Complications; (5) Stranger Reports; (6) Rendezvous; (7) Problems Encountered in Air Intercept Control; and (8) Training Set-Ups. Each of the eight levels was divided into units, which were formed on the basis of a topic or cluster of skills that could be taught in a series of related lessons (Granberry, Halley, Kerry, King, and Regelson, 1981a).

Throughout his training, the ACE trainee received information from both visual displays (e.g., alphanumeric data and simulated radar) and audio presentation (e.g., speech communications from simulated tactical personnel). He answered discrete questions by means of the student-station keyboard, and performed his intercepts using both voice commands and manual actions (e.g., discrete keying and position tracking with a trackball).

The vocabulary used in ACE is given in Table 1. The AIC trainee had to learn how and when to use this vocabulary to accomplish his control task. It should be noted that the most frequently-used communications dealt with relative positions of aircraft, including a series of numbers (digit strings), such as "Bogey, two four zero, twenty five" (240 degrees, 25 miles).

The ACE curriculum was enlarged to include some of the instructional material presented in the first week of classroom lecture, as well as some advanced material not presented in Synthetics. These differences in curricula between ACE and Synthetics influenced the transfer of training study, which will be discussed later.

Documentation was provided to support ACE training. The Instructor's Guide (Granberry, et al., 1981a) was particularly useful to both ACE instructors and trainees because it defined the system functions, the syllabus subdivisions, and explained guidelines and procedures to be followed. The Student Guide (Granberry, Halley, Kerr, King, and Regelson, 1981b) provided an introduction to ACE and a description of the system for the trainee.

TABLE 1. ACE TRAINING VOCABULARY  
(From Granberry, et al., 1982)

- 
0. Zero
  1. One
  2. Two
  3. Three
  4. Four
  5. Five
  6. Six
  7. Seven
  8. Eight
  9. Nine
  10. Ten
  11. Eleven
  12. Twelve
  13. Thirteen
  14. Fourteen
  15. Fifteen
  16. Sixteen
  17. Seventeen
  18. Eighteen
  19. Nineteen
  20. Twenty
  21. Thirty
  22. Forty
  23. Fifty
  24. Sixty
  25. Digits (0-9) as a group
  26. 10-19 as a group
  27. 20, 30, 40, 50, 60
  28. Silver Hawk
  29. Crackerjack
  30. C/S; P/S/V XXX
  31. C/S; Port XXX
  32. C/S; Starboard XXX
  33. C/S; Vector XXX
  34. C/S; P/S/V/ XXX; For Bogey
  35. Station XXX; YY
  36. Bogey XXX; YY
  37. Bogey Tracking XXX; Speed Point X
  38. Bogey Tracking XXX; Speed Point X
  39. Bogey Tracking XXX; Speed Point X
  40. C/S(;) Mark your TACAN
  41. C/S(;) What State?
  42. Roger State XXX
  43. C/S(;) State XXX
-

TABLE 1  
(Continued)

- 
44. I Have Control of C/S
  45. C/S(;) On Station
  46. C/S(;) Breaking Away
  47. Splash/Heads-up 1/2 Bogey(s)
  48. Splash One Bogey
  49. Heads-up Two Bogey
  50. Splash One Bogey
  51. Heads-up Two Bogeys
  52. Bogey(s) Single/Multiple Altitude Y Thousand
  53. Bogey Single Altitude Y Thousand
  54. Bogeys Multiple Altitude Y Thousand
  55. Bogey Single Altitude Y Thousand
  56. Bogey Jinking Left/Right
  57. Bogey Jinking Left
  58. Bogey Jinking Right
  59. Bogey Splitting
  60. Roger; Bogey Tracking XXX
  61. Negative; Bogey XXX; YY
  62. Stranger XXX; YY
  63. Stranger Tracking XXX; Angels Y
  64. Stranger Opening
  65. C/S(;) Tighten Turn/Ease Turn
  66. C/S(;) Tighten Turn
  67. C/S(;) Ease Turn
  68. C/S(;) Radio Check Over
  69. Bogey/Fighter in the Dark
  70. Bogey in the Dark
  71. Fighter in the Dark
  72. C/S(;) My Octopus is Bent
  73. C/S(;) Emergency
  74. C/S P/S/V XXX; For Rendezvous
  75. C/S; Detach Port/Detach Starboard XXX; For Separation
  76. C/S; Detach Port XXX; For Separation
  77. C/S; Detach Starboard XXX; For Separation
  78. C/S; Detach P/S XXX; For Separation
  79. C/S; Continue XXX
  80. C/S; Breakaway XXX
  81. C/S(;) Angels Y
  82. C/S; C/S XXX; YY
  83. C/S(;) Steady
  84. C/S(;) Lost Communications Intentions
  85. Roger Lost Comm
  86. C/S; P/S/V XXX; As Bogey
  87. Roger
-

TABLE 1  
(Continued)

- 
- 88. Negative
  - 89. Say Again
  - 90. Correction
  - 91. AIC 1

NOTE: Voice data are collected for underlined phrases only (except for elements 0-27, which are all collected).

---

## ACE EVALUATION OBJECTIVES AND CONSTRAINTS

The objectives of the ACE evaluation evolved from the original Statement of Work promulgated by the NAVTRAEQUIPCEN. These objectives were multiple, and included assessment of new technologies for training, identification of the strengths and weaknesses of the prototype system, cost analyses of the overall system as well as the new technologies comprising it, derivation of principles for future system design, and recommendations for the future application of these technologies to the training of air intercept controllers. The multiplicity of objectives sometimes resulted in conflicting approaches in conducting the evaluation. The primary conflict was between the evaluation of individual technologies regarding their potential for use in future training systems and the evaluation of the ACE prototype as an effective trainer. Promising technologies (or even excellent ones) do not necessarily ensure training effectiveness in an experimental prototype system; similarly, a marginally effective prototype may obscure the actual performance capabilities of a given technology.

Evaluation criteria appropriate for an operational training system may not be appropriate for an experimental prototype system comprising high risk technologies which have been integrated for the first time. This type of training system is largely a demonstration of new technologies and their possible applications. More stringent evaluation criteria should be applied to an operational training system than to an experimental system intended to demonstrate new technology applications. Canyon's approach was to evaluate training effectiveness in a manner similar to that used for an operational training system, but, when the prototype system fell short of operational effectiveness, to provide information about how the new technologies might be integrated more effectively in the development of an operational system.

The evaluation was further constrained by the condition of the system at the time of delivery: ACE was unfinished. It needed software debugging and more development time before it was ready for an evaluation. Nevertheless, the evaluation proceeded as scheduled, apparently because no resources were available for further development.

One final constraint was that the curriculum taught by ACE was different from that taught at the AIC school. This difference made it difficult to compare the performance of trainees from the two groups. The curricula differed for two reasons: (1) Logicon's attempt to improve the training through the Instructional System Development (ISD) approach, and (2) changes in the AIC school curriculum during the period of ACE development.

## SECTION II

### ANALYSES OF ACE TRAINING FEATURES AND DESIGN

#### OVERVIEW

**OBJECTIVES.** The objectives of these analyses were to estimate how the following factors contributed to the training effectiveness of ACE: automated instruction; problem simulation; speech system; instructor's role; and temporal factors. Issues of user acceptance also were addressed.

**APPROACHES.** Information for the analysis of training features and system design was obtained by observation, interview, and questionnaire survey.

Observation. The Canyon project team observed trainees interacting with ACE for a period of six months. Each ACE trainee was observed at several points during training. Additionally, Dr. Douglas Chatfield, a consultant on the project, performed an in-depth observational analysis of one trainee. The job/role of the instructor in ACE also was observed.

Trainee Interviews. ACE trainees were interviewed formally after completing the course. In addition, informal discussions were held periodically with each trainee during his training on ACE.

Instructor Interviews. The three ACE instructors were interviewed intensively during the evaluation period. Their concepts and opinions were an important influence on the ACE evaluation. In addition to the ACE instructors, interviews were conducted with the officers in charge of AIC training and with the Senior Chief who headed the instructional staff.

Exit Questionnaire. A brief questionnaire was developed by Canyon to assess the attitudes and opinions of the ACE trainees following their experience with the system. This Exit Questionnaire (see Appendix A) contained 14 questions pertaining to ACE and its constituent technologies and features.

#### AUTOMATED INSTRUCTION

The analyses of ACE automated instruction focused on the system features intended to reduce (or replace) the instructor's workload. The features which were analyzed included instructional modes, videodisc, computer assisted instruction (CAI), adaptive syllabus control, and performance measurement and feedback.

**INSTRUCTIONAL MODES.** Three instructional modes were used in ACE: interactive teaching (IAT), commented practice (CP), and free practice (FP). As the trainee progressed through the eight-level syllabus, he was continually shifted among these three modes of operation by the ACE courseware.

INTERACTIVE TEACHING (IAT). The IAT mode was used for teaching. Video disc presentations and CAI were used to present new concepts and procedures to the trainee. These two technologies will be discussed individually. In general, the IAT mode was well conceived and performed adequately. Continued development and revision of the instruction would be necessary in an operational version of ACE.

Physically separating the IAT portion of ACE from the remaining simulation/practice portion was suggested by an ACE instructor. This design would enable two trainees to use the system simultaneously - one on IAT and one practicing intercepts.

COMMENTED PRACTICE (CP). The CP mode of ACE was a disaster. The concept of "errorless training" was good: prevent the trainee from practicing errors by freezing a problem when he makes errors on the newly learned behavior, then provide instructive feedback and allow him to try again. In operation, however, the CP mode was completely unacceptable because it so frequently (if not usually) froze the problem needlessly, and then forced the trainee to start the entire intercept again. To an AIC, an intercept is the unit of work. Preventing the ACE trainees from completing an intercept was considered to be a form of punishment. The restarts were particularly unpopular (and unwarranted) when they were caused by a series of speech recognition errors, which were interpreted by ACE as performance errors. Both instructors and trainees suggested that the CP intercepts continue after performance feedback is provided during the freeze. They preferred some discontinuity rather than completely restarting the problem.

The CP mode could be a viable technique if used judiciously. A successful CP mode could be achieved only by an accurate performance measurement system, immune to speech recognition errors. Until such measurement accuracy is attained, attempts to apply techniques such as the CP mode should be avoided. Unwarranted punishment of the trainee erodes user acceptance.

The ACE instructors frequently preempted the CP mode and advanced the trainees to the associated FP mode. Their rationale was that too much training time was wasted in CP by unnecessary freezes and restarts.

FREE PRACTICE (FP). The FP mode was the strength of ACE. It allowed the trainee to practice his skills by providing appropriate intercept scenarios. ACE probably would have been more effective if more FP intercepts had been provided. Some old adages are proven with time, such as "practice makes perfect." In the case of ACE, "FP makes perfect" would have been a good system design principle. Unfortunately, the number of intercepts per unit time was limited in ACE. This point will be discussed in more detail in this section, under the heading "Temporal Factors in ACE Training."

The FP intercepts were based on the speech-interactive simulation, followed by the presentation of performance feedback. The weak links of the FP mode were: (1) simulation/scenario discrepancies caused by speech recognition errors; and (2) lengthy, and often fallacious, performance feedback presented to the trainee. Both of these issues will be discussed in more detail.

ACE instructors suggested that some FP intercepts be practiced as a continuing series, rather than as discrete events. In some portions of Synthetics and in "lives," the AIC trainee has continuous control of two friendly aircraft practicing intercepts. The trainee cannot stop after one intercept. Some ACE trainees had difficulty in "lives" because they tended to relax and relinquish control at the end of an intercept. Real aircraft keep flying. An operational version of ACE would require practice in continuous control through a series of intercepts.

VIDEODISC. The videodisc in ACE was added to the design rather late in the development of the system (Grady, 1982). Furthermore, it was one of the earliest discs cut by MCA Discovision for use in training. Many lessons were learned by the ACE developers from this effort (Halley 1981; reproduced in Grady, 1982). Videodisc appears to be a powerful medium for presenting audiovisual information. Both still frame and motion sequences were used in ACE to motivate, illustrate a point, and teach procedures.

Videodisc was a successful instructional medium in ACE, but not without its problems: inappropriate material sometimes was presented, the audio was out of sync with the video, or entire segments of instruction presented in a few seconds. These problems apparently stemmed from hardware bugs that may not be indicative of current videodisc technology.

One of the strengths of videodisc for training is the capability for rapid access to stored audiovisual information. ACE trainees, however, were given little control over information access. The videodisc control was associated with the courseware for each lesson.

The videodisc picture of an AIC's console display often provided inadequate visual resolution of critical features. Perhaps such examples could be given on the actual console display, rather than on the videodisc monitor.

Further development of videodisc instruction would be a valuable contribution to training technology, in ACE as well as many other applications. One issue to be resolved about the use of videodisc technology in Navy training is its cost-effectiveness relative to the anticipated curriculum changes. A videodisc master is expensive; how soon will it be out-of-date in a Navy training system?

Computer Assisted Instruction (CAI). Interactive instruction at the ACE student station was accomplished through the use of a standard computer-terminal keyboard and a visual display unit (VDU). The CAI generally was good. Some ACE trainees (usually E-5s) thought that the

material was too detailed and elementary. For example, "port" and "starboard" are defined at one point in a CAI session. Is this necessary for Navy E-5s? One good feature of the CAI was that the menu driven system gave some degree of control to the trainee to continue, review, take coffee breaks, etc.

CAI is not a particularly new technology. It is certainly a good process for teaching the knowledge required to execute a skill such as air intercept control. The "goodness" of the instruction depends on the strength of the instructional design team and regular input from subject matter experts (SMEs).

The instruction in ACE and Synthetics emphasized different aspects of the AIC's job. This does not reflect CAI technology; rather, it represents the philosophy of the design team and their SME. The emphasis in ACE was on combat intercepts and communications with the combat team, such as the ships weapons coordinator (SWC). By contrast, the emphasis in Synthetics was on providing support for practice intercepts with two friendly aircraft. The ACE emphasis could be considered more appropriate to the AIC's job in combat, while the Synthetics emphasis was more typical of the AIC's role during peacetime. Decisions about instructional emphasis must be considered carefully when committing the program to software/courseware and CAI. Software-based instructional systems have more inertia (and require more money to change) than human instructors.

**SYLLABUS CONTROL.** The issues of syllabus control concern individualized instruction and its implementation. The ACE documentation often uses the term "adaptive" in referring to the ACE curriculum. Whether or not ACE was an adaptive system may be a matter of definition. According to our expert consultant, ACE was not adaptive (see Chatfield and Gidcumb, 1977; Chatfield, Klein and Coons, 1981; Chatfield, Marshall and Gidcumb, 1979).

ACE did have two features that were intended to tailor the instruction to the progress and capabilities of the individual trainee: "challenges" and remediation. In the end, ACE would have performed better if it had been simpler. The challenge and automated remediation features detracted from the system's performance.

**Challenges.** ACE trainees were given the opportunity to exempt portions of the instruction by challenging the system. This nice idea was rendered inoperative because the vocabulary required to execute the test problem had not been sampled by the speech system or practiced by the trainee. A further drawback to the challenge procedure was that the combination of speech recognition errors and frequent invalid performance measurement errors made it difficult, if not impossible, to pass the challenge test.

In practice, trainees were kept moving through the ACE syllabus in a timely manner through the use of instructor overrides to the syllabus control.

Automated Remediation. Suggesting a review to a student having difficulty is clearly a good concept. The automated remediation in ACE was intended to do just that. Unfortunately, the abundance of performance errors detected by the PMVs meant that ACE trainees frequently were remediated without cause. Undeserved remediation wastes time and leads to motivational problems.

The ACE trainees and instructors also suggested that remediation need not repeat the speech sampling process for those lessons being repeated. If the trainee has been achieving good recognition accuracy with those vocabulary items, why begin the remediation process with a lengthy procedure of unneeded speech sampling? Perhaps the trainee could be shown a confusion index of vocabulary items, and encouraged to provide new samples of problem items (Cotton and McCauley, 1982). A simpler solution would be to simply suggest to the trainee that he "retrain" any vocabulary items that he deems necessary.

Although the ACE documentation refers to the "extensive diagnosis" and "sophisticated prescriptive" capabilities of ACE, it was difficult to observe these features because the ACE instructors found it necessary to manually control syllabus progress. Apparently, the combined problems of speech recognition and PMV errors did not provide sufficiently accurate input to the syllabus control (diagnosis and prescription) subsystem to enable it to function properly.

The remediation process in ACE created a considerable workload for the instructors. They were frequently called upon to override the system to prevent an unwanted remediation. This type of problem is not indicative of the state-of-the-art in training technology. Undoubtedly, it reflects the previously stated fact that ACE did not have the benefit of a test, evaluate and revise (TEAR) cycle, prior to this evaluation.

Sources of Syllabus Control Problems. The two major contributors to faulty syllabus control were performance measurement and speech recognition. The automated performance measurement system needed further development. Besides having its own problems, the PMV system often was given inaccurate input from the speech understanding system. A critique of the actual syllabus control logic in ACE is not possible because it was inundated with inaccurate input from the PMVs (which received erroneous input from speech recognition) as shown in Figure 2.

Again, ACE attempted to demonstrate a combination of high-risk technologies in an experimental prototype system. This is the purpose of Research and Development (R&D). In an operational system, however, it might be prudent to develop a simple linear syllabus, without automated syllabus control. The instructor (and possibly the trainee) could be given control over acceleration or remediation decisions. This conservative design strategy enables a voice-interactive training system to be effective from the outset. Later addition of automated syllabus control is always possible. It is this type of lesson that can be learned from R&D

prototypes such as ACE. At the present time automated syllabus control appears to be unwise in a system which relies heavily on accurate speech recognition and automated performance measurement.

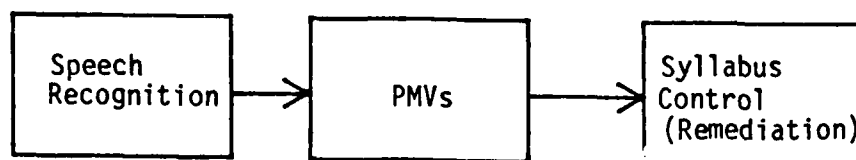


Figure 2. Error Amplification Pathway to Syllabus Control

PERFORMANCE MEASUREMENT AND FEEDBACK. The automated performance measurement in ACE was accomplished by 84 performance measurement variables (PMVs). Examples of the PMVs were given earlier. They were based on the timing and accuracy of a verbal or manual action with respect to the Radar simulation on the NTDS console display. An arbitrary number of points was subtracted from 100 when the trainee's actions were late or inaccurate.

The instructional staff of FLECOMBATRACENPAC reported that the highest priority "fixes" necessary to improve ACE were the PMVs and speech recognition accuracy. The Canyon observers concurred, and projected that better performance measurement would generate large gains in training effectiveness.

Speech recognition errors were only part of the problem. Accurate speech recognition is necessary but not sufficient for precise performance measurement in ACE.

The scoring process was based on timing criteria that, in many cases, needed revision. Listings of the criteria were reviewed by SMEs during ACE development, but timing factors, particularly, are difficult to assess from listings. Although some adjustments were made in the PMV criteria just prior to the evaluation, a considerable amount of improvement in the ACE performance measurement system is necessary.

Inaccurate performance measurement in ACE detracted from training effectiveness in two major ways - incorrect performance feedback and inappropriate remediation.

Performance Feedback. At the end of a free practice problem, the trainee would observe a long series of error messages on a small visual display unit which was added to the simulated NTDS console. Each error message would report the time of the infraction, i.e., how many seconds into the problem, and a statement of the error. There were several difficulties with the process. The most important problem was the large number of error messages. While the number varied, depending on speech recognition accuracy, syllabus level and (hopefully) the trainee's performance, a typical number of error messages was 25.

Trainees were confronted with such a long sequence of error messages that they usually could not ascertain the relationship between the scenario events (including their own actions) and the error messages. The time tag with each message provided only a rough key to the scenario events in question.

The typical trainee's response to the performance feedback was to learn to ignore most of it. This was adaptive behavior, under the circumstances. Too great a workload was imposed in attempting to (1) associate the time and error message with his own action (or omission), and (2) decide whether the message was valid. Trainees typically knew when they had committed an error during an intercept. When they watched the feedback, they tended to search the message sequence to find their known error(s), detect errors they were unaware of, and reject the remaining messages as system faults. The ACE instructors estimated that, on an average problem, less than half of the error messages were valid.

ACE trainees and instructors also commented that the repetition of common error messages detracted from their impact. For example, ACE might give the same error message a dozen times after one problem; each occurrence preceded by a different time. The instructors suggested that the system could amalgamate such messages by saying "several errors of the following type were detected..."

In the three-hour sample of ACE training activity obtained by Dr. Chatfield, 29% of the total time was spent performing intercepts, and 25% was spent on performance feedback. The trainee required approximately three minutes to read the feedback displays following each intercept. Of the 14 intercepts sampled during a three-hour period, all but 2 presented some incorrect or inappropriate feedback information, as judged by the trainee and (usually) the instructor.

The sources of erroneous feedback in ACE have been mentioned: overly strict criteria; faulty timing criteria; unrefined scoring rules; arbitrary passing scores; and speech recognition errors. The effects of the resulting proliferation of error messages can be summarized as follows: reduced user confidence in self and/or system; increased cognitive workload of trainees; training time wasted in presentation of error messages; faulty freezes caused in CP mode; inappropriate remediation assigned; and increased instructor workload in overriding remediation. In summary, erroneous feedback leads to ineffective training.

These problems are not fatal for ACE. They do reflect a lack of sufficient development time and resources for test and revision. In addition to at least one TEAR cycle, a performance measurement system designed with a data base structure is recommended. The ACE PMVs purportedly were designed, at least in part, on a data base structure. This allows the numerical criteria to be changed easily, without changing the software for the performance measurement process. In ACE, however, the data base was not accessible to the instructors or other FLECOMBATRACENPAC personnel, thereby preventing its continual refinement.

