DOCUMENT RESUME

ED 073 237 VT 018 574

AUTHOR Wheaton, George R.; Mirabella, Angelo

TITLE Effects of Task Index Variations on Training Effectiveness Criteria.

INSTITUTION Naval Training Equipment Center, Orlando, Fla.
REPORT NO NAVTRAEQUIPCEN-71-C-0059-1
PUB DATE Oct 72
NOTE 94p.; Prepared under contract by American Institutes for Research, Silver Spring, Md.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Educational Programs; Feasibility Studies; *Job Training; Performance Criteria; Predictor Variables; Simulators; *Skill Development; *Task Performance; Trainees; Trainers; Training Laboratories; *Training Techniques; *Transfer of Training

ABSTRACT

A feasibility study was undertaken as part of a program to develop quantitative techniques for prescribing the design and use of training systems. As the second step in this program, the present study attempted to: (1) refine quantitative indices employed during the early research, (2) conduct laboratory research on the effects which task index variation have on training criteria, and (3) support the laboratory results with data gathered in the field. In one laboratory study, effects of variations in task indices on skill acquisition of a set-up task were examined, while in a companion effort, preliminary data were collected on relationships between task index variations and performance during transfer of training. In the field study quantitative task index data were related to ratio estimates provided by instructors on four training effectiveness criteria. Significant multiple correlations were obtained between task indices and speed and accuracy of performance during skill acquisition. Predictor patterns changed over time and between criteria. Set-up task speed was predicted early in training, while errors made were predicted later during acquisition. Similar but more provisional relationships were found during transfer of training. The results continue to indicate that quantitative task index data can be predictively related to training criteria. (Author/SB)

Technical Report: NAVTRAEQUIPCEN 71-C-0059-1

EFFECTS OF TASK INDEX VARIATIONS ON TRAINING EFFECTIVENESS CRITERIA

George R. Wheaton
Angelo Mirabella

American Institutes for Research
Silver Spring, Maryland

Contract N61339-71-C-0059
NAVTRAEQUIPCEN Task No. 1752-02

October 1972

DOD Distribution Statement

Approved for public release; distribution unlimited.

NAVAL TRAINING EQUIPMENT CENTER
ORLANDO, FLORIDA 32813

Technical Report: NAVTRAEQUIPCEN 71-C-0059-1

EFFECTS OF TASK INDEX VARIATIONS ON TRAINING EFFECTIVENESS CRITERIA

ABSTRACT

A feasibility study was undertaken as part of a program to develop quantitative techniques for prescribing the design and use of training systems. As the second step in this program, the present study attempted to: (1) refine quantitative indices employed during earlier research; (2) conduct laboratory research on the effects which task index variations have on training criteria; and (3) support the laboratory results with data gathered in the field.

Two laboratory investigations and a field study were conducted. In the first laboratory study, effects of variations in task indices on skill acquisition of a set-up task were examined. In a companion effort, preliminary data were collected on relationships between task index variations and performance during transfer of training. In the field study quantitative task index data, descriptive of a variety of sonar trainers and sonar trainee tasks, were related to ratio estimates provided by instructors on four training effectiveness criteria.

Significant multiple correlations were obtained between task indices and speed and accuracy of performance during skill acquisition. Predictor patterns changed over time and between criteria. Set-up task speed was predicted early in training, while errors made were predicted later during acquisition. Similar but more provisional relationships were found during transfer of training. Speed and, in particular, accuracy of performance during transfer bore consistent relationships to task index values. Support for these general findings was obtained in the field. Significant relationships were established between instructors' judgments of training criteria and trainee subtask index values.

The results continue to indicate that quantitative task index data can be predictively related to training criteria. Further development appears warranted. Future research should extend the laboratory findings especially for transfer of training, and should seek to generalize these results to field settings through the collection of performance data.

FOREWORD

PURPOSE

The objective of this research project is to develop quantitative indices of the characteristics of instructors' and trainees' tasks so that the effectiveness of a given amount and type of training on a given task can be predicted. The results of this research should lead to greater accuracy in establishing the human performance requirements in a training system, greater accuracy in human factors design recommendations, and improved instructor station design.

ACCOMPLISHMENTS

In the first phase of this research project, the feasibility of an initial set of quantitative indices in describing the trainee tasks on three sonar operator training devices was demonstrated.

In addition, the feasibility of using quantitative task characteristic indices to predict performance was tested by describing the characteristics of tracking tasks appearing in the experimental literature and predicting tracking performance. (The AD number for ordering the technical report which describes the first phase from the National Technical Information Service, Department of Commerce, Springfield, Va., 22151, is AD 722423.)

In the second phase of this research project—which this technical report describes—the objective was to determine the relationships between systematic variations in quantitative task characteristic indices and performance measures. This was successfully accomplished by learning and transfer experiments in the laboratory and a field validation exercise.

Strong relationships between performance measures and variations in task indices (representing various configurations of synthetic trainer tasks) were obtained. Further, the transfer experiment resulted in data which suggest the feasibility of predicting transfer effects from quantitative task indices. Finally, the data of the field study validated much of the laboratory data.