#### SIMULATION

TEC CONSOLE. The simulation of intercepts was a strong point of ACE. The simulated NTDS console, named the Training Enhancement Console (TEC), was a mock-up of the actual equipment minus certain features such as a mode roller. The controls were well designed, and operated better than the real thing according to some reports. The back-lit pushbuttons were cleverly used for instructional purposes during appropriate parts of the syllabus. In fact, the Canyon observers preferred the instructional effects on the TEC to the use of the adjacent videodisc to show a picture of the TEC.

The TEC display was an excellent reproduction of the actual NTDS display. In short, physical and functional reproduction of the characteristics of the NTDS console were well done.

Similarly, the NTDS program functions were simulated well. The only criticisms heard from the users were that a newer version of NTDS was now available.

INTERCEPT SCENARIOS. The intercept scenarios were considered quite acceptable by the ACE instructors. The variation in starting locations and heading of the aircraft was good, as were their general motion characteristics.

PILOT/AIRCRAFT LATENCY. The major problem with the intercept simulation was the occasional long response latency of the aircraft in response to verbal transmissions from the trainee. The delays were associated with the speech subsystem and seemed to occur for both a verbal response and a turn of the simulated pilot/aircraft. The exact source of the time lag is unknown to Canyon.

The effect of the delayed response was increased problem difficulty for the trainee. It would be similar to introducing a delay of several seconds between a steering wheel and the front wheels of a car. To make matters worse, the problem seemed intermittent. Thus, when a trainee would give a transmission requiring a turn, such as "c/s vector 320 for bogey," he could be confronted with no response for several seconds. His dilemma then, was whether to wait or repeat the transmission. Meanwhile his cognitive roadmap for the sequence of required actions was held in abeyance until he received some feedback on the state of the system with respect to

his last transmission. In short, the delay was not only a temporary nuisance, but it also hampered prosecution of the intercept and interfered with training.

Some insight into the causes of these delays is given by Grady (1982). It seems to be associated with peculiarities of the syntax of connected speech recognizers. As a simplistic example, consider two transmissions "c/s vector XXX" and c/s vector XXX for bogey." If the first transmission is given, the ACE speech understanding software apparently waits some time to determine whether a legal suffix will be appended, such as "for bogey." It is difficult for the observer to understand, however, why a delay of approximately three seconds should be created. Whatever the causes, the response of the pilot aircraft (whether verbal or visual) must occur in a realistic timeframe in order to maintain the continuity and rhythm of the simulated intercept.

PILOT/AIRCRAFT RESPONSE ERRORS. The speech system was responsible for the other major difficulty in ACE simulation, namely, failure of the pilot/aircraft to "understand" a transmission and respond accordingly. Examples of this type of problem would be to turn in the wrong direction or to turn the wrong amount in the correct direction. The frequency of occurrence of this problem was mentioned in the analysis of operational significance of digit recognition errors. Speech recognition errors in ACE tended to make the intercept more difficult. They had the effect of introducing unpredictable responses of the pilot/aircraft. Although a certain amount of such variability might be appropriate for the advanced trainee (as when a real pilot misrecognizes a transmission), this effect in ACE reflected a lack of instructional control over problem difficulty. The AIC trainee's task is difficult enough without being confronted by a capricious simulated pilot.

SIMULATION SUMMARY. Other than the problems of occasional delayed or inappropriate response of the simulated pilot/aircraft, the ACE NTDS-intercept scenario simulation was good. Trainees should have spent more time engaged in it.

#### SPEECH SYSTEM

SPEECH INPUT/OUTPUT (I/O) AND INTERACTIVE SIMULATION. Speech I/O for interactive simulation is unquestionably a sound concept. It eliminates the need for another person to close the communications loop in support of training. In the case of ACE, it entirely eliminated the need for a pseudo-pilot to control the response of the pilot/aircraft. Cotton and McCauley (1982) have reviewed speech interactive training system concepts, and suggested some general guidelines. The complexities of successful implementation are evidenced in ACE, particularly when one recognizes that no organization has had more experience or expertise in applying speech technology in training systems than the developers of ACE.

ACE SPEECH RECOGNITION ISSUES. A number of alternative design decisions confront system developers at each step in designing a complex speech-interactive system. Some of the major issues are the required accuracy, estimating/projecting current technology, speaker stylization requirements, speech data collection (SDC) techniques, recognition feedback procedures, and user training. Each of these issues will be discussed with respect to its implementation in ACE.

Recognition Accuracy. The data, reported in Section III, indicate considerable room for improvement in the recognition accuracy achieved in ACE. Only 44% of the 1765 transmissions sampled were recognized correctly. This level of accuracy would be unacceptable for an operational training system. It is interesting to note, however, that the "gut reaction" of trainees, instructors, and the Canyon observers/evaluators was that ACE performed its training function rather well, considering this low recognition accuracy.

There is no simple fix for the recognition accuracy problem. All of the following discussion issues, however, have direct bearing on recognition performance.

Speech Sampling Techniques. Speaker-dependent recognizers (over 95% of the current market) require the new user to produce a sample of his (or her) voice for each pre-defined word or phrase. Most recognizers require 4 to 10 samples of each vocabulary item. One of the advantages of the NEC DP-100 was its requirement for only one to four samples of each item. The sampled speech data are stored as a template for comparison to later speech input. When the comparison produces a sufficiently similar acoustical pattern, the speech is "recognized." It is imperative that the original speech samples be obtained under conditions identical (or nearly so) to those under which speech recognition is attempted. The shorthand way to describe this is to "sample in context." This was not done in ACE.

The ACE trainee tended to be relaxed when he was engaged in CAI, reading instructional material, or providing speech samples during SDC. When conducting an intercept, he tended to be more alert, more animated, and under greater stress. It is not surprising that his voice characteristics changed. To maximize the probability of correct speech recognition, the voice patterns used as the basis of recognition should be obtained in conditions which closely approximate the actual task. Although no hard data are available, we submit that recognition accuracy in ACE would have been substantially greater if speech sampling had been done in the context of conducting an intercept.

Prompts. Speech sampling was achieved in ACE by both visual and auditory prompts. The screen displayed a message, for example, instructing the trainee to repeat the following phrase: "Splash One Bogey," followed immediately by an auditory prompt issued by the speech generation system "Splash One Bogey." At this point, the trainee would say the phrase.

In the opinion of the evaluators, there are two drawbacks to this procedure. First, the trainee was not producing the speech sample in the context of conducting an intercept. Second, prompts by speech generation should be avoided. The goal is to sample the ideosyncratic way in which the trainee produces the speech while doing his job, not to influence the inflection or prosody of his speech. The auditory prompt may temporarily influence his speech pattern, but based on only a few repetitions it is unlikely to transfer to his speech behavior in the conduct of intercepts a few minutes, hours, or days later. Some unintentional mimicking of the auditory prompt is likely. To the extent that the mimicking does not become a lasting behavioral characteristic, the speech sample is distorted relative to the trainee's normal speech, and recognition accuracy is compromised.

Speech Stylization and Connected Word Recognition (CWR). The NEC DP-100 in ACE was reportedly the second unit delivered to this country. No American manufacturer produced a CWR unit during the ACE timeframe. Because the CWR technology was new, the techniques for its application were unfamiliar. The primary advantage of a CWR system is to provide flexibility in speech stylization. Pauses may or may not be inserted between vocabulary items. With the DP-100, up to five words/phrases could be spoken without pausing. However, this advantage may have worked to ACE's disadvantage in that the few remaining stylization requirements may not have been taught or practiced sufficiently.

The three speech stylization requirements in ACE were:

1. Always pause after the initial aircraft call sign  
Example: "Crackerjack... tighten turn"
2. Never add pauses. Pause only where indicated in the prompts you see and hear.
3. Always pause after a sequence of three digits.  
Example: "Stranger 165... 8"

(from Granberry, et al., 1981a)

Training in the Use of Speech Recognition. The three speech stylization requirements in ACE seemed simple enough. However, merely presenting them to the trainee and conducting a brief practice period in conjunction with speech sampling is not enough. The stylization requirements must be practiced and internalized by the trainee. Subsequent use of speech recognition in the context of performing an intercept did not provide adequate feedback for the trainee to determine when he failed to comply with stylization requirements. Failure to comply resulted in the usual problems associated with speech recognition errors discussed previously, i.e., the pilot/aircraft responses were erratic and long lists of error messages were presented after the problem. It was Canyon's observation that compliance with the stylization requirements tended to reduce the

rejection rate, although the recognition accuracy of "accepted" phrases may not have been greatly improved. We do not have data to support this observation.

Observation also indicated that failures to comply with stylization requirements were rarely detected by either the trainee or the instructor. The resulting poor recognition accuracy was assumed by the user to be indicative of the "state-of-the-art" of speech recognition. This assumption is true only to the extent that automated speech recognition is seen as an entire system, not just a black box that matches acoustic patterns. The entire system includes the following functions:

- o development cycles of vocabulary syntax definition (including pause requirements)
- o training the speaker how to comply with speech stylization requirements (including sufficient practice periods)
- o collecting speech samples in the context of the task (including stress-inducing factors)

Unfortunately, ACE did none of the above. The vocabulary definition, including pause requirements, was not tested on a population of AIC trainees and revised accordingly. Very little time was spent on training the user in the art and discipline of achieving high recognition accuracy. Post-problem feedback to the user was inadequate for him to make inferences about his speech behavior. Speech samples were obtained in a relatively benign atmosphere, while recognition was attempted under more stressful operational circumstances.

Immediate Speech Feedback. Previous research by Breaux (1976) indicated that the immediate presentation of alphanumeric information about "what was recognized" may be distracting to a trainee engaged in a voice-control task. In systems obtaining high recognition accuracy (circa 90%) this constraint may be particularly applicable.

However, the ACE trainees may have benefitted from knowing what was recognized, i.e., what the simulation system was responding to. Particularly in an R&D prototype, the optional capability to view the recognized output immediately would have been beneficial. This capability may have reduced the confusion caused by the sometimes mysterious behavior of the pilot/aircraft such as turning to a heading not called for by the trainee. Also, the immediate feedback capability may have helped the trainee shape his verbal behavior to conform with stylization constraints.

Speech Test and Retrain. The option to test the recognition accuracy of any vocabulary item was available in ACE. This provided an opportunity for the user to test items which seemed to be causing difficulty and to retrain (create new samples for) those items, if desired. This process allowed

the trainee to exercise the speech recognition system. It provided immediate visual feedback results, allowing the trainee to achieve consistency and to test the limits of speech variability tolerated by the system.

The test mode was introduced by ACE early in the syllabus, and it was encouraged in print (Granberry et al., 1981a). It was not used enough, however. More emphasis should be placed on its frequent use, both in the automated instruction and by the human instructors. A system design feature that would encourage the use of the test mode would be to automatically notify the trainee when the recognition system is having difficulty, i.e., when it is rejecting more than X% of the speech inputs.

**SPEECH GENERATION.** Two types of speech generation were used in ACE, synthesis and digitized recording and playback. The technology of speech generation is much less risky than speech recognition. No data were collected on the intelligibility of the speech generation in ACE. The synthesized speech (by VOTRAX) was sufficiently intelligible, particularly given the strong contextual cues. Some trainees remarked about the odd-sounding voice, but they soon adapted to it, and the synthesized speech became just another information source. Intelligibility is the key to system performance, but "realism" of the speech does contribute to user acceptance.

Digitized speech was used to simulate the role of SWC. This speech sounded more like normal speech and received better acceptance by the trainees.

The only problem related to the speech generation came from occasional time delays before initiating the message. This problem has been discussed previously. Delays of three or more seconds are not indicative of the technology, but of the system accessing the speech message. The temporal characteristics of speech communications are important for simulation and training in tasks such as AIC, and the system design should eliminate awkward delays.

Identifying speakers with different voices was a good feature of ACE. The trainees soon learned to distinguish SWC and CAP by the sound of their "voices."

The successful use of speech generation in ACE demonstrated the compatibility of speech recognition and generation technologies in the simulation of a communication network.

#### INSTRUCTOR'S ROLE

**INSTRUCTOR AS SYSTEM MANAGER/SUPERVISOR.** The ACE technologies were intended to assume many of the traditional roles of the instructor, such as teaching, performance monitoring, providing feedback, making syllabus decisions (individualizing instruction) and keeping records. The introduction of an automated training system, however, requires that the

instructor assume new roles including system initiation and management, tutoring for special problems, and general supervision of training.

**INSTRUCTOR TRAINING.** The shift in the instructor's role implies new responsibilities and, with them, the need for learning how to perform new tasks. The ACE instructors were experienced AICs (E-6 to E-7), but inexperienced computer operators. As stated in the final report on ACE, training the instructors how to operate the system was considered very important (Grady, 1982). In practice, however, the ACE instructors were given only one day of training. This was clearly inadequate, considering the complexity of the system.

The ACE instructors acquired the necessary skills over a period of many months of trial and error, supplemented by telephone calls for assistance and informal information exchanges during on-site support visits by the contractor. This situation was not optimal from the standpoint of achieving the support of the instructors for the project. If an operational version of ACE is introduced, a complete training program should be developed for the instructors. In the prototype ACE, the instructor training appeared to suffer because of resource shortfalls during system development.

**AUTOMATION AND INSTRUCTOR WORKLOAD.** The automated technologies in ACE were intended to reduce instructor workload, and enable improvement of the 1:1 instructor-trainee ratio found in the normal Synthetics method of training. In practice, however, the workload of the ACE instructors was reduced little, if any. An instructor's presence was necessary nearly continually to provide summary performance feedback and to override faulty syllabus decisions, particularly the frequent inappropriate remediation.

If ACE had received the benefit of a TEAR cycle, and many of the problems had been worked out of the system, the ACE instructor may have had a very light workload. The instructors hypothesized that they could potentially handle four to six AIC students with a fully operational ACE. In the prototype ACE, however, the instructor's participation was required to the extent that the related problems of speech recognition errors, faulty PMVs, and inappropriate syllabus control plagued the system.

**DEGREE OF INSTRUCTOR CONTROL.** ACE allowed the instructor a considerable amount of control over system operations. He could direct the trainee to any part of the syllabus. However, access was restricted to the beginning of a training unit. The instructor could not direct the training to a particular lesson within the training unit. Furthermore, the menu-driven process for achieving syllabus control was cumbersome, according to the ACE instructors.

In other facets of system control, the instructors felt that they were prevented from making constructive changes. This was particularly true in the area of PMV criteria. Instructors were not allowed to access the criteria data base. It is a difficult problem to define the extent of control and software access to allocate to the instructors in a prototype

system. The somewhat disgruntled attitude of the ACE instructors on this issue may have been influenced by the lack of in-depth training on system operation provided to them.

#### TEMPORAL FACTORS IN ACE TRAINING

A SAMPLE OF ACE TRAINING. Canyon's consultant, Dr. Chatfield, conducted a detailed analysis of the time spent by one trainee on various types of activities during a three hour sample of ACE training.

The total time that the trainee spent interacting with ACE was 159 minutes, after excluding break times. Table 2 summarizes the breakdown of time apportioned to the various categories of interaction with ACE.

TABLE 2. TIME ALLOCATION IN A SAMPLE of ACE TRAINING

| <u>Activity</u>           | <u>Time (Min.)</u> | <u>%</u> |
|---------------------------|--------------------|----------|
| Intercepts                | 46                 | 29       |
| Speech Data<br>Collection | 19                 | 12       |
| Feedback                  | 39                 | 25       |
| Down Time                 | 13                 | 8        |
| CPU Communication         | 19                 | 12       |
| Set Up                    | 18                 | 11       |
| Tutorial                  | 5                  | 3        |
| Total                     | 159                | 100      |

This analysis is based on the concept that AIC training involves primarily skill acquisition rather than simply knowledge acquisition. The training of these skills focuses on the intercept as the basic unit of practice.

The first category of interest is, of course, the intercepts themselves. During this three-hour block the trainee performed fourteen intercepts. Many of them, however, were not run to completion. They were either frozen because of errors, or aborted because of some malfunction. The trainee spent only 46 minutes on the intercepts themselves. This amounts to 29% of the total time. Thus, 71% of the trainee's time was spent in other categories of activities.

In a corresponding three-hour period, the trainee in Synthetics practiced about 30 intercepts to ACE's 14. These comparisons were tenuous, however, since the trainees were not at comparable points in the syllabus. There are six other categories of activities, shown in the table, which account for the differential rate of intercept presentation.

During the three-hour sample of training, the trainee was required to engage in speech sampling nine times for a total of 19 minutes. Thirty-nine minutes out of the total period were allocated to the reading of feedback displays. The trainee took roughly three minutes to read the feedback displays at the end of each intercept.

During the three hour sampling period, the system was down only once for a total of 13 minutes. Nine times, however, the trainee was essentially put into an idling mode, for a total of 19 minutes, during which the system indicated that the CPUs were communicating with each other. Also, a total of 18 minutes were spent while the trainee or the instructor typed in information for setting parameters and various other indices for the next intercept.

Another category of trainee activity included five minutes of tutorial, in which the system provided instruction through the use of CAI and the videodisc monitor. There was also approximately one minute (not recorded in the tabulations) in which the trainee attempted to replay an intercept. However, the replay option was not functioning at the time, and it simply didn't work.

The second column of figures in the table gives the percentage breakdowns of the various categories of the total 159 minutes training time. As can be seen, 29%, or less than one third of the trainee's time, was spent in actual practice on an intercept. Almost as much time was spent on reviewing feedback displays as was spent in practice. Four categories (speech data collection, feedback, CPU communication, and instructional setups) accounted for a total of 95 minutes related to the 46 minutes of intercept practice. These activities took roughly twice as long as the actual amount of time spent on intercepts. Thus, for every hour of practice on intercepts, two hours are required in other support activities.

The type and amount of feedback could probably be improved easily, and could result in a large time savings. As was observed, the trainee currently ignores much of the feedback. Improvement of the PMVs, in terms of the criteria for scoring and the actual measurement techniques themselves, would contribute substantially to more meaningful feedback and better time management.

In summary, it would seem very important to reduce the amount of time that a trainee spends in non-intercept practice categories. It is likely that, in any skill acquisition such as this, sheer practice is going to be one of the most important determiners of training efficiency. Although based on a small sample, the data indicate that the ACE trainee spends an inordinate amount of time in non-practice activities.

PILOT/AIRCRAFT RESPONSE TIME. The simulated intercepts were degraded by the frequent occurrence of a long (three seconds or more) delay in visual or audio response by the pilot aircraft. This delayed response increased the time stress on the trainee who was impatiently waiting for the chance to proceed with a sequence of actions.

These delays were unrealistic, according to the ACE instructors, and detracted from the execution of the intercept, which can be considered as a timed sequence of operational communications.

ERROR FEEDBACK TIME. The presentation of very long lists of error messages in ACE wasted the trainee's time. Because of the large number of invalid messages, the trainees almost entirely ignored the feedback after the first day or two on ACE. Providing a one-to-one correspondence between PMV failure (a score below on arbitrary criterion score) and feedback is not necessary. Summary information on strengths and weaknesses would be much more beneficial to the trainee.

If there were a measure of time well spent in automated training, ACE would have received a low rating. Performing an air intercept is a skill, and the more practice, the more skilled the performance. It is understandable that system bugs will exist in a prototype trainer. But, observation of the training process over many trainees led to the feeling that the temporal interface (to coin a term) between the system and the trainee had not been considered in a cohesive manner.

The major culprits in time consumption were: extended error-feedback messages; inappropriate remediation; unnecessary speech data collection for appropriate remediation; override system to avoid remediation; and pauses of several minutes while "the computers speak to each other." The most difficult job, as an observer, was to stay awake for the four intercepts per hour, each of which lasted an average of three or four minutes.

## USER ACCEPTANCE

The users of ACE were the trainees and instructors. In a larger context, the administrative personnel at FLECOMBATRACENPAC and the instructors not assigned to ACE also represented the user community.

The trainees generally liked the concepts behind ACE but were unimpressed with its performance. Their attitudes were expressed in the Exit Questionnaire. An example of the questionnaire and a summary of the ratings are given in Appendix A. It is interesting to note that the ratings on the question "What do you think of the ACE trainer...?" ranged from poor to excellent. Of the seven sampled, three trainees rated it "Poor," three "Good," and one rated it "Excellent."

These mixed reviews were clarified in the exit interviews. The essence of the mixed reaction was whether the trainee focused on the R&D concepts in ACE or on the implementation problems. Almost invariably, the trainees liked the concepts and believed that the technologies could be made to work, given a little "clean-up" time. Ultimately, this ambivalence also typified the attitudes of the ACE instructors and the Canyon observers. The technological concepts were demonstrated satisfactorily to make "believers" out of the users, but the prototype system had so many "bugs" and problems that it was totally unacceptable as an operational trainer in its present state.

## SECTION III

## SPEECH RECOGNITION ACCURACY EVALUATION

## INTRODUCTION

The simulated intercepts on ACE were completely automated through the use of interactive speech technology and pilot/aircraft modeling. Each intercept involved a series of radio transmissions to and from the AIC trainee. Correct recognition and understanding of the AIC's radio transmissions enables the simulated pilot/aircraft to respond (maneuver) appropriately as the intercept progresses. Errors in recognizing these transmissions lead either to a failure to respond or an incorrect (inappropriate) response of the pilot/aircraft model. The proper transmission of numbers is particularly critical to the AIC's task because bearing and range information are both important and frequent. Table 3 gives a typical set of transmissions in an intercept. The frequent occurrence of bearing and range information is evident in this example.

TABLE 3. A TYPICAL SPEECH-INTERACTION SCENARIO

| <u>PHRASE NO.</u> | <u>TRAINEE SAYS</u>             | <u>ACE/PILOT (OR SWC) SAYS</u>                                                 |
|-------------------|---------------------------------|--------------------------------------------------------------------------------|
| 91 *              | AIC-1                           | "SILVERHAWK ** AIRBORNE<br>FOR "CONTROL" (SWC)                                 |
| 87 *              | ROGER                           | "RUTH, THIS IS SILVERHAWK<br>ON MIRAMAR'S 180, 12,<br>ANGELS 20, OVER" (Pilot) |
| 87 *              | ROGER                           |                                                                                |
| 30                | SILVERHAWK, STARBOARD<br>275    | "ROGER, 275"                                                                   |
| 41 *              | SILVERHAWK, WHAT STATE?         | "STATE 224"                                                                    |
| 42                | ROGER, STATE 224                |                                                                                |
| 44 *              | I HAVE CONTROL OF<br>SILVERHAWK | "SWC AYE"                                                                      |
| 43                | SILVERHAWK, STATE 224           | "VERY WELL"                                                                    |
|                   | (Continued)                     |                                                                                |

TABLE 3  
(Continued)

| <u>PHRASE NO.</u> | <u>TRAINEE SAYS</u>                   | <u>ACE/PILOT (OR SWC) SAYS</u> |
|-------------------|---------------------------------------|--------------------------------|
| 35                | STATION 275, 36                       | -                              |
| 62                | STRANGER 270, 5                       | "LOOKING"                      |
| 35                | STATION 279, 34                       | -                              |
| 62                | STRANGER 270, 5                       | -                              |
| 34                | SILVERHAWK, PORT 195<br>FOR BOGEY     | "ROGER, 195"                   |
| 36                | BOGEY 195, 18                         | -                              |
| 37                | BOGEY TRACKING 290,<br>SPEED POINT 6  | -                              |
| 52                | BOGEY SINGLE,<br>ALTITUDE 18 THOUSAND | -                              |
| 30                | SILVERHAWK, STARBOARD 205             | "ROGER 205"                    |
| 36                | BOGEY 204, 6                          | "JUDY"<br>"FOX 1, BREAKAWAY"   |
| 80                | SILVERHAWK, BREAKAWAY 060             | "ROGER 060"                    |

\* Indicates phrases not containing digit strings

\*\* SILVERHAWK is the aircraft call sign (c/s)

The NEC DP-100 was chosen as the speech recognition unit for ACE because it was purported to be a sophisticated, state-of-the-art device, capable of recognizing limited connected speech. This capability allows the user to speak several previously-sampled words or phrases without pausing between them. Most current recognizers require pauses of 100 msec or more between phrases, but this constraint was considered unacceptable for the rapid delivery of successive digits (called digit-strings) inherent in the AIC's task. The NEC DP-100 was selected primarily for its

capability to parse digit strings and still achieve accurate speech recognition (Grady, 1982). In Canyon's analysis of the performance of the ACE speech recognition system, care was taken to address separately the recognition of transmissions with and without digit strings.

The output of the NEC DP-100 was fed into speech understanding software (SUS), developed by Logicon to boost the recognition accuracy by making use of task-relevant information. The analysis of the speech recognition system included assessment of the performance of both the DP-100 recognizer and the Logicon understanding software.

#### PROCEDURE

The procedures for the speech recognition accuracy evaluation can be divided into data collection, coding, and error analysis.

**DATA COLLECTION.** Speech recognition data samples were obtained during the following segments of the ACE syllabus: 3.13, 3.19, 4.13, 5.11, 6.13, 8.14-8.16 (as defined in Granberry, et al., 1981a). For every ACE trainee, data were obtained on two to four intercept problems at each syllabus segment, depending on factors such as the trainee's success at that level. These syllabus segments were selected to encompass the entire ACE lexicon as used in a variety of training scenarios.

The data were obtained by recording the audio portion of an intercept on a cassette recorder and, subsequently, transcribing the tape verbatim. Canyon estimates the error rate from the transcription process to be considerably less than 1%. The transcript then was compared with a printed output called the ACE Speech Understanding System (SUS) Recognition Summary Report, which contained a listing of the output of both the DP-100 and the Logicon understanding software.

**DATA CODING.** All utterances of the trainee first were checked against the list of "legal" phrases, as defined in the Instructor's Handbook (Granberry et al., 1981a), which was reproduced earlier in Table 1. Any utterances not conforming to the list were dropped from the sample. Examples of "illegal" utterances were the use of words or phrases not in the ACE vocabulary, coughs and stutters, or the use of incomplete or incorrect phraseology. This error rate was approximately 3%. The remaining "legal" utterances were identified and coded by the phrase number listed in the Instructor's Handbook.

These numbered phrases were used as the basic unit of analysis of the speech understanding system because they were the operationally significant unit. It should be noted that each phrase consisted of one or more vocabulary items, which may be independently recognizable by the DP-100. If the understanding software were ignored, the probability of correctly recognizing the entire phrase would be a multiplicative function of the probability of recognizing each item, i.e.,  $p(\text{phrase ABCD}) = p(A) \times p(B) \times p(C) \times p(D)$ . If, for example, a recognition probability of  $p=.90$  is assumed for each vocabulary item in a phrase of four, then the probability

of recognizing the entire phrase correctly would be  $(.90)^4 = .65$ . The understanding software, however, attempted to generate entire legal phrases from partial recognition, thereby elevating the phrase recognition probability above the multiplicative assumption.

The ACE SUS printout provided two separate listings for each utterance made by the trainee: (1) recognition by the DP-100 (the "raw" recognition, used as input to the understanding software) and (2) the final output of the Logicon SUS. The latter output was used by the ACE system as the final datum on which system responses were based. Each utterance in the transcript of an intercept was compared word-for-word with the final output of the SUS.

Error coding was carried out in two stages:

Stage 1. SUS performance coding

To determine the extent to which Logicon's understanding software improved upon the output from the NEC DP-100, the outputs of the two listings were coded for overall accuracy according to the following scheme: DP-100 output Correct or Incorrect; understanding software output Correct, Empty, Rejected, or Incorrect. This coding scheme corresponds to the operating characteristics of the two systems. Whereas the DP-100 output (as represented in the Speech Summary Report) was a "best-guess" interpretation of acoustic pattern matching, and, as such, is either accurate (correct) or inaccurate (incorrect), the understanding software is constrained to produce a "legal" AIC transmission as output. If unable to do so, it either rejected the transmission (e.g., by reporting a "say again" message) or produced a null (empty) response, indicating that the system "heard" nothing. The data coding scheme was developed to account for this multiplicity of function.

Stage 2. Coding for recognition accuracy

Each phrase uttered by the trainee was compared with the SUS final output and coded according to the following scheme:

| <u>ERROR<br/>CODE</u> | <u>DESCRIPTION</u>                                             | <u>SUS<br/>OUTPUT</u>          |
|-----------------------|----------------------------------------------------------------|--------------------------------|
| A                     | "system deaf" error; system failed to respond                  | blank, empty                   |
| B                     | "system reject" error; system rejects the utterance            | 'say again' *                  |
| C                     | "misrecognition" error; system output is not what trainee said | incorrect or incomplete phrase |

D "digit error"; output phrase is incorrect digit correct, but contains incorrectly recognized digit(s)

\* the computer did not always send the "say again" message; in some cases the SUS called for such a message but it was never transmitted.

In addition, digit errors (Type D: phrase contains an incorrect digit) were coded in further detail on the basis of the type and number of errors made, as follows:

#### DIGIT ERROR

| <u>TYPE</u> | <u>DESCRIPTION</u>                                                                                                                                                                         |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| D.1         | single-digit error; or the number of incorrectly-recognized digits (e.g., if the trainee said "Bogey 160," and the system output was "Bogey 106," it was coded as two single-digit errors) |
| D.2         | compound-number error; (e.g., if the trainee said "Station 145, 36," and the system output was "Station 134, 48" it was coded as one compound-number error.)                               |

Note that a phrase can contain both kinds of errors, as in the last example. It has two D.1 and one D.2 errors.

Finally, digit errors (D.1) made on bearing calls (e.g., Bogey 144) were coded to determine the frequency of digit errors as a function of position (100's place, 10's place, 1's place) in the digit string. The intent here was to obtain some estimate of the seriousness of digit-recognition errors in the AIC task context. An error of "1" in the 100's place, e.g., 200 vs 100, is more serious than an error of "1" in either the 10's (112 vs 122) or 1's (111 vs 112) place when giving bearing information. The complete coding scheme, therefore, incorporates DP-100 and SUS performance, recognition accuracy for phrases, recognition accuracy for digits, and single digit errors in bearing calls by position.

#### SAMPLE DESCRIPTION

A total of 1765 acceptable transmissions were obtained from six syllabus levels, as shown in Table 4. The number of samples with and without digit strings are indicated.

TABLE 4. SAMPLE SIZE (NUMBER OF TRANSMISSIONS) BY SYLLABUS LEVEL AND TRANSMISSION TYPE (WITH/WITHOUT DIGIT STRINGS)

| TRANSMISSION<br>TYPE | SYLLABUS LEVEL |      |      |      |      |      | TOTAL |
|----------------------|----------------|------|------|------|------|------|-------|
|                      | 3.3            | 3.19 | 4.13 | 5.11 | 6.13 | 8.13 |       |
| Digit String         | 146            | 241  | 352  | 227  | 252  | 222  | 1440  |
| No Digit String      | 129            | 1    | 58   | 87   | 19   | 31   | 325   |
| Total                | 275            | 242  | 410  | 314  | 271  | 253  | 1765  |

## RESULTS

RECOGNITION AND UNDERSTANDING ACCURACY. The overall speech understanding accuracy was 777 correct/1765 total = 44%. The term understanding accuracy here refers to the final output of the speech understanding software, not the "raw" recognition accuracy of the DP-100 device. This figure, 44%, represents the operational level of speech understanding that mediated the responses of the ACE simulation system and automated performance measurement.

The "raw" recognition accuracy of the DP-100 was 419 correct/1765 total = 24%. In our sample, therefore, the Logicon understanding software boosted the overall transmission recognition accuracy (TRA) from 24% up to 44%. In terms of the number of correctly recognized transmissions, the understanding software increased the accuracy of the recognizer from 419 to 777 correct, an increase of 85%.

It must be noted that these data are based on a very stringent criterion for correct recognition, i.e., the entire transmission must be recognized correctly. For example, if a single word or digit in a transmission is not recognized, the entire transmission is counted as one error. Thus, the word by word accuracy of the recognizer could be considerably greater than 24%.

The relationship between the accuracy of the DP-100 recognizer and the Logicon understanding software is given in greater detail in Table 5.

TABLE 5. RECOGNITION ACCURACY OF THE DP-100  
AND THE SPEECH UNDERSTANDING SOFTWARE (SUS)

|                  |           | SUS OUTPUT   |              |                |
|------------------|-----------|--------------|--------------|----------------|
|                  |           | Correct      | Incorrect    | Total          |
| DP-100<br>Output | Correct   | 370<br>(21%) | 49<br>(3%)   | 419<br>(24%)   |
|                  | Incorrect | 407<br>(23%) | 939<br>(53%) | 1346<br>(76%)  |
|                  | Total     | 777<br>(44%) | 988<br>(56%) | 1765<br>(100%) |

It is evident from the table that the understanding software corrected DP-100 errors in 23% of the sample, while it created errors from correct DP-100 output on only 3% of the sample. Still, on 53% of the transmissions, the understanding software was unable to correct recognizer errors. These data indicate that the understanding software was doing a creditable job of improving the accuracy of the DP-100 recognizer, but much fertile ground remains for improving recognition accuracy.

The serious effects of recognition errors in an interactive training system such as ACE will be discussed in detail. In passing, however, it can be noted that recognition errors can cause immediate problems of inappropriate aircraft response, incorrect performance measurement leading to fallacious performance feedback at the end of a problem, and eventual unnecessary syllabus remediation.

In all subsequent discussion of speech recognition accuracy, the final output of the system is being used, i.e., the output of the SUS.

Analysis of Error Types. Contrary to what might be expected, a small minority (10%) of the speech recognition errors were caused by phrase misrecognition (insertions or confusions). A larger error rate (37%) was found for digit strings, but, most importantly, over half (53%) of the errors occurred when a transmission resulted in no output from the recognition system. Table 6 shows the results of frequency and percentage of error types.

TABLE 6. SPEECH RECOGNITION ERRORS BY TYPE

| <u>Error Type/Description</u> |   |                       | <u>Frequency</u> | <u>% of Total Errors</u> |
|-------------------------------|---|-----------------------|------------------|--------------------------|
| Rejected                      | A | No Response           | 398              | 40                       |
|                               | B | "Say Again"           | 129              | 13                       |
| Accepted                      | C | Phrase Misrecognition | 96               | 10                       |
|                               | D | Digit String Error    | 365              | 37                       |
| Total Errors                  |   |                       | 988              | 100%                     |

One way to conceptualize the data is to consider Type A and B errors as "Rejects," i.e., the system never gave a recognition output for these transmissions. The remainder of the sample can be considered "Accepted" by the system. Errors in this category would be Type C or D. Table 7 gives an analysis of the data according to this distinction between "Accepted" and "Rejected" transmissions. This classification speaks more directly to the issues under consideration since "rejects" cannot, by definition, be considered accurate or inaccurate. Although the overall recognition accuracy was 44%, that figure was improved to 63% when the "accepted" transmissions were considered.

A further breakdown of the 461 errors in the "Accepted" category showed that 96 (21%) were Type C (phrase misrecognition) while 365 (79%) were Type D (digit string) errors.

TABLE 7. RECOGNITION ACCURACY FOR  
"ACCEPTED" AND "REJECTED"  
TRANSMISSIONS

| <u>Category</u> | <u>Number of Transmissions</u> | <u>Error</u> | <u>Error Rate</u>    | <u>Recognition Accuracy</u> |
|-----------------|--------------------------------|--------------|----------------------|-----------------------------|
| Overall         | 1765                           | 988          | 56%                  | 44%                         |
| "Rejected"      | 527                            | NA           | 30% (rejection rate) | NA                          |
| "Accepted"      | 1238                           | 461          | 37%                  | 63%                         |

Similarly, a breakdown of the 527 "Rejected" transmissions showed that 398 (76%) were Type A (no response) while 129 (24%) were Type B (say again) errors. The importance of the distinction between these two types of errors is one of training effectiveness and user acceptance. The trainee knows the system did not "understand" a transmission when a "say again" is given; he remains uncertain when no response (verbal or displayed) is given.

Analysis of Digit String Errors. Many transmissions from the AIC trainee include bearing and range information, such as "Bogey two four zero twenty three." For the present analysis, the bearing information is considered a string of three digits and the range is termed a compound number. The majority (1440/1765 or 82%) of all transmissions contained a digit string. The recognition accuracy for these transmissions was 38%. Conversely, 62% of all transmissions containing digit strings were in error.

An analysis of error type for the digit string transmissions is given in Table 8, where it can be seen that half of the errors were due to rejection and half to misrecognition. Of the 50% recognition errors, 40% were digit recognition errors and only 10% non-digit errors.

Of the 1440 transmissions containing digit strings, 999 (69%) were "accepted" by the system and 441 (31%) were "rejected" (Error Types A+B). If only the "Accepted" digit string transmission are considered, the recognition accuracy rate was 55%.

TABLE 8. ERROR TYPES FOR 1440 TRANSMISSIONS  
CONTAINING DIGIT STRINGS

| <u>Error Type/Description</u> |                          | <u>Errors</u> | <u>Percentage of Total Errors</u> |
|-------------------------------|--------------------------|---------------|-----------------------------------|
| A                             | No Response              | 320           | 36%                               |
| B                             | "Say Again"              | 121           | 14%                               |
| C                             | Phrase<br>Misrecognition | 86            | 10%                               |
| D                             | Digit String<br>Error    | 365           | 40%                               |
| Total Errors                  |                          | 892           | 100%                              |

Analysis of Non-Digit String Errors. The same type of analysis was performed for transmissions not containing digit strings, to provide a basis for comparison to the assessment of the "digit-string problem." Only 325 of the 1765 transmissions contained no digit strings. Of this sample of 325, recognition errors occurred on 96, or 30%. Conversely, the recognition accuracy was 70% for non-digit string transmissions. Table 9 gives the error breakdown by type.

TABLE 9. ERROR TYPES FOR 325 TRANSMISSIONS  
CONTAINING NO DIGIT STRINGS

| <u>Error Type/Description</u> |                          | <u>Errors</u> | <u>Percentage Total Error</u> |
|-------------------------------|--------------------------|---------------|-------------------------------|
| A                             | No Response              | 78            | 81%                           |
| B                             | "Say Again"              | 6             | 6%                            |
| C                             | Phrase<br>Misrecognition | 12            | 13%                           |
| D                             | Digit String<br>Error    | 0             | 0%                            |
| Total Errors                  |                          | 96            | 100%                          |

It is interesting to note that 87% of the errors for the no-digit-string transmissions occurred because of rejection; only 13% of the errors were actual misrecognitions. Another way to state this result is that 95% of the 241 "accepted" transmissions (without digit strings) were recognized accurately. Clearly, the rejection rate is cause for concern.

Compound Number and Single-Digit Errors. Range information was frequently greater than 20 nm (nautical miles). In such cases a compound number, such as "twenty-three," was required. Observational analyses suggested particularly poor recognition accuracy with compound numbers, so an in-depth analysis was performed on them, as well as on the single digits in digit strings.

A total of 4,766 single digits were obtained in the sample, and 261 compound numbers. The SUS recognition accuracy for this sample is given in Table 10.

TABLE 10. NUMBER OF CASES CORRECT, REJECTED, AND MISRECOGNIZED FOR SINGLE DIGITS AND COMPOUND NUMBERS

| Outcome                         | Single Digits |        | Compound Numbers |        |
|---------------------------------|---------------|--------|------------------|--------|
| Rejected<br>(Type A or B error) | 1548          | (33%)  | 91               | (35%)  |
| Misrecognized                   | 490           | (10%)  | 110              | (42%)  |
| Correct                         | 2728          | (57%)  | 60               | (23%)  |
| Total                           | 4766          | (100%) | 261              | (100%) |

The 57% recognition accuracy for single digits was more than twice as good as that for compound numbers. Only 23% of the compound numbers were recognized correctly.

Again, the importance of Type A and B (rejection) errors must be considered. If only the "Accepted" samples are included in the analysis, 85% of the single digits, but only 35% of the compound numbers, were recognized correctly. The rejection rates were nearly equal for single-digit and compound numbers, but the misrecognition rate was four times higher for compound numbers.

However, single digits occurred much more frequently (4,766 versus 261), so that if these two types of number errors are pooled (N=5027), single digit errors account for 82% of all the number-recognition errors.

Operational Significance of Single-Digit Errors. The importance of the single-digit recognition errors must be judged on the basis of the accuracy of the heading calls given by the AIC trainee. For example, if the trainee says "vector two three one" and the system recognizes 232, a heading error of only one degree is introduced, and the intercept will proceed normally. However, if the system recognized 331, a gross error of 100° is introduced as the aircraft makes an incorrect turn. An error of such magnitude may render the intercept insolvable, at least in the training environment.

The following data were derived to determine the operational significance of misrecognitions in heading calls. In Table 11 the three digit positions in a heading call are signified as the 100's, 10's, and 1's position, and the frequency of errors in each position is given.

TABLE 11. SINGLE DIGIT ERRORS IN HEADING CALLS, BY POSITION

| Example: "Vector   | 3            | 4           | 0 "        |              |
|--------------------|--------------|-------------|------------|--------------|
|                    | <u>100's</u> | <u>10's</u> | <u>1's</u> | <u>Total</u> |
| Recognition Errors | 69           | 133         | 135        | 337          |
| % of Total         | 20%          | 40%         | 40%        | 100%         |

In 40% of the cases, the recognition error was less than 10 degrees, and usually not operationally critical. But in the remaining 60% of the cases the heading error ranged from 10 to hundreds of degrees in error - an unacceptable error for conducting an intercept.

Recognition Accuracy by Syllabus Level by Trainee. All the data presented so far have been pooled across syllabus levels and trainees. Table 12 shows recognition accuracy by trainee, across the six syllabus levels sampled. Although a trend toward increased recognition accuracy might be expected with practice, the data shown in Table 12 does not reflect a strong tendency toward increased recognition with advanced syllabus level. However, the ACE instructors pointed out that the difficulty of the problems at various syllabus levels tend to follow these data. For example, syllabus level 8.13 "Running Super Set-Ups" was difficult, had a larger active vocabulary, and was associated with a lower recognition accuracy. By contrast, level 6.13, "Performing a Rendezvous," was easier, had a more limited vocabulary, and resulted in a higher recognition accuracy.

Inspection of the mean recognition accuracy by trainee shows that a surprisingly small variation among the eight trainees was evidenced. The total range of mean recognition accuracy was less than 14%.

**DISCUSSION.** Accurate speech recognition is essential for the proper functioning of a speech-interactive training system like ACE. Overall, the speech recognition accuracy of ACE was inadequate to support an operational training system. The causes of the high error rates must be determined and corrected before the potential cost savings and training benefits of AST can be realized.

Several factors are suggested as critical to obtaining good recognition accuracy: (1) speech sampling in context; (2) instruction and practice on speech recognition; (3) immediate recognition feedback availability; (4) extensive development of supporting software; and (5) allocation of sufficient project resources to test evaluate and revise (TEAR) the lexicon and stylization rules based on the speech performance of trainees, not system engineers.

TABLE 12. TRANSMISSION RECOGNITION ACCURACY BY SYLLABUS LEVEL  
FOR EIGHT TRAINEES

| TRAINEE | Syllabus Level |      |      |      |      |      | Mean |
|---------|----------------|------|------|------|------|------|------|
|         | 3.13           | 3.19 | 4.13 | 5.11 | 6.13 | 6.13 |      |
| 1       | 35             | 33   | 33   | 42   | 75   | -    | 43.6 |
| 2       | 59             | -    | 24   | 56   | -    | -    | 46.3 |
| 3       | 30             | 24   | 47   | 40   | 57   | 29   | 37.8 |
| 4       | 54             | 26   | 33   | 44   | 48   | 39   | 40.7 |
| 5       | -              | 31   | -    | -    | 52   | -    | 41.5 |
| 6       | 63             | 49   | 39   | 44   | 63   | -    | 51.6 |
| 7       | 53             | 33   | 36   | 55   | 57   | 51   | 47.5 |
| 8       | 58             | 33   | -    | -    | -    | -    | 45.5 |
| MEAN    | 50.3           | 32.7 | 35.3 | 46.8 | 58.7 | 39.7 | 44.1 |

The recognition accuracy levels found in the ACE sample (24% from the recognizer and 44% from the understanding software) are not generalizable to other applications or methods of system implementation.