PLANS

The next phase of this project will investigate the generality of the findings in this technical report to a different family of training devices.

Gene S. Micheli

GENE S. MICHELI, Ph.D.
Human Factors Laboratory

ACKNOWLEDGMENTS

The authors would like to take this opportunity to acknowledge the contributions of Dr. Gene S. Micheli to this project. In his role as Project Monitor, Dr. Micheli was instrumental in making arrangements for the field portion of the effort, in addition to suggesting directions for the laboratory work. No less significantly, his interest in and general support of the research contributed to the success of the project.

The contributions of the U. S. Navy personnel are too numerous to mention. Without their interest and capable assistance the field portions of the study could not have been conducted. We would like to especially thank the officers and men who acted as judges at the following locations: HS1, Quonset Point; FAETULANT, Norfolk; Fleet Training Center, Norfolk; FBM Submarine Training Center, Charleston; and the Fleet Sonar School, Key West.

Finally, work could not have been completed without the advice and support of other individuals at AIR. We would like to extend special thanks to Mr. Warren McDowell, who with limited materials and great imagination, developed the "trainers" used in the laboratory research. Last, but by no means least, we would like to express our thanks to Ms. Jan Strasel, project secretary, for keeping the project on course.

TABLE OF CONTENTS

Section		Page
I	INTRODUCTION	1
	Background	1
	Research Objectives	2
II	METHOD	3
	Quantification of Sonar Training Devices	3
	Laboratory Validation of Indices	5
	Conceptualization and Development of the Synthetic Trainer	5
	Selection of Trainer Configurations	7
	Experimental Procedure	13
	Transfer of Training Procedure	14
	Field Validation of Indices	14
III	RESULTS	17
	Laboratory Findings	17
	Acquisition	17
	Transfer	29
	Field Findings	32
IV	DISCUSSION	36
	Prediction of Set-Up Task Skill Acquisition	36
	Prediction of Set-Up Task Transfer of Training	37
	Prediction of Judged Training Effectiveness	38
	Conclusions and Implications for Research	39
REFERENCES		41
APPENDIX A	SQS-26CX Set-Up Subtask Operations Flow Chart	42
APPENDIX B	Task Characteristic Index Values for Trainee Subtasks Evaluated in the Field	45
APPENDIX C	Operations Flow Chart for the Complex-All Synthetic Set-Up Task	54
APPENDIX D	Instructions for Magnitude Estimates	63
APPENDIX E	Subtask Standard Flow Charts	66
APPENDIX F	Ratio Estimate Answer Sheet	71
APPENDIX G	Mean Instructor Ratio Estimates for the Four Subtasks	73
APPENDIX H	Intercorrelations of Task Index Values and Adjusted Mean Instructor Ratio Estimates	78

LIST OF TABLES

Table		Page
1	Set-Up Task Functions Identified from Task-Analytic Review.	6
2	Task Characteristic Index Values for Synthetic Trainer Tasks	11
3	Intercorrelations of Task Index Values and Mean Performance Times Across Trial Blocks for the Laboratory Tasks.	25
4	Intercorrelations of Task Index Values and Mean Errors Across Trial Blocks for the Laboratory Tasks.	26
5	Summary of Multiple Regression Analyses of Performance Time and Number of Errors for First, Middle, and Last Block of Acquisition Trials.	28
6	Summary of Multiple Regression Analyses of Instructors' Ratio Estimates: Individual Subtasks and Pooled Subtasks	35

LIST OF ILLUSTRATIONS

Figure		Page
1	Complex-All Trainer	8
2	Simple-All Trainer	9
3	Medium-All Trainer	10
4	Mean Performance Time as a Function of Trial Block During Acquisition Training for Simple-Third and Simple-None Tasks	18
5	Mean Performance Time as a Function of Trial Block During Acquisition Training for Simple-All Tasks	19
6	Mean Performance Time as a Function of Trial Block During Acquisition Training for Medium and Complex Tasks	20
7	Mean Errors as a Function of Trial Block During Acquisition Training for Simple-Third and Simple-None Tasks	21
8	Mean Errors as a Function of Trial Block During Acquisition Training for Simple-All Tasks	22
9	Mean Errors as a Function of Trial Block During Acquisition Training for Medium and Complex Tasks	23
10	Mean Performance Time as a Function of Trial Block During Transfer to Medium-All Task	30
11	Mean Errors as a Function of Trial Block During Transfer to Medium-All Task	31

SECTION I

INTRODUCTION

One of the most difficult and complex problems confronting individuals responsible for training is the design and development of effective training devices. In military settings, where complex simulators and trainers often provide the basis for instruction, the problem is particularly acute. During development of these complex devices, options are nearly always available with respect to the design of trainee and instructor stations. Given such options, however, there is seldom any solid basis for choosing among them in terms of their relative effectiveness. Faced with alternative designs for the trainee's station, one finds it hard to specify with confidence those which will prove most effective in promoting rapid acquisition of skills and/or positive transfer to the operational situation. Similarly, given alternatives in design of the instructor's station, one may have difficulty in identifying those which will enable instructor personnel to function most effectively in carrying out their duties.