The problems of speech recognition error in a highly interconnected system like ACE are illustrated in Figure 3.

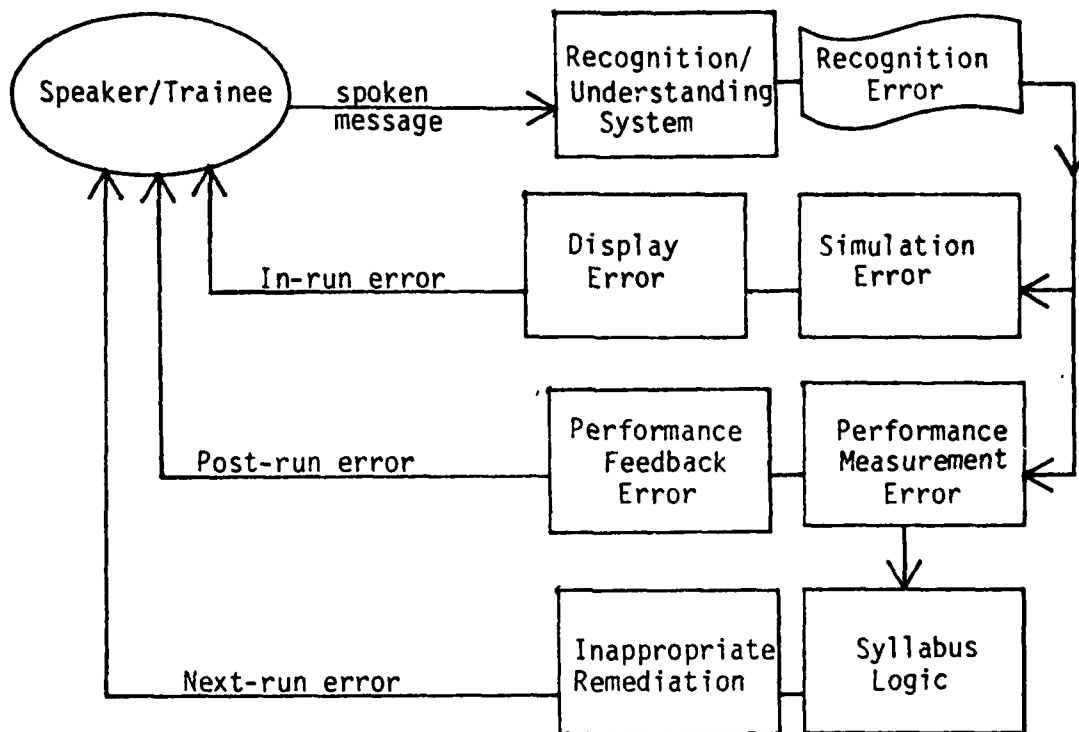


Figure 3. Three Categories of Training Errors Can be Caused by Speech Recognition Errors

Recognition errors have a negative effect on the problem simulation, the accuracy of the trainee's performance feedback, and the appropriateness of syllabus decisions such as remediation. Because many of the ACE subsystems are dependent on accurate speech recognition, various strategies and techniques for improving the accuracy should be attempted. Suggestions for such improvements will be discussed in the final section of this report.

## SECTION IV

## STUDIES OF PERFORMANCE MEASUREMENT AND TRANSFER OF TRAINING

## PERFORMANCE MEASUREMENT VALIDATION STUDY

INTRODUCTION. A description of the ACE Performance Measurement Variables (PMVs) is given in the ACE Instructor's Handbook as follows:

"A PMV is a statement of expected learner behavior which includes the:

- . conditions under which it should occur
- . timing and/or frequency of occurrence
- . point values (out of a total of 100) which are assigned to certain aspects of the behavior" (Granberry, et al., 1981a).

A total of 87 PMVs were listed in the Instructor's Handbook. Only a subset of these was active for a particular problem, depending on the behavioral objectives and the task demands, as defined by the trainee's position in the course curriculum.

Two examples of PMVs are given below. The examples were selected from those active in syllabus levels 6.13 and 8.3, selected for the PM Validation study.

---

Syllabus  
Level 6

---

| Number | Performance<br>Measurement<br>Variable (PMV) | Component<br>Parts                                                                                                                              | Point<br>Value                                     |
|--------|----------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------|
| 40     | TRANSMIT VECTORS<br>FOR RENDEZVOUS           | Respond to the CAP<br>message "REQUEST<br>RENDEZVOUS WITH<br>CRACKERJACK" within<br>10 seconds with the<br>message "SILVER HAWK<br>VECTOR xxx." | 60<br>(30 for<br>response<br>within 15<br>seconds) |

(Cont.)

---

## (PMV Examples Continued)

| Number              | Performance<br>Measurement<br>Variable (PMV) | Component<br>Parts                                                                                                                                                                                                                             | Point<br>Value                                                          |
|---------------------|----------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| 40 (cont.)          |                                              | Respond to the CAP message "REQUEST RENDEZVOUS WITH CRACKERJACK" within 24 seconds after the CAP request with the message "CRACKERJACK VECTOR xxx."                                                                                            | 25                                                                      |
|                     |                                              | Silver Hawk vector must be within +/- 10 degrees of the bearing from the CAP to the MAC.                                                                                                                                                       | 15                                                                      |
| Syllabus<br>Level 8 |                                              |                                                                                                                                                                                                                                                |                                                                         |
| 9*                  | TRANSMIT CONTINUING BOGEY BEARING AND RANGE  | Transmit the message "BOGEY xxx, yy" within 10 seconds after the sweep passes the bogey position. The maximum score is given if the transmissions are made 3 out of 5 times and the AIC does not miss making a transmission 2 sweeps in a row. | 70<br>(The score is decreased in proportion to transmission omissions.) |

(from Granberry, et al., 1981a)

The score yielded by a PMV is dependent on at least the following parameters: the timing criteria (e.g., within 10 seconds); the frequency criteria (e.g., 3 out of 5 times); and the assigned point value. These parameters were set through estimation based on task analysis and the recommendations of a subject-matter expert (SME).

PROCEDURE. Validation of the ACE performance measurement system, and evaluation of the utility of the PMVs used in that system, were accomplished by comparing samples of computer grading with independent grading by an instructor. Each trainee was graded on a total of eight runs, four intercept problems in each of two syllabus levels (6.13 and 8.13). The grading was done by an instructor using a specially-devised form (Appendix B). The computer also graded each run according to its built-in protocols. The instructor's independent assessment of trainee performance was compared with the computer's evaluation of the same problems. Assuming that the structured ratings of the experienced instructors were indicative of the trainee's proficiency, the correlation between instructor scores and computer scores is an index of the validity of the automated performance measurement system. In addition, records were kept of the trainees' performance scores on each of the PMVs to enable evaluation of the pass/fail rates of the PMVs within the two syllabus levels sampled. The mean of the set of PMV scores which were activated for each intercept was calculated as an index of performance for that intercept. Additionally, the instructor's rating was obtained for each intercept. Both methods of measurement were based on a 100 point scale.

RESULTS AND DISCUSSION. The correlation between 40 sets of PMV scores and instructor ratings was  $r = .35$  ( $p < .05$ ,  $df = 38$ ). This low (but marginally significant) correlation indicates that the two measurement methods were largely measuring different phenomena. If the instructor ratings are assumed to be the criterion for measuring AIC proficiency, then it must be concluded that the ACE PMVs have limited validity and can be considered only weak indices of AIC performance.

Comparison of the mean scores from the two measurement methods showed that the instructor ratings were, on the average, nearly 20 points higher than the PMV scores. This difference probably was due to arbitrary weightings used in both scoring methods, but it also may reflect the influence of speech recognition errors on ACE PMVs. The important point is the relationship between the two sets of scores, identified by the correlation of  $r = .35$ , which is not affected by the overall mean difference reported above.

The number of times a particular PMV was passed in the sampled intercepts is shown in Table 13. Different PMVs were used in the two syllabus levels (6.13 and 8.13). Pass rates ranged from 0% on PMVs #8, 46, 65, and 67 to a high of 87% on PMV #42. Clearly, PMVs which are never passed need revision (although the pass rate for a PMV does not necessarily indicate its validity).

TABLE 13. "PASS" PERCENTAGE FIGURES FOR SPECIFIC  
PMVS MEASURED DURING PMV VALIDATIONS STUDY

| <u>Syllabus<br/>Level</u> | <u>PMV NO.</u> | <u>Passed/Attempts</u> | <u>Pass %</u> |
|---------------------------|----------------|------------------------|---------------|
| 6.13                      | 40             | 5/19                   | 26            |
| "                         | 41             | 14/19                  | 74            |
| "                         | 42             | 13/15                  | 87            |
| "                         | 43             | 8/19                   | 42            |
| "                         | 44             | 6/17                   | 35            |
| "                         | 45             | 3/17                   | 18            |
| "                         | 46             | 0/12                   | 0             |
| 8.13                      | 1              | 11/21                  | 52            |
| "                         | 7              | 5/9                    | 56            |
| "                         | 8              | 0/17                   | 0             |
| "                         | 9              | 2/21                   | 10            |
| "                         | 22             | 1/17                   | 6             |
| "                         | 56             | 18/21                  | 86            |
| "                         | 65             | 0/3                    | 0             |
| "                         | 66             | 1/7                    | 14            |
| "                         | 67             | 0/7                    | 0             |
| "                         | 68             | 2/3                    | 67            |
| "                         | 70             | 7/17                   | 41            |
| "                         | 71             | 5/19                   | 26            |
| "                         | 72             | 8/18                   | 44            |
| "                         | 76             | 10/21                  | 48            |
| "                         | 79             | 9/21                   | 43            |

The data obtained in the PMV Validation study indicate that the PMVs need revision or further development to provide adequate measures of trainee performance. This conclusion is supported by the interviews with the ACE instructors. Apparently, the developers of ACE had insufficient time and resources available to do a thorough test, evaluation and revision (TEAR) of the PMVs with actual AIC trainees prior to the delivery of ACE to FLECOMBATRACENPAC. These results should not be interpreted as an indictment of automated performance measurement (APM) processes. We feel strongly that acceptable APM can be achieved for the AIC task and other similar tasks.

Speech recognition errors contributed directly to PMV inaccuracy. Most recognition errors resulted in lower scores on at least one PMV because the trainee was not credited with giving a proper command within a specified time limit. Sometimes one recognition error led to a series of

PMV penalties. Improved accuracy in speech recognition would result in substantial improvements in PMV validity. (See the section headed "Speech Recognition Accuracy Evaluation"). As a worst-case example, it is obvious that the PMVs cannot accurately measure performance in Syllabus Level 8, where the speech recognition accuracy was less than 40%.

Was it justifiable to use the structured ratings of instructors as a criterion for validity of the PMS? The AIC instructors constantly conduct structured performance ratings during AIC training and they have extensive experience in evaluating AIC task performance parameters using a structured rating format. Furthermore, the instructor ratings were demonstrably reliable in their evaluation of trainee performance: the correlations were moderately high among the instructor ratings on PMV test (PMVINSTR), overall instructor ratings at the end of AIC training (INSTRATE), and overall school grade (SCHLPERF). These data are shown in Table 14.

TABLE 14. CORRELATION MATRIX OF INSTRUCTOR RATINGS  
AND SCHOOL PERFORMANCE

|          | PMVINSTR | SCHLPERF |
|----------|----------|----------|
| INSTRATE | .79*     | .68*     |
| SCHLPERF | .60      |          |

\*  $p < .05$

The evidence supports the conclusion that the ACE performance measurement system, as it is currently configured, does not provide a valid measure of the trainee's overall performance.

This result is not surprising, considering the quality of the major input to the PMVs, namely, speech recognition. Until the problem of speech recognition errors can be reduced, a detailed analysis of the ACE automated performance measurement system is inconclusive. The inadequate measurement may be due to recognition errors, the PMVs, or a combination of both.

#### TRANSFER OF TRAINING (TOT) TEST

INTRODUCTION. The ideal TOT test would assess the transfer of training from ACE to operational performance as an AIC. This was not feasible within the time and cost constraints of the present program. Therefore, a criterion test of AIC proficiency was designed specifically for the TOT study. The objective was to determine the training effectiveness of ACE relative to the traditional training received in Synthetics.

The wisdom of pursuing a TOT test was open to question, considering that ACE had all the problems discussed previously in this report. Nevertheless, the TOT study was conducted as planned.

**PROCEDURE.** The TOT criterion test was developed on the basis of three primary objectives: (1) to be representative of the operational job of the AIC; (2) to be "equally unfair" to trainees from ACE and Synthetics (regarding their respective curricula); and (3) to be scored in an objective and unbiased manner.

**Test Development.** The AIC instructors, with Canyon personnel acting as facilitators, generated a list of candidate AIC operational skills for representation on the TOT test. These skills were integrated into five types of intercepts or events to be included in the TOT test. However, the curricula differed in ACE and Synthetics, which led to the principle of a TOT test that was "equally unfair" to both groups. As a result, the TOT test included two types of tasks for each trainee: those on which he had been trained and those on which he had received little or no training.

The test consisted of 10 intercept tasks arranged as two consecutive runs in each of the following event categories:

1. Fleet-type NCI. A Nearest-Collision Intercept (also called a Forward-Quarter intercept) performed using a single aircraft and a simulated bogey. This type of intercept is taught by ACE but not in Synthetics.
2. Training-Type NCI. A Forward-Quarter intercept using two aircraft, one of which is designated (by the AIC) to play the role of the bogey, the other the interceptor. This type of intercept is the standard in the Synthetics curriculum, and is taught briefly in ACE (where it is called "Training Set-Ups").
3. Emergencies. During the course of two additional Training-type NCIs the trainee encountered an emergency situation requiring immediate and specific action. In the first run, the AIC lost radio communication with one of his aircraft (NORDO). In the second run, the NTDS system went down (simulated loss of NTDS program) and the AIC was required to complete the intercept in Mode 3, using only the radar blips and a grease pencil to finish the exercise. Training for these emergencies is received to some extent in both ACE and Synthetics.
4. Tanker joinups. This event simulated a refueling mission in which one aircraft is designated as the "tanker" and is required to join with the fighter in order to commence refueling. The two training methods provided approximately equal (very little) training for this event.

5. Stern conversions. In a Nearest-Collision Intercept Conversion (NCIC) the fighter aircraft is maneuvered to a position behind the bogey in order to fire a missile up the bogey's tailpipe. Neither the ACE nor the Synthetics trainee had received any training on this intercept method.

Subjects. The Synthetics trainees were selected from each consecutive class based on his progress. The criteria were that the training received to date should match as closely as possible that received by the ACE trainee, both in terms of quantity and of content. The use of this criterion reduced the size of the Synthetics sample, because appropriate candidates were not always available when needed.

Of the nine ACE trainees tested, two failed to complete the test for various reasons. Another ACE trainee failed to progress to the point where he could be tested.

The final group sizes were six trainees in ACE and six in Synthetics.

Sample Equivalence. Data obtained from the Entry Questionnaire and Personnel Records were analyzed to determine, post hoc, the equivalence of the trainees assigned to the ACE and Synthetics training groups. On 11 of the 12 scales contained in the Entry Questionnaire the difference between groups was not significant. Similarly, three sets of scores from personnel records showed no difference between groups, namely, the OS ASVAB score, mean ASVAB score, and CIC score. These results support the conclusion that the trainees assigned to the ACE and Synthetics groups for the TOT test were equivalent.

Test Administration. Each ACE trainee received approximately one half day of transition training to familiarize him with equipment and procedural differences between the ACE system and the operational training system. ACE trainees were tested every two weeks; Synthetics trainees were tested on alternate weeks.

Although the trainees in both groups were required to perform exactly the same set of tasks, the order in which they were presented was different. The TOT test was structured so that the trainee began with the task with which he was most familiar (i.e., Fleet-type NCI for the ACE trainee and Training-type NCI for the Synthetics trainee) and performed progressively less familiar tasks to the end. The order of events for ACE trainees was: Fleet-type NCI, Training-type NCI, Emergencies, Tanker joinups, and Stern conversions. The order for Synthetics trainees was: Training-type NCI, Emergencies, Fleet-type NCI, Tanker joinups, and Stern conversions.

The TOT test utilized two connected NTDS AIC consoles and the support of the simulation services of the Problem Control and Evaluation (PC&E)

pseudo-pilots. This equipment configuration was identical to that used in the Synthetic training, except that the testing took place in the "lives" equipment spaces.

The instructor in charge played an active role in the setup and control of the intercept scenarios and events in the TOT test. He conducted the test exercise from the AIC console linked to the trainee's console. The instructor generated and manipulated symbols which appeared on the trainee's display (e.g., bogeys, strangers, etc.) and communicated both with the trainee (to simulate SWC, for example) and with the PC&E pseudo-pilot. The instructor also was able to communicate directly with PC&E without the trainee's knowledge in order to simulate events requiring action by the pseudo-pilot. A member of the Canyon team assisted in the conduct of the test exercise.

The test was graded by an impartial, highly qualified and experienced AIC (not a member of the AIC instructor staff) positioned to observe the trainee's actions as well as the console display. The test was graded using a structured grading instrument (Appendix C) developed for this purpose.

All test administrators were equipped with radio headsets which allowed them to monitor communications between trainee, instructor and PC&E pseudo-pilots.

Test Protocol. The conduct of the test was rigidly standardized to ensure equivalent treatment for all trainees. Every test event was specified in a written protocol (see Appendix D). For example, the origin and path taken by bogeys and strangers were defined by bearing and range from a standard reference point. This procedure ensured uniformity of events across trainees and instructors. In addition, all communications between instructor and trainee were proscribed by the written protocol, the only allowed deviations being in case of equipment problems or other unforeseen circumstances.

Debrief. Following completion of the TOT test, each trainee was debriefed by the Instructor, the grader, and the Canyon researcher. The purpose of the test was explained and the trainee was assured that, regardless of his performance, the test would have no bearing on his subsequent training. Since the trainee had completed over an hour's training on the intercept exercises, however, the instructor and grader thoroughly debriefed the trainee on all aspects of his test performance. The objective of this debriefing was to utilize the test exercise as an instructional device to supplement normal training.

RESULTS AND DISCUSSION. The scores on the five types of intercepts in the TOT test are given in Table 15 for the six trainees in each group. Each score represents the mean performance on two intercepts. An analysis of variance (ANOVA) was performed on these data and is summarized in Table 16. Although the mean difference between the two groups was 23 points, that difference was not statistically significant,  $F(1,10) = 2.80$ ,  $p > .10$ . These

data are plotted in Figure 4 where it is clear that the performance of ACE trainees was nearly equal to that of Synthetics trainees on Fleet-type NCIs, but 20 to 35 points lower than Synthetics trainees on the remaining four event types.

TABLE 15. TOT TEST SCORES FOR TWO TRAINING GROUPS  
BY TYPE OF INTERCEPT

|            |   | TYPE OF INTERCEPT |             |                 |      |      | X    |
|------------|---|-------------------|-------------|-----------------|------|------|------|
|            |   | ACE<br>NCI        | TRNG<br>NCI | NORDO<br>MODE 3 | TNKR | NCIC |      |
| Trainee    |   |                   |             |                 |      |      |      |
| ACE        | 1 | 38*               | 54          | 57              | 36   | 51   | 47.2 |
|            | 2 | 18                | 23          | 10              | 0    | 0    | 10.2 |
|            | 3 | 75                | 79          | 62              | 5    | 58   | 55.8 |
|            | 4 | 32                | 26          | 13              | 24   | 25   | 24   |
|            | 5 | 75                | 35          | 0               | 30   | 31   | 34.5 |
|            | 6 | 73                | 79          | 90              | 97   | 57   | 79.2 |
| X          |   | 51.8              | 49.3        | 38.7            | 32.0 | 37.0 | 41.8 |
| Synthetics | 1 | 44                | 76          | 80              | 42   | 33   | 55   |
|            | 2 | 59                | 81          | 77              | 84   | 97   | 79.6 |
|            | 3 | 81                | 63          | 78              | 80   | 66   | 73.6 |
|            | 4 | 32                | 12          | 36              | 2    | 29   | 22.2 |
|            | 5 | 76                | 100         | 90              | 71   | 90   | 85.4 |
|            | 6 | 41                | 87          | 82              | 97   | 60   | 73.4 |
| X          |   | 55.5              | 69.8        | 73.8            | 62.7 | 62.5 | 64.9 |

\* Data are percents, computed as total points awarded divided by total points possible, and are based on mean performance on two intercepts.

Based on the results of the ANOVA, no significant difference in training transfer was found between the ACE and Synthetic groups. As seen in the figure, however, the data suggest better performance by the Synthetics group.

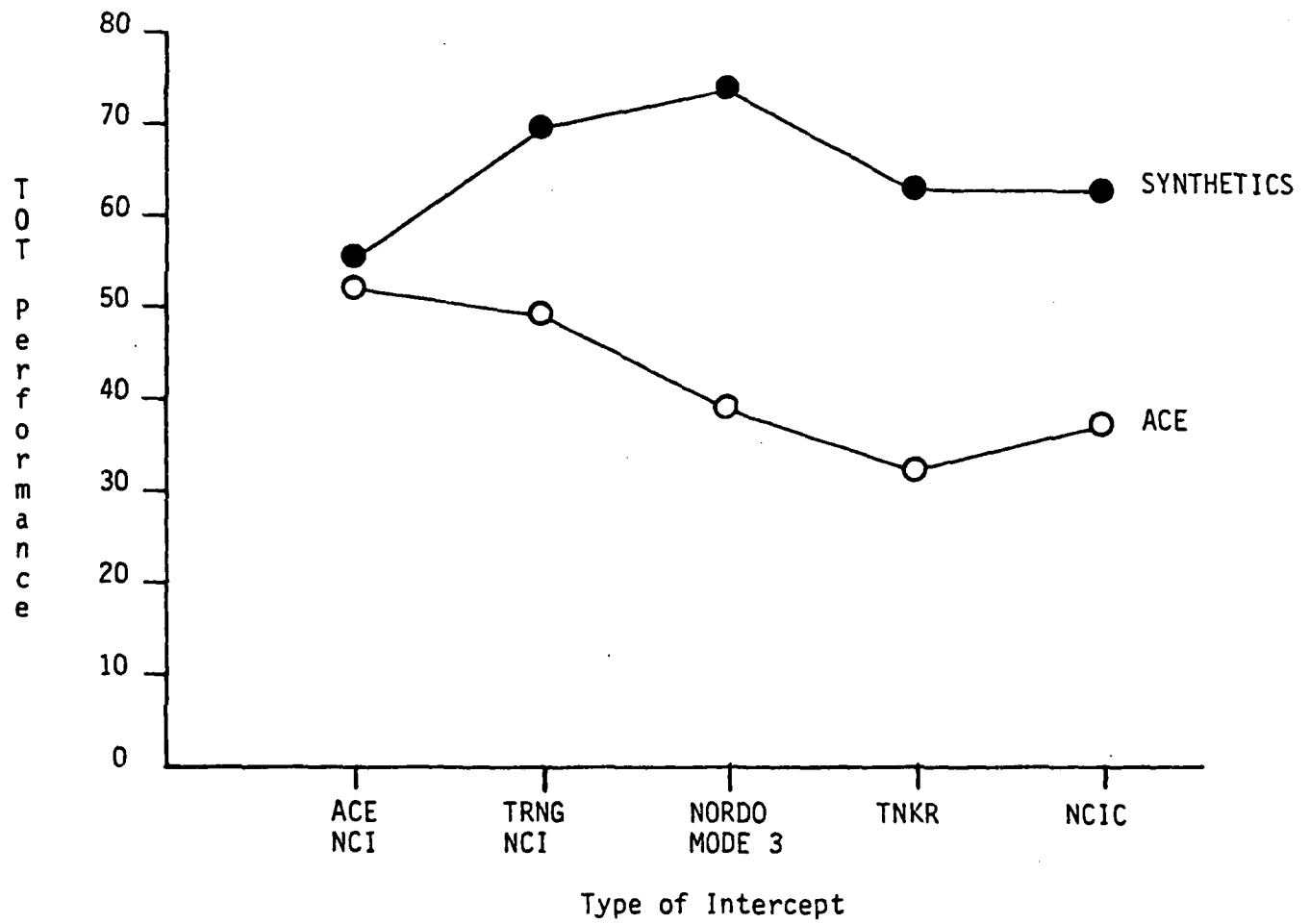


Figure 4. TOT Study: Mean Performance Scores by Type of Intercept and Training Group

The power of the statistical analysis was limited by the small sample size (N=6 in each group), and only a very large mean difference would be found statistically significant. Efforts to increase the sample size were thwarted by limited resources.

TABLE 16. ANALYSIS OF VARIANCE SUMMARY  
TABLE: TOT RESULTS

| <u>Source</u>         | <u>df</u> | <u>SS</u> | <u>MS</u> | <u>F</u> | <u>p</u> |
|-----------------------|-----------|-----------|-----------|----------|----------|
| Mean                  | 1         | 170453.4  | 170453.4  | 59.87    | .000     |
| Group <sup>1</sup>    | 1         | 7981.1    | 7981.1    | 2.80     | .125     |
| Run Type <sup>2</sup> | 4         | 1145.8    | 286.4     | 0.98     | .430     |
| Group X Run           | 4         | 1781.8    | 445.4     | 1.52     | .214     |
| Trainee               | 10        | 28472.3   | 2847.2    |          |          |
| Run X Trainee         | 40        | 11713.7   | 292.8     |          |          |

1. Group = ACE, Synthetics

2. Run Type = FLT NCI, TRNG NCI, EMERGENCY, TANKERS, NCIC

## SECTION V

## COST-BENEFIT ANALYSIS

## THE COST-BENEFITS OF ACE

A major objective of this evaluation was to estimate the potential cost-benefit of ACE as a substitute for the current air intercept controller system. Based on the findings presented in the previous chapters, we must assume, at the present time, that the training effectiveness of the two system alternatives is equivalent. Therefore, there is no direct comparison possible of the relative cost-effectiveness of the two systems. On the other hand, it is reasonably clear that there are certain benefits to be gained from an operational version of the ACE as well as certain costs of acquiring it. While it is difficult to estimate the precise costs of an acceptable automated system, the basic configuration is clear as are the general cost consequences associated with it (cf., Fischhoff, 1977).

**PERIOD OF PERFORMANCE.** While it is customary to estimate life cycle costs over a 20-year period, it does not seem desirable to do so in this particular case. Potential changes in technology and associated operational AIC tasks alone make that period of time questionable. Based on operational estimates, it has been assumed here that a 10-year cycle time is about as far as should be estimated. Therefore, a 10-year life cycle costing time period has been assumed for this analysis. In actual fact, assuming a continuation of the existing peacetime training demands, the training requirements seem very clear for the next five years (1982-1987) with increasing uncertainty afterward. Should the training demands shift to wartime conditions, an entirely new situation would, of course, prevail. In the meantime, predictions from existing conditions seem rather more clear than most such estimating situations.

**THE STUDENT LOAD.** The primary driving function for this analysis is the number of anticipated students which must be trained as air intercept controllers, that is, the determination of anticipated training load (cf., String and Orlansky, 1977, p.68). For the training installation in which this evaluation was conducted, all current data project an average student load of four. With a new class beginning every two weeks, this equates to a total of about 100 students per year.\* Current planning for AIC training is based on this estimate of the student load, and it will be used here. At any given time, the variability of student load may be expected to range from zero to six. But, it is apparently acceptable to assume that any system must be capable of handling at least four students at any given time.

\* This is approximately half of the total Navy AIC training requirement which is accomplished by two installations. The analysis here is for one of these two training facilities, but could be easily expanded to cover both, if the total Navy training requirements for AICs were of interest.

SYSTEM SUBSTITUTION. The ACE is, in a sense, a substitution for the existing training system. But, as presently configured, it is not a complete replacement for the total AIC training program. Figure 5 briefly sketches the training sequence as it is currently devised. It may be seen that the ACE technologies are considered here only for the second major element in the AIC training program, i.e., the Synthetics simulation. Since the evaluation was done within this context, the cost-benefit analysis follows the same model.

Future ACE systems could be expanded beyond the scope shown in Figure 5. For example, it is possible to substitute current classroom training (or at least some parts of it) with the ACE CAI sub-system. Further, it could be feasible to project live targets on the ACE operator's console. In short, it is not impossible to conceive of an expanded ACE version that could serve as a complete substitute for the current AIC training program. Further expansion could include AIC "conventionals" (non-NTDS) and Anti Submarine Air Controller (ASAC) training. Such an expansion of ACE, however, is not possible with the existing hardware-software configuration. Accordingly, the present analysis concentrates on ACE as a substitute for the current Synthetics step in the training program. However, this is in fact the major element of the training stream where ACE is probably most effective.

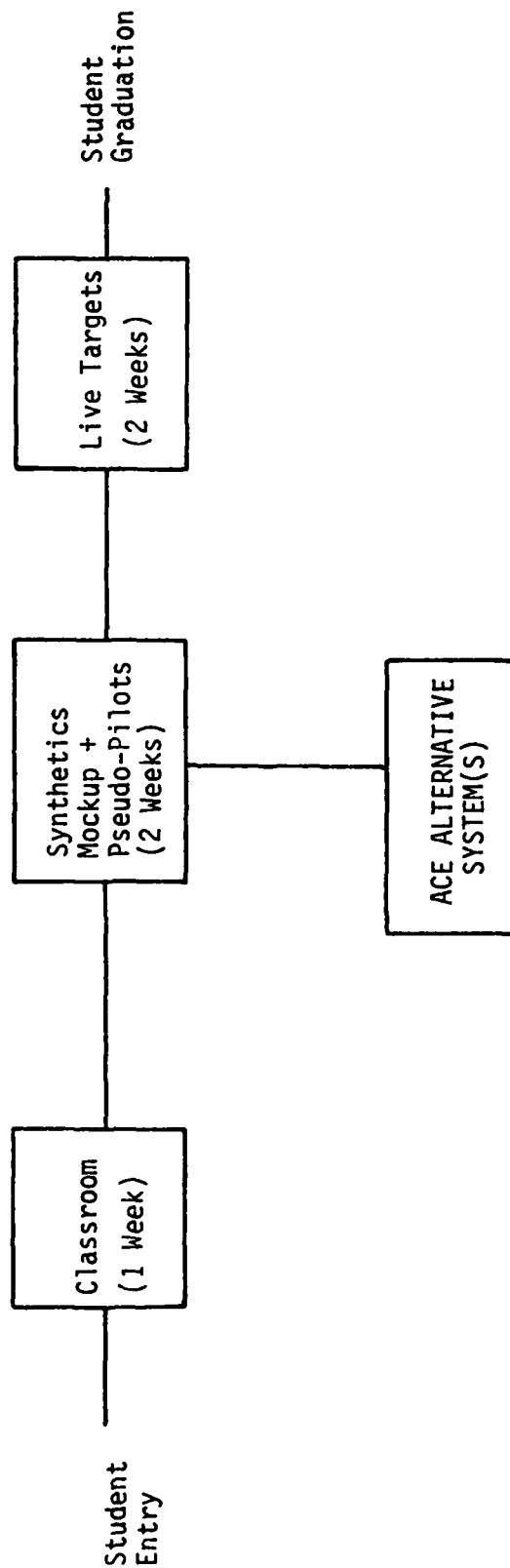
#### PRINCIPAL COST COMPONENTS

For the cost-benefit analysis, a number of significant cost components must be estimated for both the current and the ACE system (cf. String and Orlansky, 1977, p.84). These include:

1. Assigned officer personnel
2. Assigned enlisted instructor personnel
3. Pseudo-pilots
4. Civilian support personnel
5. Military maintenance personnel
6. Civilian maintenance personnel
7. Maintenance and facilities costs

We are restricted here to those specific costs in each of these categories that may be assigned specifically to the Synthetics training segment shown in Figure 5.

ACE ACQUISITION COSTS. A major factor for the ACE system is the cost of acquiring an operational ACE configuration. It is assumed here that four stand-alone ACE consoles would be required to handle the anticipated student load. At an estimated \$400,000 per console, the total



#### Instructional Features

1. Speech Generation
2. Video disc CAI
3. No pseudo-pilots
4. Automated Performance Measurement

Figure 5. Current AIC and ACE Alternative System Training Configuration

(Note: This is not necessarily the optimum training flow, and changes may be instituted by FLECOMBATRACENPAC concerning training for non-NTDS familiarization)

hardware/software cost for acquisition would be \$1,600,000. Further, the results of this study clearly show that changes in the system are required; the cost for reprogramming and software changes has been estimated at approximately \$400,000. Thus, the estimated total acquisition cost for an operational ACE configuration is \$2,000,000.

**PERSONNEL SAVINGS.** A major anticipated benefit resulting from the ACE system is in personnel savings. The question is: What specific labor categories are affected and to what degree? Table 17 shows an estimate of labor categories and changes that may occur within the introduction of the ACE system. These estimates are based on: (1) analysis of the task loads; (2) discussions with the AIC training personnel; and (3) conservative estimates of a working installation. These figures are based on a working force without backup for equipment down-time or supplementary personnel.

TABLE 17. PERSONNEL LOADING FOR CURRENT AND ACE SYSTEMS

| Labor Category        | <u>Current System</u> |                          | <u>ACE System</u> |            |
|-----------------------|-----------------------|--------------------------|-------------------|------------|
|                       | Number                | Level                    | Number            | Level      |
| Officer               | 2                     | O-4<br>O-3               | 1                 | O-3        |
| Enlisted Instructors  | 4                     | E-9<br>E-8<br>E-7<br>E-6 | 2                 | E-8<br>E-7 |
| Pseudo-Pilots         | 6                     | E-3                      | 0                 | 0          |
| Civilian Support      | 1                     | GS-3                     | 0                 | 0          |
| Maintenance Personnel |                       |                          |                   |            |
| Military              | 2                     | E-7<br>E-6               | 2                 | E-7<br>E-6 |
| Civilian              | 1                     | GS-11                    | 1                 | GS-11      |

It is apparent from Table 17 that significant savings in the areas of enlisted instructors and pseudo-pilots are anticipated by the introduction of ACE. On the other hand, no savings are anticipated in maintenance personnel.

Comparative 10-Year Operating Costs. It is now possible to make an estimated comparison of the 10-year operating costs of the current and ACE systems. This comparison is shown in Table 18. As may be seen in Table 18, there are eight principal cost components in the comparison. For estimating the 10-year costs of military personnel, we have used the enlisted life cycle cost projections generated by Koehler and Turney (1981) and for officer costs the estimates provided by Koehler (1980). As may be seen in these sources, a 10 percent discount rate is assumed as required by Office of Management and Budget Circular A-94.

TABLE 18. COMPARATIVE 10-YEAR OPERATING COSTS

| Cost Component              | Current System                           | ACE System                   |
|-----------------------------|------------------------------------------|------------------------------|
| Officer Personnel*          | \$ 234,207<br>186,340                    | \$ -<br>186,340              |
| Instructor Personnel **     | 215,701<br>194,059<br>170,287<br>150,274 | -<br>194,059<br>170,287<br>- |
| Pseudo-Pilots**             | 597,162                                  | -                            |
| Civilian Support            | 110,000                                  | -                            |
| Military Maint. Personnel** | 183,292<br>158,121                       | 183,292<br>158,121           |
| Civilian Maint. Personnel   | 270,000                                  | 270,000                      |
| Repair Parts                | 600,000                                  | 600,000                      |
| Utilities and Facilities    | 1,000,000                                | 1,000,000                    |
| TOTAL                       | \$4,069,443                              | \$2,762,009                  |

\* Based on data from Koehler (1980)

\*\* Based on data from Koehler and Turney (1981)

Estimating repair, utility, and facility costs is very difficult. Cost data for the current training system, are available and these data have been used to generate estimates. However, we have little reliability data on the ACE system,\* and, hence, no way of confidently predicting maintenance and spare parts requirements. For this analysis, therefore, we have assumed that the two systems are equivalent with respect to maintenance requirements.

The estimates provided in Table 18 suggest that a significant operating cost savings could be obtained from ACE. The savings over a 10-year period would be \$1,307,344. In addition a significant "value" would be the reduced requirement for highly trained military personnel.

**AMORTIZATION OF ACQUISITION COSTS.** Unfortunately, the savings in operating costs must be compared with the acquisition costs of the ACE system. As noted before in this Chapter, it is estimated that ACE acquisition cost may approximate \$2,000,000. Thus, the acquisition costs could not be amortized over a 10-year period; indeed, some 15 years would be needed to absorb the estimated acquisition costs.

On the other hand, the cost of computer technology is dropping so radically in many areas that the question might be raised: What console cost would present a favorable amortization situation? Holding reprogramming costs constant at \$400,000 and continuing to assume four ACE consoles, the ACE hardware/software cost would have to be reduced almost in half (to \$907,340) to amortize acquisition costs over a 10-year operating period.

#### REDUCING AUTOMATED INSTRUCTION FEATURES

The ACE prototype system is designed to provide a substantial number of automated features for training with two general objectives: (1) eliminate the need for human pseudo-pilots; and (2) enhance instructor effectiveness and reduce instructor workload. One might consider a limited ACE configuration which concentrates only on the first goal. This limited ACE system would include only the speech interaction system and would not include the videodisc, CAI, and automated performance measurement system.

It is possible to iterate the analysis shown in Table 18 for a comparison of 10-year operating costs with the current system and a limited speech interaction ACE subsystem. Table 19 shows this comparison.

---

\* Some reliability data are available from experience gained in very early system development (cf., Grady, 1982, Appendix B) and probably are poor estimators of operational reliability. Even so, "...the MTBF was 25 days and the MTTR was 2 days.

TABLE 19. COMPARATIVE 10-YEAR OPERATING COSTS: LIMITED ACE SYSTEM

| Cost Component            | Current System                           | Limited ACE                              |
|---------------------------|------------------------------------------|------------------------------------------|
| Officer Personnel         | \$ 234,207<br>186,340                    | \$ 234,207<br>186,340                    |
| Instructor Personnel      | 215,701<br>194,059<br>170,287<br>150,274 | 215,701<br>194,059<br>170,287<br>150,274 |
| Pseudo-Pilots             | 597,162                                  | -                                        |
| Civilian Support          | 110,000                                  | -                                        |
| Military Maint. Personnel | 183,292<br>158,121                       | 183,292<br>158,121                       |
| Civilian Maint. Personnel | 270,000                                  | 270,000                                  |
| Repair Parts              | 600,000                                  | 600,000                                  |
| Utilities and Facilities  | 1,000,000                                | 1,000,000                                |
| TOTAL                     | \$4,069,443                              | \$3,362,281                              |

The relative totals in Table 19 suggest a 10-year savings of about \$707,162 simply by eliminating the pseudo-pilots and the civilian support. This seems to be a significant savings, and has the added value that the training system can eliminate the very difficult requirement of continuously providing human pseudo-pilots.

AMORTIZATION OF LIMITED ACE ACQUISITION COSTS. It is difficult to estimate a reasonable cost figure for a speech interaction simulation system for this application; the technology is changing very rapidly, competition is intense, and costs are closely-held proprietary information. Another way of approaching the problem is to estimate the maximum allowable cost for four speech interaction systems as a function of the amortization period:

1. For a five-year amortization period, the savings available from operating costs would be \$353,581. That would place the per-unit speech interaction system cost at about \$88,395.

2. For a 10-year amortization period, the savings available from operating costs would be \$707,162. The per-unit cost of a four unit system then would be \$176,791.

The current technology costs are already generally within the range of these figures. This assumes that the software from the current ACE prototype system would be usable in developing the four speech interaction systems.

## SECTION VI

## CONCLUSIONS AND RECOMMENDATIONS

ACE was a complex system that integrated several emerging technologies into an experimental prototype training system for AICs. It was a smorgasboard of new techniques aimed at reducing the manloading of the instructional staff, while providing effective, standardized AIC training. How effectively it accomplished its objectives depends on one's perception of the objectives, as stated at the outset - was it to demonstrate the potential of new technologies for training or to accomplish effective AIC training by means of an automated system?

## CONCLUSIONS

TRAINING TECHNOLOGIES DEMONSTRATED. ACE successfully demonstrated the potential benefits of several new technologies for training. Interactive simulation of the AIC task was accomplished through the use of computer speech recognition, simulation of the operational NTDS console, and simulation of the scenarios and response characteristics of the pilot/aircraft and other tactical personnel. This interactive simulation demonstrated the capability to eliminate the human pseudo-pilot, who currently is required to support AIC training.

The conceptual design of ACE demonstrated how an instructor's workload could be reduced by automating several of his responsibilities, such as instruction, performance evaluation, performance feedback, syllabus control, and record keeping.

Some of the potential benefits of automated instruction are to allow trainees to progress at their own rate, provide consistent standardized training over extended periods of time, and enable trainees to use the system at any time of day or night without the need for extensive personnel support.

ACE demonstrated the potential for combining new technologies to provide the instruction and simulation environment necessary for the acquisition of knowledge and skill through practice.

The best features of ACE were:

- o Instructional presentation
- o Training Enhancement Console
- o Multi-media video
- o Simulated intercept scenarios

FURTHER DEVELOPMENT REQUIRED. Incomplete development of ACE left a number of problems to be resolved. Because it was a highly connected system, errors in one new technology were passed along, and perhaps amplified, by other technologies or subsystems. Thus, at the time of the evaluation, ACE was performing considerably below its potential.

The major problem areas, as discussed in detail elsewhere in the report, were the following:

- o Speech recognition accuracy
- o Performance measurement accuracy
- o Commented practice freezes
- o Performance feedback techniques

These problem areas were not independent. Improved performance of the speech recognition system and the PMVs would ameliorate many of ACE's problems. Suggestions for improvements have been given earlier in this report.

COST BENEFIT. The major cost benefits provided by the ACE technologies are due to the reduction of personnel costs, particularly, elimination of the pseudo-pilot position. When applied only to the two-week Synthetics portion of AIC training, the development costs would appear to be amortized over a period of 10 to 15 years. If the ACE technologies were applied to related areas of training such as "conventionals" and ASAC, the cost benefits would be greater because of shared development costs.

SPEECH RECOGNITION IN TRAINING. Is speech recognition technology sufficiently advanced to apply it to sophisticated, operational training systems? Based on the ACE data, we must respond, "not quite." Several fundamental design principles must be established regarding human factors issues in speech recognition systems. It is likely that ACE would have performed considerably better if: (1) speech data collection had been done in the context of the AIC job, (2) synthesized speech had not been used for prompts during speech data collection, (3) vocabulary and pausing constraints had been tested and revised, and (4) more training had been given to naive users on how (and how not) to speak to the system. It is likely that the speech recognition technology is sufficiently advanced for many training applications, but that extreme care must be given to the human factors implementation of the technology for each specific application.

Interactive training applications of AST are particularly difficult because the users are: (1) short term (e.g., nine to ten days on ACE); (2) naive with respect to speech recognition technology; (3) naive with respect to the vocabulary and the task; and (4) subjected to the stresses of interactive voice control and performance evaluation.

Because of these peculiar difficulties associated with AST in interactive training, system development should include frequent checks of speech system performance with the actual user population.

DEVELOPMENT OF SPEECH INTERACTIVE TRAINING SYSTEMS. Much has been learned from the development of experimental prototype systems such as PARTS and ACE. In proceeding to operational training systems, however, a plan for the implementation of speech technology is required. Perhaps the phased integration of the associated technologies would be the optimal strategy, with a series of TEAR cycles as the level of automation is increased. For example: (1) develop the system using speech recognition to drive the interactive simulation (replacing pseudo-pilots); (2) deliver and use the system at the school while gathering information on needed revisions and collecting data on instructors' grading and syllabus control procedures; (3) make necessary revisions, expand the role of speech to support automated performance measurement, and provide syllabus control options to the users; and (4) repeat the cycle of using the system while evaluating it, making revisions as needed. This final step would be a continuing process, to keep the system current with changing instructional needs, thereby extending its life cycle.

Advances in basic R&D are necessary to extend the fruitful applications of speech recognition technology. The basic thrust of the research should be to develop hardware, software, or applications techniques that will result in high recognition accuracy under a wide range of conditions, including unpracticed speakers in stressful task situations.

#### RECOMMENDATIONS

SPEECH TECHNOLOGY VIABLE. Automated speech technology should be considered a viable candidate for application to AIC training, including "conventionals," and anti-submarine air controllers (ASAC), as well as NTDS-based AIC. This recommendation is based on the fact that, despite the implementation problems of AST in ACE, (1) the technology is advancing rapidly, (2) lessons have been learned from the ACE experience, and (3) advancements at other facilities are applicable (cf., Naval Ocean Systems Center, Naval Air Development Center, Naval Air Test Center, Naval Surface Weapons Center, and various civilian organizations).

ACE IS AN R&D RESOURCE. ACE should be used for further R&D on speech technology applications for Navy training. This recommendation is based on the potential for significant advances in system performance that are likely to accrue from the investment of relatively limited additional resources. ACE represents a substantial investment in very promising new technologies for training. Large improvements in system performance could be obtained from further development of the speech recognition and performance measurement systems. The techniques learned from these improvements are directly relevant to future implementation of these technologies in operational training systems. Return on the investment in ACE can best be realized by investigating techniques for improving system performance. Suggestions have been made in this report for potential R&D on ACE improvements.

PHASED INTEGRATION OF SPEECH TECHNOLOGY. Future applications of automated speech technology for interactive training systems should be based on a plan of phased integration as discussed in the conclusion section of this report.

REFERENCES AND BIBLIOGRAPHY

- Breaux, R. Training Characteristics of the Automated Adaptive Ground Controlled Approach Radar Controller Training System (GCA-CTS). Technical Note NAVTRAEQUIPCEN TN-52. Naval Training Equipment Center, Orlando, FL, July 1976.
- Breaux, R., Curran, M., and Huff, E. (Eds.) Voice Technology for Interactive Real-Time Command/Control System Application. NASA-Ames Research Center, Moffett Field, CA, 1978.
- Breaux, R. and Goldstein, I. Developments of Machine Speech Understanding for Automated Instructional Systems. In Eighth NTEC/Industry Conference Proceedings. Naval Training Equipment Center, Orlando, FL, November 1975.
- Brightman, T. and Crook, S. Exploring Practical Speech I/O. Mini-Micro Systems, May 1982, 291-303.
- Chatfield, D.C. and Gidcumb, C.F. Optimization Techniques for Automated Adaptive Training Systems. Technical Report NAVTRAEQUIPCEN 77-M-0575. Naval Training Equipment Center, Orlando, FL, 1977.
- Chatfield, D.C. Klein, G.L. and Coons, D. Instruct: An Example of the Role of Artificial Intelligence in Voice-Based Training Systems. NAVTRAEQUIPCEN 80-C-0061-1. Naval Training Equipment Center, Orlando, FL, September 1981.
- Chatfield, D.C., Marshall, P.H., and Gidcumb, C.F. Instructor Model Characteristics for Automated Speech Technology. Technical Report NAVTRAEQUIPCEN 79-C-9985. Naval Training Equipment Center, Orlando, FL, 1979.
- Clark, W., Halley, R., Regelson, E., Slemon, G., Ver Steeg, R., Functional Design for Air Intercept Controller Prototype Training System. Technical Report NAVTRAEQUIPCEN 78-C-0182.8. Naval Training Equipment Center, Orlando, FL, November, 1979.
- Cotton, J.C. and McCauley, M.E. Voice Technology Design Guides for Navy Training Systems. Technical Report NAVTRAEQUIPCEN 80-C-0057-1. Naval Training Equipment Center, Orlando, FL, January, 1981.
- Dixon, N.R. and Martin, T.B. (Eds.) Automatic Speech and Speaker Recognition. New York: IEEE Press, 1979.
- Evans, C. The Micro Millenium. New York: Washington Square Press, 1979.
- Fischhoff, B. The Art of Cost-Benefit Analysis. DARPA: Perceptronics Technical Report PTR-1042-77-2, February, 1977.

- Grady, M.W. Air Intercept Controller Prototype Training System. Technical Report NAVTRAEQUIPCEN 78-C-0182-14. Naval Training Equipment Center, Orlando, February, 1982.
- Grady, M.W., Halley, R., and Nowell, L.H. Ordinal Syllabus for Air Intercept Controller Prototype Training System. Technical Report NAVTRAEQUIPCEN 78-C-0182-3. Naval Training Equipment Center, Orlando, June, 1981.
- Granberry, R.D., Halley, R., Kerr, D., King, M.R., and Regelson, E.C. Prototype Equipment Instructor Handbook for ACE (Air Intercept Controller Prototype Training System). Technical Report NAVTRAEQUIPCEN 78-C-0182-10. Naval Training Equipment Center, Orlando, June 1981a.
- Granberry, R.D., Halley, R., Kerr, D., King, M.R., and Regelson, E.C. Prototype Equipment Student Guide for ACE (Air Intercept Controller Prototype Training System). Technical Report NAVTRAEQUIPCEN 78-C-0182-9. Naval Training Equipment Center, Orlando, 1981b.
- Halley, R. Training Effectiveness Report for Air Intercept Controller Prototype Training System. Technical Report NAVTRAEQUIPCEN 78-C-0182-12. Naval Training Equipment Center, Orlando, June, 1981.
- Halley, R., Hooks, J.T., Lankford, H.G., and Nowell, L.H. Behavioral Objectives for Air Intercept Controller Prototype Training System. Technical Report NAVTRAEQUIPCEN 78-C-0182-1. Naval Training Equipment Center, Orlando, FL, February, 1979.
- Halley, R., Hooks, J.T., Lankford, H.G., and Nowell, L.H. Objectives Hierarchy for Air Intercept Controller Prototype Training Systems. Technical Report NAVTRAEQUIPCEN 78-C-0182-2. Naval Training Equipment Center, Orlando, FL, 1981.
- Halley, R., King, M.R., and Regelson, E.C. Functional Requirements for Air Intercept Controller Prototype Training System. Technical Report NAVTRAEQUIPCEN 78-C-0182-4. Naval Training Equipment Center, Orlando, FL, 1979.
- Halley, R., King, M. and Regelson, E.C. Prototype Test Plan and Demonstration Report for Air Intercept Controller Prototype Training System. Technical Report NAVTRAEQUIPCEN 78-C-0182-11. Naval Training Equipment Center, Orlando, December, 1980.
- Hicklin, M., Barber, G., Bollenbacher, J., Grady, M., Harry, D., Meyn, C., and Slemon, G. Ground Controlled Approach Controller Training System (GCA-CTS). Technical Report NAVTRAEQUIPCEN 77-C-0162-6. Naval Training Equipment Center, Orlando, April 1980.

- Joplin, L. Interactive Logs for Automated Speech Technology. Technical Report NAVTRAEQUIPCEN 77-C-0162-6. Naval Training Equipment Center, Orlando, April 1980.
- Kaplan, G. Words into Action: I. IEEE Spectrum, June 1980, 22-26.
- Koehler, E.A. Life Cycle Navy Officer Billet Costs -- FY80. U.S. Navy: NPRDC SR 80-18, May 1980.
- Koehler, E.A. and Turney, R.F. Life Cycle Navy Enlisted Billet Costs -- FY81. U.S. Navy: NPRDC SR 81-22, July 1981.
- Lea, W.A. (Editor). Trends in Speech Recognition, Englewood Cliffs: Prentice-Hall, 1980.
- Marx, R.J. Videodisc-Based Training: Does it Make Economic Sense? Training/HRD, March, 1982, 56-65.
- McCauley, M.E. and Semple, C.A. Precision Approach Radar Training System (PARTS) Training Effectiveness Evaluation. Technical Report NAVTRAEQUIPCEN 79-C-0042-1. Naval Training Equipment Center, Orlando, August, 1980.
- Michaelis, P.R. and Wiggins, R.H. Speech Synthesizers: The Basics. Human Factors Society Bulletin, 1981, Vol. 24(4), 1-2.
- Poock, G.K. Experiments with Voice Input for Command and Control. Technical Report NPS-55-80-016, Naval Postgraduate School, Monterey, CA, 1980.
- Simpson, C.A. and Williams, D.H. Response Time Effects of Alerting Tone and Sernantic Context for Synthesized Voice Cockpit Warnings. Human Factors, 1980, 22(3), 319-330.
- String, J. and Orlansky, J. Cost Effectiveness of Flight Simulators for Military Training, Vol. II: Estimating Costs of Training in Simulators and Aircraft. IDA Paper P-1275. Arlington, VA: Defense Advanced Projects Agency, August, 1977.
- Toffler, A. The Third Wave. New York: William Morrow, 1980.
- U.S. Navy, Curriculum Outline for Air Intercept Controller Conventional Qualification/Air Intercept Controller NTDS Qualification. Fleet Combat Training Center Pacific, San Diego, CA, July, 1980.
- Van Hemel, P.E., Van Hemel, S.B., King, W.J., and Breaux, R. Training Implications of Airborne Applications of Automated Speech Recognition Technology. Technical Report NAVTRAEQUIPCEN 80-C-0009-0155-1. Naval Training Equipment Center, Orlando, October, 1980.

ACRONYMS

AAW - Anti Air Warfare  
ACE - Air Controller Exerciser  
AIC - Air Intercept Controller  
ANOVA - Analysis of Variance  
ASAC - Anti Submarine Air Controller  
AST - Automated Speech Technology  
CAI - Computer Aided Instruction  
CAP - Combat Air Patrol  
CIC - Combat Information Center  
CP - Commented Practice  
C/S - Call Sign  
CSR - Connected Speech Recognition  
FLECOMATRACENPAC - Fleet Combat Training Center Pacific  
FP - Free Practice  
GCA-CTS - Ground Controlled Approach-Controller Training System  
(Same as PARTS)  
IAT - Interactive Teaching  
I/O - Input/Output  
IWR - Isolated Word Recognition  
MTBF - Mean Time Between Failure  
MTTR - Mean Time to Replacement  
NAVTRAEQUIPCEN - Naval Training Equipment Center  
NCI - Nearest Collision Intercept  
NCIC - Nearest Collision Intercept Conversion

NEC - Nippon Electric Company  
NM - Nautical Mile  
NORDO - No Radio  
NTDS - Naval Tactical Data System  
OS - Operations Specialist  
PACTS-AIC - Prototype Air Controller Training System for Air Intercept  
Controllers (same as ACE)  
PARTS - Precision Approach Radar Training System  
PC&E - Problem Control and Evaluation  
PMV - Performance Measurement Variable  
R&D - Research and Development  
RFP - Request for Proposal  
SDC - Speech Data Collection  
SME - Subject Matter Expert  
SOW - Statement of Work  
SPEECH I/O - Speech Input/Output  
SUS - Speech Understanding Software  
SWC - Ships Weapons Coordinator  
TEAR - Test, Evaluate and Revise  
TEC - Training Enhancement Console  
TOT - Transfer of Training  
VDU - Visual Display Unit

APPENDIX A  
THE EXIT QUESTIONNAIRE

## QUESTIONNAIRE

This questionnaire is part of the research project to evaluate the ACE trainer. The answers to these questions will not be used on your official Navy records or effect your career in the Navy.

Please answer the following questions using the numbers 1 to 5. You should answer with your opinion and/or experience. There is no right answer.

- 1 - Terrible
- 2 - Poor
- 3 - Neutral
- 4 - Good
- 5 - Excellent

- 3 1. Your training is not over, but at this point how well do you think you are doing?
- 3 2. How would you rate the help you received from your instructors during your time on the ACE trainer?
- 3 3. What did you think of the voice used by the computer in ACE.
- 2 4. How well did the ACE system recognize your voice?
- 3 5. How do you feel about practicing your AIC skills with a computer rather than real people?
- 3 6. How do you feel the instructors treated you compared to students not on ACE?
- 4 7. Overall what do you think of the ACE trainer?
- 4 8. How do you feel about becoming an AIC at this point?
- 4 9. What do you think of computer training systems in general?
- 3 10. How would you evaluate the performance evaluation information that ACE presented to you after each practice run?
- 4 11. What would be your evaluation of the training you have received on ACE?
- 3 12. The ACE computerized performance measurement is?
- 3 13. What did you think of the video presentations in ACE?
- 2 14. How was the voice response time in the ACE trainer?

## EXIT QUESTIONNAIRE DATA SUMMARY

### MEAN RATINGS BY QUESTION

| <u>Question Number</u> | <u>Sample Size</u> | <u>Mean Rating</u> |
|------------------------|--------------------|--------------------|
| 1                      | 7                  | 3.28               |
| 2                      | 7                  | 4.14               |
| 3                      | 7                  | 3.00               |
| 4                      | 7                  | 2.43               |
| 5                      | 7                  | 3.00               |
| 6                      | 7                  | 3.42               |
| 7                      | 7                  | 3.28               |
| 8                      | 7                  | 4.14               |
| 9                      | 7                  | 4.00               |
| 10                     | 7                  | 2.43               |
| 11                     | 7                  | 3.57               |
| 12                     | 6                  | 2.83               |
| 13                     | 6                  | 4.16               |
| 14                     | 6                  | 2.00               |

### EXIT QUESTIONNAIRE HIGHLIGHTS

The Exit Questionnaire revealed the attitudes of the trainees with respect to the various features of ACE. Four items received an average rating below 3 (Neutral): Recognition Accuracy (2.43), Performance Feedback Information (2.43), Computerized Performance Measurement (2.83), and Voice Response Time (2.0). All of these issues have been discussed in the body of the report.

By contrast, three items received an average rating of at least 4 (Good): Help Received From Instructors (4.14), Opinion of Computer Training System in General (4.0), and Video Presentations in ACE (4.16).

NAVTRAEQUIPCEN 81-C-0055-1

APPENDIX B

PERFORMANCE MEASUREMENT  
VALIDATION STUDY: GRADING FORM

ACE GRADING SHEET  
FORWARD QUARTER

INSTRUCTOR MEYER  
STUDENT

| SEG # | SET UP |      | COM'S |       | CONTROL |     |     |         | FINAL    | TOTAL  | SAFETY | TOTAL                     |
|-------|--------|------|-------|-------|---------|-----|-----|---------|----------|--------|--------|---------------------------|
|       | A&B    | TURN | PPOI  | STATE | C/A     | T/S | B/D | RESPOND | SEP AREA | RECOMP | TRACK  |                           |
| 13    | 5      | 5    | 5     | 2     | 5       | 3   | 10  | 5       | 10       | 10     | 10     | MP-2=20<br>2-3=10<br>3+=0 |
| 21    | 5      | 5    | 5     | 0     | 5       | 3   | 8   | 5       | 5        | 10     | 5      | 20                        |
|       |        |      |       |       |         |     |     |         |          |        |        | 86                        |
|       |        |      |       |       |         |     |     |         |          |        |        | -40                       |

900 NOT BREAKAWAY PORTION STABBOARD

AREA NOT IN

OUT OF AREA

POOR TRACKING

A/C RESPONSE TO HEADINGS

TOO MUCH SEP/OUT OF AREA DUE TO

|    |   |   |   |   |   |   |   |   |    |   |   |    |     |   |
|----|---|---|---|---|---|---|---|---|----|---|---|----|-----|---|
| 13 | 5 | 5 | 5 | 2 | 0 | 0 | 5 | 0 | 10 | 0 | 0 | 32 | -40 | 0 |
|----|---|---|---|---|---|---|---|---|----|---|---|----|-----|---|

TOO MANY POINTS ON SCOPES

DIDN'T LATCH CS HEADING OF 107 VICE 307

DOESN'T LISTEN TO A/C

OUT OF AREA DUE TO BAD HEADING (107)

CHANGED BOG HEAD TO 140

|  |   |   |   |   |   |   |   |   |   |    |   |    |    |     |    |
|--|---|---|---|---|---|---|---|---|---|----|---|----|----|-----|----|
|  | 0 | 0 | 5 | 0 | 5 | 3 | 8 | 0 | 0 | 10 | 5 | 20 | 66 | -40 | 26 |
|--|---|---|---|---|---|---|---|---|---|----|---|----|----|-----|----|

DOESN'T WAIT FOR A/C TO RESPOND

86

COPIES OF

#2

NOT USING PREPARED PLANNING

|   |   |   |   |   |   |   |   |   |    |    |   |    |    |   |    |
|---|---|---|---|---|---|---|---|---|----|----|---|----|----|---|----|
| 3 | 5 | 5 | 5 | 2 | 5 | 3 | 8 | 3 | 10 | 10 | 5 | 20 | 91 | - | 91 |
|---|---|---|---|---|---|---|---|---|----|----|---|----|----|---|----|

25

point 146 as target

STILL NOT LISTENING TO A/C

will the form

TRACK BETTER

|   |   |   |   |   |   |   |   |   |    |    |   |    |    |   |    |
|---|---|---|---|---|---|---|---|---|----|----|---|----|----|---|----|
| 4 | 5 | 5 | 5 | 2 | 5 | 3 | 8 | 0 | 10 | 10 | 8 | 20 | 91 | - | 91 |
|---|---|---|---|---|---|---|---|---|----|----|---|----|----|---|----|

TOLD CS TO BREAKAWAY 125

DIDN'T LET SH. REPLY TO HEADING (twice)

STILL NOT TRACKING AS REQUESTED

# ACE GRADING SHEET

JOIN UPS

STUDENT:

INSTRUCTOR: M E YER

| RUN # | SET UP  | COMMS   |         | CONTROL |     |     |     |        |         |      | FINAL TOTAL | SAFETY AREA | TOTAL |       |
|-------|---------|---------|---------|---------|-----|-----|-----|--------|---------|------|-------------|-------------|-------|-------|
|       | LAT SEP | VEC CAP | VEC MAC | C2M     | M2C | MAC | ALT | VECTOR | LAT SEP | SIDE |             |             |       | CK PT |
|       | 15      | 5       |         | 5       | 5   | 5   |     | 5      | 10      | 10   | 15          | 20          | 100   | 85    |
| 1     | 15      | 5       | 5       | 5       | 0   | 5   | 5   | 5      | 10      | 10   | 15          | 10          | 85    |       |

use 4 miles for lat sep. (ake actual 4 turns well fore!)  
(all runs) go ahead and give c2m calls on lat sep line!

040  
265

TURN

|   |    |   |   |   |   |   |   |   |    |    |    |   |    |
|---|----|---|---|---|---|---|---|---|----|----|----|---|----|
| 2 | 15 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 10 | 10 | 15 | 0 | 80 |
|---|----|---|---|---|---|---|---|---|----|----|----|---|----|

10 used 4 miles for lat sep.

OR

turned inside CAP TRACK

070/4 from CAP-0 MAC

105

CS TURNED FIRST/ T/E DID W HAT A E

GOULD TO ASPECT?

|   |    |   |   |   |   |   |   |   |    |    |    |   |    |
|---|----|---|---|---|---|---|---|---|----|----|----|---|----|
| 3 | 15 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 10 | 10 | 15 | 0 | 80 |
|---|----|---|---|---|---|---|---|---|----|----|----|---|----|

100 used 4 miles lat sep

CAP

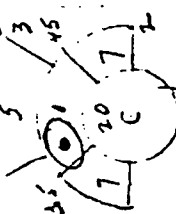
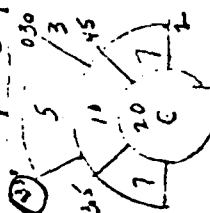
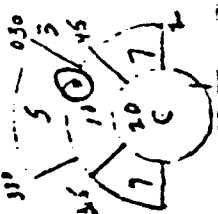
145

|   |    |   |   |   |   |   |   |   |    |    |    |    |    |
|---|----|---|---|---|---|---|---|---|----|----|----|----|----|
| 4 | 15 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 10 | 10 | 15 | 10 | 90 |
|---|----|---|---|---|---|---|---|---|----|----|----|----|----|

410 new 4 miles lat sep

CS any 20 vice 22

040



APPENDIX C  
SAMPLE GRADING SHEET  
FOR TOT-TEST

DATE: 8 SEPT 81

START TIME: 0640

STOP TIME: 1110

NO STUDENT GRADE SHEET

31 INSTRUCTOR: N. J. Sp

PILOT: Dave CIV

STUDENT:

GRADER: QICM CROSS

CHECK IN: CONTACT 110 STATE L NORDO L VECTOR TO AREA L

| TYPE RUN     | SET UP  | RECOMPS | COMMS | CONSOLE MECHANICS | SVC INTERFACE | BOGEY JINK | STRANGERS | AREA CONTROL |
|--------------|---------|---------|-------|-------------------|---------------|------------|-----------|--------------|
| 1. 10(TING)  | - 0 + 5 | * 6 *   | * 4 * | * 8 *             | - 0 + 1       | - 0 + 3    | - 0 + 5   | - 0 + 1      |
| 2. 10(TING)  | * 22 *  | * 6 *   | * 4 * | * 8 *             | - 0 + 1       | - 0 + 3    | * 20 *    | * 10 *       |
| 3. 10(TING)  | * 10 *  | * 6 *   | * 4 * | * 8 *             | - 0 + 1       | - 0 + 3    | * 20 *    | * 10 *       |
| 4. 10(TING)  | * 20 *  | * 12 *  | * 4 * | * 8 *             | - 0 + 1       | - 0 + 3    | * 20 *    | * 20 *       |
| 5. 10(TING)  | * 20 *  | * 6 *   | * 4 * | * 8 *             | - 0 + 1       | - 0 + 3    | * 20 *    | * 20 *       |
| 6. 10(TING)  | * 20 *  | * 12 *  | * 4 * | * 8 *             | - 0 + 1       | - 0 + 3    | * 20 *    | * 20 *       |
| 7. 10(TING)  | * 20 *  | * 12 *  | * 4 * | * 8 *             | - 0 + 1       | - 0 + 3    | * 20 *    | * 20 *       |
| 8. 10(TING)  | * 20 *  | * 12 *  | * 4 * | * 8 *             | - 0 + 1       | - 0 + 3    | * 20 *    | * 20 *       |
| 9. 10(TING)  | * 20 *  | * 12 *  | * 4 * | * 8 *             | - 0 + 1       | - 0 + 3    | * 20 *    | * 20 *       |
| 10. 10(TING) | * 20 *  | * 12 *  | * 4 * | * 8 *             | - 0 + 1       | - 0 + 3    | * 20 *    | * 20 *       |

COMMENTS:

APPENDIX D  
TOT TEST PROTOCOL

TOT Test Schedule

| MON am                        | TUES pm                           | THUR pm                          | TUES am                       | FRI am                | WED am                 |
|-------------------------------|-----------------------------------|----------------------------------|-------------------------------|-----------------------|------------------------|
| <u>Class</u><br><u>Starts</u> | <u>Ready for</u><br><u>Mockup</u> | <u>NCI/2B</u><br><u>Complete</u> | <u>ACE</u><br><u>Complete</u> | ***<br><u>N6 Test</u> | ***<br><u>ACE Test</u> |
| AUG 3                         | AUG 11                            | AUG 20                           | AUG 25                        | AUG 21                | AUG 26                 |
| AUG 17                        | AUG 25                            | SEP 3                            | SEP 8                         | SEP 4                 | SEP 9                  |
| AUG 31                        | SEP 8                             | SEP 17                           | SEP 22                        | SEP 18                | SEP 23                 |
| SEP 14                        | SEP 22                            | OCT 1                            | OCT 6                         | OCT 2                 | OCT 7                  |
| SEP 28                        | OCT 6                             | OCT 15                           | OCT 20                        | OCT 16                | OCT 21                 |
| OCT 13                        | OCT 20                            | OCT 29                           | NOV 3                         | OCT 30                | NOV 4                  |
| OCT 26                        | NOV 3                             | NOV 12                           | OCT 17                        | OCT 13                | NOV 18                 |
| NOV 9                         | NOV 17                            | NOV 26                           | DEC 1                         | NOV 25                | DEC 2                  |
| NOV 23                        | DEC 1                             | DEC 10                           | DEC 15                        | DEC 11                | DEC 16                 |
| DEC 7                         | DEC 15                            |                                  |                               |                       |                        |

## Performance Evaluation Briefing

The performance evaluation test you are about to commence was developed for the purpose of measuring the level of proficiency achieved by a student trained on the Air Controller Exerciser (ACE) as compared to a student trained by conventional methods with a pseudo-pilot (N6). The results of this study will have significant impact on the future of both air-intercept (AIC) and anti-submarine (ASAC) air-control training. Therefore, it is important that you do your best on this test, even though it is not a part of the regular training program at this school. The results of this evaluation will be retained in your student record here at the school and will serve as a measure of your progress up to this point in your training.

This exercise consists of 10 intercepts and will take approximately 90 minutes to complete. You will be asked to run nearest-collision intercepts (NCI's) in a fleet-type environment (in response to a SWC's engagement orders), NCI's in a training-type (N6) environment, tanker join-ups using method 2B, and nearest-collision intercept conversions (NCIC's). At the beginning of each run, the instructor in charge of the exercise will inform you as to which kind of intercept you are to execute, so you won't have to keep track of that yourself. During the course of the exercise you may also encounter various events such as loss of NTIS program, jinking bogey, or loss of radio communications. These events may occur without warning and may either come alone or in various combinations.

by now you are aware that you will be asked to demonstrate some skills which you have not yet been taught at this school. Don't let this worry you! This exercise is designed to test your flexibility as well as your level of training; it is expected that you will not be able to perform perfectly on every event in the test exercise. The main thing is to give it your best shot, using all the knowledge and skill you have accumulated so far. You might take this opportunity to consider the things presented in this briefing, and think about what actions you might take if you are faced with any unfamiliar events. As an example, consider the following: you are familiar with the tanker join-up method 2B, in which the tanker is turned in front of the fighter. In this exercise you will be asked to execute an NCIC, in which the fighter is turned in behind the bogey. No other instruction for this type of intercept will be given to you, so you will have to figure out the geometry based on your previous training. The important thing to remember is that your performance on unfamiliar events will not affect your training evaluation at this school. Just do it as best you can.

If events occur during the exercise that you feel should be brought to someone's attention if you were doing this in the real world (i.e. SwC, Evaluator, Track Sup, Watch Officer, AIC Sup, etc.) then pass that information on to the instructor conducting the exercise. He will simulate any watch stations you feel it is necessary to communicate with. Also, any questions or comments you have should be directed to the instructor in charge. Do not ask the exercise evaluator!

One more thing: you may feel that this exercise is conducted in a very formal, inflexible and impersonal manner. The events in this exercise are rigidly controlled in order to ensure that all students who take this test receive exactly the same treatment; this allows us to compare your results with those of a student who may take the test months later. So, if you find this exercise a bit formal, it is because it was designed that way, not because the person in charge decided to conduct it that way.

If you have any questions, please ask them now.

TOT Transition Training, ACE

1. Date/time:

Training to be carried out TUES afternoon following student's completion of the ACE curriculum on the same TUES morning.

2. Duration:

Three (3) regular training periods, approximately 50 minutes each, for a total of three (3) hours minus breaks.

3. Content outline:

Period I: Familiarization, begin Forward Quarters (FQ).

Period II: Continue FQ.

Period III: Continue FQ, add Tanker joinups 2B .

4. Content details:

Period I: (1). Familiarize the student with mock-up operations and differences from the ACE trainer. Areas to be covered are:

- a. A/C pickup and check-in; use of TACAN for identification.
- b. Data entry functions (e.g. entering a CAP station).
- c. Data-access functions (e.g. moving the PPIRO).
- d. AVS differences (e.g. A/C symbol may not come to the proper heading after a turn is made).
- e. Getting Edge of Lope (bearing/range) to points in space.
- f. DRO differences (e.g. ordered head changes with updates).

(2). Run FQ intercepts as time allows.

Period II: Full period of FQ.

Period III: Half period of FQ; half period of 2B Tanker Joinups.

PC & E Briefing Sheet for  
Code 31 TOT Test

1. Schedule: The tests are intended to be run on alternate Wednesdays and Fridays, according to the enclosed schedule. You will be notified in advance of any changes.
2. Personnel: The following Code 31 personnel are in charge of conducting the TOT Test:

OSCS Earnev

OSCS Vincent

CSC MacPherson

CSC Meyer

CS1 Nielsen

3. Pre-test briefing: At the beginning of each test exercise, the Instructor in charge will establish contact with you ( the PC&E pseudo-pilot assigned to that test run). He will identify himself and indicate which version of the test ( ACE or N6 ) is to be run; he will also ask your name, briefly discuss procedures, and answer any questions
4. Procedures: This test exercise is organized into "Events", with each event being a particular type of intercept run under closely controlled conditions. There are two different versions of the test: the "ACE" version and the

"K6" version; each version contains the same events, but they occur in a different order, and the communications are slightly different as well.

The Instructor controls the exercise through communications with the student and through manipulation of the NTDS consoles. The Instructor will communicate with the student over the comm "net"-- this means that you ( the pseudo-pilot ) will hear all conversations between the Instructor and the student, as well as the regular communications between you and the student AIC.

In order to minimize possible confusion, a list of the events in each test exercise, along with a list of the communications which the Instructor will be making, is enclosed in this briefing. You should look over these lists before the exercise begins, and ask the Instructor to clarify any issues during the pre-test briefing. If you will follow along with the check sheet as the exercise progresses you should have a good idea of what's going on, and you should have no problems with the unconventional communications which will be coming over the net.

IMPORTANT NOTE:

If anything occurs during the exercise which you do not understand, or which appears to be wrong, DO NOT attempt to contact the Instructor over the net; ask your supervisor!

Instructor Notes for TOT Test

1. In the event of a student error which, if left uncorrected, would result in serious delay and/or failure to continue the exercise (e.g. student loses his A/C), the instructor shall intervene in a timely manner and provide such assistance as may be required to correct the problem.
2. A particular event/run terminates when one of the following occurs:
  - a. the exercise is successfully completed
  - b. the exercise cannot be successfully completed as ordered
  - c. the exercise is unsuccessfully completed  
( e.g. A/C pass beam of each other on a FQ run)
3. The following symbols in the protocol have special meaning:  

|        |                                                          |
|--------|----------------------------------------------------------|
| o-->   | Defines the timing of an Instructor action.              |
| ___ *  | Describes an action to be taken by the Instructor.       |
| ___ [] | Describes a communication to be given by the Instructor. |
4. It is recommended that the Instructor check each action/event on the protocol with a grease pencil to avoid confusion.

AD-A123 289 TRAINING EVALUATION OF AN AUTOMATED TRAINING SYSTEM FOR 2/2

242

AIR INTERCEPT CON. (U) CANYON RESEARCH GROUP INC

WESTLAKE VILLAGE CA M E MCCAULEY ET AL. DEC 82

UNCLASSIFIED

CRG-TR-82-015 NAVTRAEQUIPC-81-C-0055-1

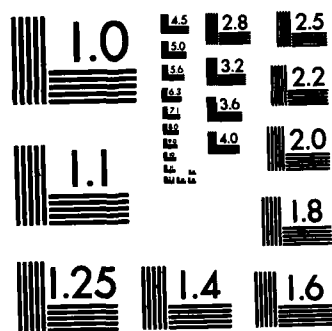
F/G 5/9

NL

END

© 2004 MHC

100



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

TOT Test Communications Check Sheet

ACE Students

Events in the ACE TOT test are:

1. Fleet NCI.
2. Fleet NCI with bogey jink.
3. Training NCI, head-on, 1 stranger.
4. Training NCI, FC, 2 strangers.
5. Training NCI, FC, with MODE 3.
6. Training NCI, FC, with NORDO.
7. Tanker join-up, method 2B.
8. Tanker join-up, method 2B.
9. NCIC.
10. NCIC.

EVENT 0. "Initialization"

\_\_\_ \* Brief PC&E pilot: instructor's name, ACE or N6 test, etc.

\_\_\_ [] "I will act as SWC and, occasionally, as the pilot. In addition, I will simulate all watch stations you might need to talk with. Any communications you feel are necessary can be made to me over the net. All communications from me, including SWC engagement orders, will also be made over the net. Do you have any questions?"

\_\_\_ [] "For the first part of this exercise you will be using one A/C only; use F4 symbols."

\_\_\_ [] SWC: " A/C # \_\_\_ is airborne for your control.  
" Establish contact with your aircraft."

\_\_\_ [] "Offset your scope with CAP STA at the center; set your range scale to 32 miles and KEEP IT THERE. YOU ARE TO MAINTAIN 32 mile RANGE UNTIL YOU ARE INSTRUCTED TO DO OTHERWISE!"

Note: Monitor this closely.

\*\*\* END EVENT 0. \*\*\*

---

EVENT 1. " FLEET NCI #1 "

\_\_\_ [] "Vector your A/C to the CAP STA and prepare to run a  
Fleet-type NCI."

\_\_\_ [] " SWC Engage Order "

\_\_\_ [] Call " Contact, ... "

\_\_\_ [] Call " Judy, Fox 1, Breakaway "

\*\*\* END EVENT 1. \*\*\*

---

EVENT 2. " FLEET NCI #2 "

\_\_\_ [] " SWC Engage Order "

\_\_\_ [] Call " Contact, ... "

\_\_\_ [] Call " Judy, Fox 1, Breakaway "

\*\*\* END EVENT 2. \*\*\*

---

EVENT 3. "TRNG NCI, HEAD-ON, 1 STRANGER "

\_\_\_ [] "Vector your A/C to Ref Point "X".

"You may now use the 64 mile scale.

"Pick up A/C # \_\_\_\_\_ at MIRAMAR and vector it to  
Ref Point "X" also.

" You are now controlling 2 A/C."

\_\_\_ [] "Run a HEAD-ON TRNG-TYPE NCI.

" All exercises from now on are to be conducted in the  
AIC area, using the 32 mile scale.

" Set your scope now."

\*\*\* END EVENT 3. \*\*\*

---

EVENT 4. " TRNG NCI, FQ, 2 STRANGERS "

\_\_\_ [] " Run a FQ TRNG-TYPE NCI."

\*\*\* END EVENT 4. \*\*\*

---

the third orbit,

\_\_\_ [] To PC&E (over the net) :

" End of lost-comms/NORCO procedure; A/C # \_\_\_\_\_  
can now accept radio transmissions."

\_\_\_ [] To Student: " You may now resume control of A/C # \_\_\_\_\_.

" Leave that A/C in orbit and vector your other  
A/C to the CAP STA I pointed out earlier;  
" You are now controlling ONE A/C only."

\*\*\* END EVENT 4. \*\*\*

---

EVENT 5. " FLEET NCI #1 "

\_\_\_ [] "Offset your scope with CAP STA at the center; set your  
range scale to 32 miles and KEEP IT THERE. YOU ARE TO  
MAINTAIN 32 mile RANGE UNTIL YOU ARE INSTRUCTED TO DO  
OTHERWISE!"

\_\_\_ [] " Prepare to run a FLT-TYPE NCI."

\_\_\_ [] " SWC Engage Order "

\_\_\_ [] Call " Contact, ... "

\_\_\_ [] Call " Judy, Fox 1, Breakaway "

EVENT 5. " TRNG NCI, MODE 3 "

\_\_\_ [] " Run another TRNG-TYPE NCI. "

\*\*\* END EVENT 5. \*\*\*

---

EVENT 6. " TRNG NCI, NORDO "

\_\_\_ [] " Run another TRNG-TYPE NCI. "

o--> At 15 miles separation on Breakaway:

\_\_\_ \* Call PC&E supervisor:

" SOUTHBOUND A/C (# \_\_\_\_\_) has lost communications.

" Anchor Port and ignore all transmissions until I give  
the end-of-NORDO order over the net."

o--> When either a) Student recognizes NORDO situation and turns  
other A/C to proper heading,  
or b) Student has failed to recognize the NORDO by  
the third orbit,

\_\_\_ [] To PC&E (over the net) :

" End of lost-comms/NORDO procedure; A/C # \_\_\_\_\_  
can now accept radio transmissions.

\_\_\_ [] To Student: " You may now resume control of A/C # \_\_\_\_\_."

\*\*\* END EVENT 6. \*\*\*

---

EVENT 7. " TANKER JOINUP, 2B, #1 "

o--> When A/C is within 5 miles of CAP STA:

\_\_\_ [] "Vector your A/C back to the NORTH-EAST quarter of  
the AIC area.

"You may now use the 64 mile scale."

o--> When A/C is at the border of the AIC area:

\_\_\_ [] " The remainder of this exercise will be conducted in  
the AIC area.

" Set your scope to 32 mile scale and leave it there."

" Resume control of A/C # \_\_\_\_\_ (the one in orbit).

\_\_\_ [] SWC: " (Low side) A/C # \_\_\_\_\_ has requested a  
rendezvous to take on fuel.

" Take A/C # \_\_\_\_\_ out of orbit and  
run a tanker joinup, or rendezvous, using  
method 2B.

" (High side) A/C # \_\_\_\_\_ is the MAC. "

o--> At completion/termination of the run:

\*\*\* END EVENT 7. \*\*\*

---

EVENT 8. " TANKER JOINUP, 2B, #2 "

\_\_\_ [ ] " Run another method 2B joinup/rendezvous. "

o--> At completion/termination of run :

\*\*\* END EVENT 8. \*\*\*

---

EVENT 9. " NCIC #1 "

\_\_\_ [] " Now I want you to run an NCIC--Nearest Collision  
Intercept Conversion.

" Remember: an NCIC is like a method 2B joinup, except  
that you are to turn the fighter in behind the ~~tanker~~ <sup>Bogey</sup>.

" For the NCIC, (Low side) A/C # \_\_\_\_\_ is the fighter. "

o--> At completion/termination of the run:

\*\*\* END EVENT 9. \*\*\* \_\_\_\_\_

EVENT 10. " NCIC #2 "

\_\_\_ [] " OK, now run another NCIC. "

o--> At completion/termination of run:

\*\*\* END EVENT 10. \*\*\* \_\_\_\_\_

\_\_\_ [] " The exercise is now completed. Place both A/C in port orbit and  
in port orbit and take a break. "

\_\_\_ [] To PC&E: " The test exercise is finished and we no longer  
require the services of your pilot. Thank you."

\_\_\_\_\_  
\_\_\_\_\_

TOT Test Protocol Check Sheet

ACE Students

Events in the ACE TOT test are:

1. Fleet NCI.
2. Fleet NCI with bogey jink.
3. Training NCI, head-on, 1 stranger.
4. Training NCI, FQ, 2 strangers.
5. Training NCI, FQ, with MODE 3.
6. Training NCI, FQ, with NORDO.
7. Tanker join-up, method 2B.
8. Tanker join-up, method 2B.
9. NCIC.
10. NCIC.

EVENT 0. "Initialization"

\_\_\_ \* Brief PC&E pilot: instructor's name, ACE or N6 test, etc.

\_\_\_ \* Identify following points for student:

1. Miramar A/C area

2. CAP STA = 1T1 CH # 100

3. Ref Point "X" = POINT WHISKEY

\_\_\_ \* Check that both consoles set for:

1. Radio channel \_\_\_\_\_

2. Radar channel \_\_\_\_\_

\_\_\_ [] "I will act as SWC and, occasionally, as the pilot. In addition, I will simulate all watch stations you might need to talk with. Any communications you feel are necessary can be made to me over the net. All communications from me, including SWC engagement orders, will also be made over the net. Do you have any questions?"

\_\_\_ \* Mark Eogey Points 1 and 2 on your scope:

Bogey Point 1: from CH# 100, CH # 100 :

bearing = 310 M

range = 60 mi

Bogey Point 2: from CH# 100,

bearing = 015 M

range = 40 mi

\_\_\_ [] "For the first part of this exercise you will be using

one A/C only; use F4 symbols."

\_\_\_ [] SWC: " A/C # \_\_\_ is airborne for your control.

" Establish contact with your aircraft."

o--> When A/C has checked in and student has symbol built:

\_\_\_ [] "Offset your scope with CAP STA at the center; set your range scale to 32 miles and KEEP IT THERE. YOU ARE TO MAINTAIN 32 mile RANGE UNTIL YOU ARE INSTRUCTED TO DO OTHERWISE!"

Note: Monitor this closely.

\_\_\_ \* Build Eogey 1 symbol at Eogey Point 1.

\*\*\* END EVENT 0. \*\*\*

---

EVENT 1. " FLEET NCI #1 "

\_\_\_ [] "Vector your A/C to the CAP STA and prepare to run a  
Fleet-type NCI."

o--> When student gives his A/C vector for CAP STA

\_\_\_ \* Initiate Eogev 1: Heading = 130  
Speed = 360  
Altitude = 36,000

o--> When bogey appears on student's scope  
(range 30 miles to CAP STA)

\_\_\_ [] " SWC Engage Order "

\_\_\_ \* Build Eogev 2 at Bogey Point 2.

o--> When Eogev 1--CAP = 10 miles:

\_\_\_ [] Call " Contact, ... "

o--> When Eogev 1--CAP = 5 miles:

\_\_\_ [] Call " Judv, Fox 1, Breakaway "

\_\_\_ \* Drop track, Eogev 1.

\*\*\* END EVENT 1. \*\*\*

EVENT 2. " FLEET NCI #2 "

o--> When student has vectored CAP back to CAP STA:

\_\_\_ \* Initiate Bogev 2: Heading = 200  
 Speed = 360  
 Altitude = 36,000

o--> When Bogey 2 appears on student's scope  
 (range 30 miles to CAP STA):

\_\_\_ [ ] " SWC Engage Order "

o--> When Bogey 2--CAP = 20 miles:

\_\_\_ \* BOGEY JINK: Heading change = 200 → 170

o--> When Bogey 2--CAP = 10 miles:

\_\_\_ [ ] Call " Contact, ... "

o--> When Bogev 2--CAP = 5 miles:

\_\_\_ [ ] Call " Judy, Fox 1, Breakaway "

\_\_\_ \* Drop track, Bogev 2.

\*\*\* END EVENT 2. \*\*\*

EVENT 3. "TRNG NCI, HEAD-ON, 1 STRANGER "

o--> When A/C is within 5 miles of CAP STA:

\_\_\_ [] "Vector your A/C to Ref Point "X".

"You may now use the 64 mile scale.

"Pick up A/C # \_\_\_\_\_ at MIRAMAR and vector it to  
Ref Point "X" also.

" You are now controlling 2 A/C."

o--> When lead A/C is within 5 miles of Ref Point "X":

\_\_\_ [] "Run a HEAD-ON TRNG-TYPE NCI.

" All exercises from now on are to be conducted in the  
AIC area, using the 32 mile scale.

" Set your scope now."

o--> When A/C have been turned for separation:

\_\_\_ \* Initiate Stranger 1, using UNKNOWN symbol:

Origin = CH #100 240 - 40

Heading = 120

Speed = 300

Altitude = 18,000

o--> At completion/termination of intercept:

\_\_\_ \* Drop track, Stranger 1.

\*\*\* END EVENT 3. \*\*\*

EVENT 4. " TRNG NCI, FQ, 2 STRANGERS "

\_\_\_ [ ] " Run a FQ TRNG-TYPE NCI."

o--> When student gives breakaway headings for separation:

\_\_\_ \* Initiate Stranger 2 and Stranger 3:

| Stranger 2       | CH # 100 | Stranger 3       |
|------------------|----------|------------------|
| <u>250 M, 40</u> | Origin   | <u>215 M, 20</u> |
| <u>090 T</u>     | Heading  | <u>240</u>       |
| <u>300</u>       | Speed    | <u>300</u>       |
| <u>18,000</u>    | Altitude | <u>18,000</u>    |

o--> At completion/termination of run:

\_\_\_ \* Drop track, strangers 2 & 3.

\*\*\* END EVENT 4. \*\*\*

---

EVENT 5. " TRNG NCI, MODE 3 "

\_\_\_ [ ] " Run another TRNG-TYPE NCI. "

o--> At 10 miles separation on Breakaway:

\_\_\_ \* TURN OFF NTDS SYMBCLOGY

NOTE: INSTRUCTOR TO KEEP SYMBOL WITH VIDEO (i.e. track A/C)  
WHILE MODE 3 IS IN PROGRESS.

o--> At completion/termination of run :

\_\_\_ \* TURN NTDS BACK ON.

\*\*\* END EVENT 5. \*\*\*

---

EVENT 6. " TRNG NCI, NORDO "

\_\_\_ [] " Run another TRNG-TYPE NCI. "

o--> At 15 miles separation on Breakaway:

\_\_\_ \* Call PC&E supervisor:

" SOUTHBOUND A/C (# \_\_\_\_\_) has lost communications.

" Anchor Port and ignore all transmissions until I give  
the end-of-NORDO order over the net."

o--> When either a) Student recognizes NORDO situation and turns  
other A/C to proper heading,  
or b) Student has failed to recognize the NORDO by  
the third orbit,

\_\_\_ [] To PC&E (over the net) :

" End of lost-comms/NORDO procedure; A/C # \_\_\_\_\_  
can now accept radio transmissions.

\_\_\_ [] To Student: " You may now resume control of A/C # \_\_\_\_\_."

\*\*\* END EVENT 6. \*\*\*

---

EVENT 7. " TANKER JOINUP, 2B, #1 "

\_\_\_ [] SwC: " (Low side) A/C # \_\_\_\_\_ has requested a rendezvous to take on fuel.

" Run a tanker joinup, or rendezvous, using method 2B.

" (High side) A/C # \_\_\_\_\_ is the MAC. "

o--> At completion/termination of the run:

\*\*\* END EVENT 7. \*\*\*

---

EVENT 8. " TANKER JOINUP, 2B, #2 "

\_\_\_ [] " Run another method 2B joinup/rendezvous. "

o--> At completion/termination of run :

\*\*\* END EVENT 8. \*\*\*

---

EVENT 9. " NCIC #1 "

\_\_\_ [] " Now I want you to run an NCIC--Nearest Collision Intercept Conversion.

" Remember: an NCIC is like a method 2B joinup, except that you are to turn the fighter in behind the ~~tanker~~ <sup>Bogey</sup>.

" For the NCIC, (Low side) A/C # \_\_\_\_\_ is the fighter. "

o--> At completion/termination of the run:

\*\*\* END EVENT 9. \*\*\* \_\_\_\_\_

EVENT 10. " NCIC #2 "

\_\_\_ [] " CK, now run another NCIC. "

o--> At completion/termination of run:

\*\*\* END EVENT 10. \*\*\* \_\_\_\_\_

\_\_\_ [] " The exercise is now completed. Place both A/C in port orbit and in port orbit and take a break. "

\_\_\_ [] To PC&E: " The test exercise is finished and we no longer require the services of your pilot. Thank you."

\_\_\_\_\_

\_\_\_\_\_

DISTRIBUTION

Commanding Officer  
Naval Training Equipment Center  
Orlando, FL 32813

Defense Technical Information Center  
Cameron Station 12  
Alexandria, VA 22310

(All others receive one copy)

Prof. Alphonse Chapanis, Director  
Communications Research Laboratory  
Suites 302 & 303; 7402 York Road  
Baltimore, MD 21204

Mr. Leon Lerman  
Lockheed Missiles & Space Co.  
85-10, Bldg. 182  
Sunnyvale, CA 94086

Dr. Carol A. Simpson  
Psycho-Linguistic Research Assoc.  
2055 Sterling Avenue  
Menlo Park, CA 94025

Dr. John Welch  
Threshold Technology, Inc.  
1829 Underwood Blvd.  
Delran, NJ 08033

CDR Charles W. Hutchins  
Naval Air Systems Command  
(Air 340F)  
Washington, DC 20361

Dr. Wayne A. Lea  
Speech Communications Research Lab.  
806 W. Adams Blvd.  
Los Angeles, CA 90007

Dr. Robert A. North  
Honeywell SRC  
2600 Ridgway Parkway  
Minneapolis, MN 55413

Mr. Leon R. Harrison  
NASA-Ames Research Center  
Moffett Field, CA 94035

Dr. Bruce Lowerre  
Hewlett-Packard  
1501 Page Mill Road  
Palo Alto, CA 94304

Mr. Robert R. Lynchard  
Eagle Technology, Inc.  
3165 McCrory Place  
Orlando, FL 32803

Dr. Robert C. Williges  
Dept. of Industrial Engineering  
and Operations Research  
Virginia Polytechnic Institute  
Blacksburg, VA 24061

Dr. David S. Pallett  
Institute of Computer Sciences  
and Technology  
National Bureau of Standards  
Washington, DC 20234

Dr. Jesse Orlansky  
Sciences and Technology Division  
Institute for Defense Analyses  
1801 North Beauregard St.  
Alexandria, VA 22311

CDR Joseph Funaro, Code 602  
Naval Air Development Center  
Human Factors Engineering Branch  
Warminster, PA 18974

NAVTRAEQUIPCEN 81-C-0055-1

Douglas Chatfield, Ph.D.  
Behavioral Eval. & Training Systems  
5517 74th St.  
Lubbock, TX 79424

Dr. Donald W. Connolly  
US DOT, FAA  
NAFEC ANA-230  
Atlantic City, NJ 08405

Dr. Thomas P. Moran  
Xerox, Palo Alto Research Center  
3333 Coyote Hill Road  
Palo Alto, CA 94304

Mr. Melvin I. Strieb  
3535 NASA, Rd. 1, Apt. 82  
Seabrook, TX 77586

Dr. William S. Meisel  
Technology Service Corp.  
2950 31st Street  
Santa Monica, CA 90405

Dr. Mark F. Medress  
ITT, Defense Communications Division  
9999 Business Park Avenue  
San Diego, CA 92131

Dr. John Makhoul  
Bolt, Beranek & Newman, Inc.  
50 Moulton Street  
Cambridge, MA 02138

Dr. George R. Doddington  
Texas Instruments, Inc.  
P.O. Box 225936  
Dallas, TX 75243

Mr. J. Michael Nye  
Marketing Consultants, International  
100 W. Washington St., Suite 214  
Hagerstown, MD 21740

Dr. Arnold M. Craft  
US Postal Service, R&D Labs  
11711 Parklawn Dr.  
Rockville, MD 20852

Dr. John Ruth  
General Dynamics Corp.  
P.O. Box 748  
Mail Zone 1352  
Fort Worth, TX 76101

Dr. Marshall Farr  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

Dr. John D. Fort, Jr.  
Code 13  
NPRDC  
San Diego, CA 92152

Dr. James McMichael  
Code 14  
NPRDC  
San Diego, CA 92152

Mr. Robert Smith  
CNO (OP 98)  
Washington, DC 20350

Dr. Clyde Brictson  
Dunlap & Associates, Inc.  
920 Kline St., Suite 203  
La Jolla, CA 92037

Mr. P. J. Andrews  
SEA 61R2  
Naval Sea Systems Command  
Room 880, Crystal Plaza 6  
Washington, DC 20360

NAVTRAEQUIPCEN 81-C-0055-1

Dr. David Lambert (Code 823)  
Naval Ocean System Center  
271 Catalina Blvd  
San Diego, CA 92152

John Sigona  
Dept. of Transportation  
DOT/TSC-533  
Kendal Square  
Cambridge, MA 02142

Mr. John Martins, Jr.  
Project Engineer  
Naval Underwater Systems Center  
New London Laboratory MC 315  
New London, CT 06320

Mr. David Hadden  
US Army Electronics Command  
Advanced Systems Design and Dev. Div.  
Chief, Computer Techs & Dev. Team  
Ft. Monmouth, NJ 07703

Mr. Lockwood Reed  
US Army Avionics R&D Activity  
DAVAA-E  
Ft. Monmouth, NJ 07703

Dr. Donald W. Connolly  
Research Psychologist  
Federal Aviation Administration  
FAA NAFEC ANA-230 Bldg 3  
Atlantic City, NJ 08405

Dr. Bruno Beek  
Rome Air Development Center  
Griffiss Air Force Base  
Rome, NY 13441

LT Jeff Woodard  
RADC/IRAA  
Griffiss AFB  
Rome, NY 13441

Commanding Officer  
Navy Fleet Material Support Office  
P.O. Box 2010  
Attn: Ralph Cleveland, Code 9333  
Mechanicsburg, PA 17055

Mr. Leahmond Tyre  
Fleet Material Support Office  
Code 9333  
Mechanicsburg, PA 17055

CDR P. M. Curran  
Naval Air Development Center  
Technology Development Branch  
Code 602  
Warminster, PA 18974

Mr. Robert Larr  
Naval Air Development Center  
Code 8143  
Warminster, PA 18974

Dr. Christian Skriver  
Naval Air Development Center  
Code 6021  
Warminster, PA 18974

Mr. William E. Gibbons  
Naval Air Development Center  
Warminster, PA 18974

Mr. Frank R. Previti  
Naval Air Development Center  
Mail Code 4043  
Warminster, PA 18974

Dr. Julie A. Hopson  
Naval Air Development Center  
Code 6021  
Warminster, PA 18974

NAVTRAEQUIPCEN 81-C-0055-11

LCDR Steve Harris  
Naval Air Development Center  
Code 6021  
Warminster, PA 18974

Charles R. Lueck, Jr.  
United States Postal Service  
Process Control Systems Test Facility  
9201 Edgeworth Drive  
Washington, DC 20027

Mr. Ben Wallis  
Computer Analyst  
David Taylor Naval Ship R&D Center  
Bethesda, MD 20084

OUSDR&E (R&AT) (E&LS)  
CDR Paul R. Chatelier  
Washington, DC 20301

Chief of Naval Operations  
OP-39T  
Washington, DC 20350

Mr. Hal Murray  
Naval Air Systems Command  
Code 53343B  
Building JP-2, Room 610  
Washington, DC 20360

Naval Air Systems Command  
Code 5313A  
Attn: LT Thomas M. Mitchell  
Washington, DC 20361

Commander  
Naval Air Systems Command  
Air 340F  
Washington, DC 20361

Commander  
Naval Air Systems Command  
AIR 413F  
Washington, DC 20361

Mr. Ernest E. Poor  
Naval Air Systems Command  
Air 413B  
Room 336  
Washington, DC 20361

F. Leuking  
Naval Air Systems Command  
AIR-360A  
JP-1, Room 612  
Washington, DC 20361

CDR Richard S. Gibson  
Bureau of Medicine and Surgery  
Head, Aerospace Psychology Branch  
Code 3C13  
Washington, DC 02372

Dr. Sam Schiflett  
Naval Air Test Center  
SY 721  
Patuxent River, MD 20670

Director, National Security Agency  
9800 Savage Road  
Attn: T. W. Page, R54, FANX II  
Ft. George G. Meade, MD 20755

Dr. John F. Boehm  
Director, National Security Agency  
9800 Savage Road  
Attn: R-542, Boehm  
Ft. George G. Meade, MD 20755

Mr. Harold C. Glass  
US Postal Lab  
11711 Parklawn Drive  
Rockville, MD 20852

Mr. John T. Masterson  
US Postal Lab  
11711 Parklawn Drive  
Rockville, MD 20852

NAVTRAEQUIPCEN 81-C-0055-1

Abraham Tersoff Gen. Manager  
US Postal Service Res. Center  
11711 Parklawn Drive  
Rockville, MD 20852

Dr. Tice De Young  
US Army Engineer Topographic Labor-  
atories Research Institute  
Ft. Belvoir, VA 22060

Defense Adv. Research Projects Agency  
Information Processing Tech. Office  
1400 Wilson Boulevard  
Arlington, VA 22209

Defense Adv. Research Projects Agency  
Cybernetics Technology Office  
1400 Wilson Boulevard  
Arlington, VA 22209

Defense Adv. Research Projects Agency  
Infor. Processing Techniques Office  
1400 Wilson Boulevard  
Arlington, VA 22209

Chief of Naval Research  
Code 458  
800 N. Quincy St.  
Arlington, VA 22217

Office of Naval Research (Code 221)  
Dir., Electromagnetic Technology  
800 N. Quincy St.  
Arlington, VA 22217

Mr. Jerry Malecki  
Office of Naval Research  
Code 455  
800 N. Quincy St.  
Arlington, VA 22217

Mr. Gordon D. Goldstein  
Office of Naval Research  
Code 437  
800 N. Quincy St.  
Arlington, VA 22217

Mr. J. Trimble  
Office of Naval Research  
Code 240  
800 N. Quincy St.  
Arlington, VA 22217

Dr. Henry M. Halff  
Office of Naval Research  
Code 458  
Arlington, VA 22217

Dr. Henry J. Dehaan  
US Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

Charles Hartz  
FLECOMTRACENLANT  
Code 02A  
Virginia Beach, VA 23416

Mr. Joe Dickinson  
US Army Applied Tech. Lab  
Ft. Eustis, VA 23662

CAPT Leslie K. Scofield  
Directorate of Training  
US Army Signal Center  
Ft. Gordon, GA 30905

Dr. James D. Mosko  
Naval Aerospace Medical Res. Lab.  
Acoustical Sciences Division  
Code L348  
Pensacola, FL 32508

MAJ Neal Morgan (LGY)  
Air Force Logistics Mgmt Center  
Bldg. 205  
Gunter AFB, AL 36114

Dr. Michael G. Sanders  
US Army Res. Inst. Field Unit  
P.O. Box 476  
Ft. Rucker, AL 36362

NAVTRAEQUIPCEN 81-C-0055-1

CWO2 Ray Priest  
Naval Air Tech. Training Center  
Code 7411  
NAS Memphis (85)  
Millington, TN 38054

Lt. Col. Robert O'Donnell  
6570 ARML-HEA  
Wright-Patterson AFB  
Dayton, OH 45322

Mr. Eric Werkowitz  
AFFDL/FGR  
Wright Patterson AFB, OH 45433

ASD/AXA  
Attn: N. R. Vivians  
Wright-Patterson AFB, OH 45433

CAPT Barry P. McFarland  
US Air Force  
ASD/ENECH  
Wright-Patterson AFB, OH 45433

Mr. Don F. McKechnie  
Research Psychologist  
AFAMRL/HEF  
Wright-Patterson AFB, OH 45433

CAPT Ronald J. Marine  
US Air Force  
ASD/AER-EX  
Wright-Patterson AFB, OH 45433

Mr. John Courtright  
AFAMRL/HEG  
Wright-Patterson AFB, OH 45433

Mr. Don Monk  
AMRL/HED  
Wright-Patterson AFB, OH 45433

Mr. Thomas J. Moore  
AFAMRL/BBA  
Wright-Patterson AFB, OH 45433

CAPT Vince Mortimer  
AFAMRL/BBM  
Wright-Patterson AFB, OH 45433

Noel P. Schwartz  
AFHRL/ASM  
Advanced Systems Division  
Wright-Patterson AFB, OH 45433

Mr. Timothy Theis  
ASD/RWR  
Building 20, Aero. B  
Wright-Patterson AFB, OH 45433

Chief of Naval Education & Training  
Liaison Office  
Human Resource Laboratory  
Flying Training Division  
Williams AFB, AZ 85224

Robert F. Lawson, CDR, USN (Ret)  
Naval Applications Engineer  
ONR Scientific Department  
1030 E. Green Street  
Pasadena, CA 91106

Dr. Keith Bromley  
Naval Ocean Systems Center  
Code 8111  
San Diego, CA 92152

Mr. Warren Lewis  
Naval Ocean Systems Center  
Human Engineering Branch  
Code 8231  
San Diego, CA 92152

Mr. Harry A. Whitted  
Code 8235  
Naval Ocean Systems Center  
271 Catalina Boulevard  
San Diego, CA 92152

NAVTRAEQUIPCEN 81-C-0055-1

Dr. Robert A. Wisher  
Navy Personnel Research and  
Development Center  
Code P309  
San Diego, CA 92152

Mr. Melvyn C. Moy  
Navy Personnel Res. & Dev. Center  
Information & Decision Processes  
Code 305  
San Diego, CA 92152

John Silva  
Naval Ocean Systems Center  
Code 823  
San Diego, CA 92152

Mr. Gary Poock  
Naval PG School  
Code 55PK  
Monterey, CA 93940

Mr. Clayton R. Coler  
Research Scientist  
NASA Ames Research Center  
Mail Stop 239-2  
Moffett Field, CA 94035

Hallie M. Funkhouser  
Technical Assistant  
NASA, Ames Research Center  
Mail Stop 293-3  
Moffett Field, CA 94035

Kinga M. Perlacki  
NASA, Ames Research Center  
Mail Stop 239-2  
Moffett Field, CA 94035

Dr. Edward Huff  
Chief, Helicopter Human Factors Ofc  
Mail Stop 239-21  
NASA, Ames Research Center  
Moffett Field, CA 94035

Dr. Robert P. Plummer  
Asst. Prof., University of Utah  
NASA, Ames Research Center  
Mail Stop 239-2  
Moffett Field, CA 94035

Mr. Robert H. Wright  
Research Psychologist  
Army Research Inst. Field Unit  
P.O. Box 476  
Ft. Rucker, AL 56362

Commander  
Naval Air Systems Command  
Air 413G  
Washington, DC 20361

Mr. Ray Satterfield  
Photographer  
Naval Air Development Center  
Warminster, PA 18974

National Aviation Facilities  
Experimental Center  
Library  
Atlantic City, NJ 08405

Chief of Naval Operations  
OP-593B  
Washington, DC 20350

Commander  
Naval Air Force  
US Pacific Fleet (Code 316)  
NAS North Island  
San Diego, CA 92135

Commander  
Training Command  
Attn: Educational Advisor  
US Pacific Fleet  
San Diego, CA 92147

NAVTRAEQUIPCEN 81-C-0055-1

Commander  
Naval Weapons Center  
Code 3154  
Attn: Mr. Curtis  
China Lake, CA 93555

Commanding Officer  
Fleet Anti-Submarine Warfare  
Training Center, Pacific  
Attn: Code 001  
San Diego, CA 92147

Commander  
Naval Air Test Center  
CT 176  
Patuxent River, MD 20670

Dr. J. D. Fletcher  
Defense Adv. Research  
Projects Agency (CTO)  
1400 Wilson Boulevard  
Arlington, VA 22209

Commanding Officer  
Naval Air Technical Training Center  
Code 104, Building S-54  
NAS Memphis (85)  
Millington, TN 38054

Mr. Walt Primas  
Chief of Naval Operations  
OP-39T  
Washington, DC 20350

Chief of Naval Operations  
OP-987H  
Attn: Dr. R.G. Smith  
Washington, DC 20350

Commander  
Naval Air Systems Command  
Technical Library  
AIR-950D  
Washington, DC 20361

Commander  
Naval Air Force  
US Pacific Fleet (Code 342)  
NAS North Island  
San Diego, CA 92135

Commander  
Naval Air Development Center  
Attn: Code 6022  
Warminster, PA 18974

CWO2 Robert R. Lynchard  
Naval Air Technical Training Cen  
Code 7411  
NAS Memphis (85)  
Millington, TN 38054

Commander  
Naval Sea Systems Command  
Attn: H. Baker, Code 6122  
Washington, DC 20362

Navy Personnel Research and  
Development Center  
Attn: McDowell  
Library, Code P201L  
San Diego, CA 92152

Chief of Naval Operations  
OP-96  
Washington, DC 20350

Commandant  
US Army Field Artillery School  
ATSF-TD-T  
Mr. Inman  
Ft. Sill, OK 73503

Commandant  
US Army Field Artillery School  
Counterfire Department  
Attn: Eugene C. Rogers  
Ft. Sill, OK 73503

NAVTRAEQUIPCEN 81-C-0055-1

Director  
US Army Human Eng. Laboratory  
Attn: DRXHE-HE (KEESEE)  
Aberdeen Proving Ground, MD 21005

USAHEL/USAAVNC  
Attn: DRXHF-FR (Dr. Hoffmann)  
P.O. Box 476  
Ft. Rucker, AL 36362

ASD/ENESS  
Attn: R.B. Kuhnen  
Wright Patterson AFB, OH 45433

Air Force Human Resources Lab  
AFHRL/LR  
Logistics Research Division  
Wright Patterson AFB, OH 45433

LT Dave Cooper  
AFHRL/OTT  
Wright Patterson AFB, OH 45433

US Air Force Human Resources Lab  
AFHRL-IT (Dr. Rockway)  
Technical Training Division  
Loury AFB, CO 80230

US Air Force Human Resources Lab  
TSZ  
Brooks AFB, TX 78235

ASD/ENETC  
Mr. R.G. Cameron  
Wright Patterson AFB, OH 45433

Headquarters 34 Tactical Airlift  
Training Group/TTDI  
Little Rock AFB, AL 72076

Headquarters  
Air Training Command, XPTI  
Attn: Mr. Goldman  
Randolph AFB, TX 78148

Commanding Officer  
Rome Air Development Center  
Library (TSLD)  
Griffiss AFB, NY 13446

Director  
Air University Library  
Maxwell AFB, AL 36100

Mr. Harold A. Kottmann  
ASD/YWE  
Wright Patterson AFB, OH 45433

AFHRL/OTO  
Attn: Mr. R.E. Coward  
Luke AFB, AZ 85309

CDR Charles Theisen  
Lauren Ridge  
R.D. #2, Box 143-8A  
New Hope, PA 18938

Chief  
ARI Field Unit  
P.O. Box 476  
Ft. Rucker, AL 36362