To deal with these and a series of allied training problems, it is essential to have data relating selected parameters of alternative designs to aspects of trainee and instructor performance. If consistent changes in these criterion measures could be demonstrated as a function of systematic manipulation of design parameters, then such information could be used to predict the effects which different console layouts, sequences of operation, etc., might have on the trainee's rate of learning or the instructor's level of performance. The ability to make such forecasts would provide sounder bases for a variety of training system design decisions including, for example, appropriate degree of simulation fidelity, trainee to instructor ratios, and part versus whole training. Equally important, accurate forecasts would aid in identifying those design tradeoffs which could be made without compromising training effectiveness.

BACKGROUND

In spite of the promise inherent in this approach, the methodology required for its implementation has been slow in developing. A major obstacle to more rapid progress has been the lack of an adequate means for describing alternative designs. Essentially, a set of indices is desired in terms of which different design configurations might be scaled quantitatively. Until such indices become available, the relationship between alternative design configurations and the different rates of learning or levels of performance associated with them cannot be meaningfully explored.

In response to this problem, the Naval Training Equipment Center (NAVTRAEQUIPCEN) initiated a program of research which was to be executed in a series of phases. The primary objectives of the first phase were to compile and to demonstrate the feasibility of applying a set of quantitative task indices. This effort, which has been described in detail elsewhere (e.g., Wheaton, Mirabella, and Farina, 1971) entailed several activities which included: (1) identifying design features of training devices which conceivably could be quantified; e.g., number of displays and controls and their arrangements; (2) exploring a variety of means for their quantification, relying primarily on indices and techniques previously developed and reported in the literature; and (3) determining the feasibility of using the assembled indices to quantify some actual training devices. To keep the scope of this effort within manageable bounds, concern was limited to features of trainee stations found in various sonar training devices. In spite of this restriction, however, it was assumed that many of the features chosen for quantification would be relevant to other types of trainee stations as well as to instructor stations. Application of the indices to four trainee tasks (i.e., set-up, detection, localization, classification), as represented in a small number of different devices, was attempted. This exercise demonstrated that most, if not all, of the indices could be used to scale quantitatively the extent and manner in which the trainee tasks differed within and across devices.

RESEARCH OBJECTIVES

As part of the larger research program and as a sequel to Phase I efforts, the present study had three objectives. The first objective was to refine the set of quantitative indices employed during the earlier research, adding new descriptors, if possible, while deleting those which proved unsatisfactory. The second objective was to conduct an investigation of the relationship between variations in quantitative indices and corresponding changes, if any, in selected criterion measures. This effort was to be conducted in a laboratory setting in order to exercise control over other variables not of immediate interest to the present study. The third and final objective was to determine whether support for relationships established in the laboratory could be provided by data collected in the field. Such support would increase confidence in the validity of the basic methodology—that of using quantitative task index information to forecast the relative effectiveness of competing designs.

The remainder of this report describes the research performed in pursuit of the three primary objectives. In the next section, Section II, the method of procedure is presented. The presentation starts with a description of how devices were quantified in the field, and proceeds to a discussion of the methods employed in laboratory and field validation studies. The results of these studies are presented in Section III. In Section IV, the final section, the results are discussed in terms of their implications for the prediction of training device effectiveness and for future research.

SECTION II

METHOD

The general approach pursued in the current research stemmed from results of the previous phase. As already indicated, the thrust of Phase I was to demonstrate that alternative design configurations could be scaled quantitatively. It remained to be established, however, that such scaling could be predictively related to learning and proficiency criterion measures. In order to provide such evidence, an approach was adopted consisting of three distinct but interrelated activities. Quantification of devices in the field was continued using a revised set of indices. The data obtained during this exercise were then used in conducting a two-pronged validation study consisting of a laboratory and a field effort.

The dual validation effort was felt necessary because of inherent limitations in either the laboratory or field approach alone. While the laboratory approach would facilitate measurement and experimental control, it would require generalization to actual field conditions. On the other hand, while the field effort would permit direct assessment of the quantitative indices, it presented the familiar problem of obtaining performance data under operational conditions. By pursuing both avenues it was hoped that their respective weaknesses could be offset.

QUANTIFICATION OF SONAR TRAINING DEVICES

Before either validation effort could be initiated, quantitative task index data were required on a sample of actual devices. These data were intended to provide guidelines for the types and ranges of design characteristics to be manipulated in the laboratory. In addition, they were to be employed directly in the anticipated field validation effort as the predictor variables. Accordingly, efforts begun during Phase I to apply the quantitative indices were continued during the present research.

Application of the indices was extended to several devices not examined during the earlier work. Altogether, 13 different trainee stations were quantified including: the 14E10/3 at Quonset Point, Rhode Island; the 14B31B (AQA-1 and ASA-20 stations), 14E14, and X14A2 at Norfolk, Virginia; the 21A39/2 (OA1283, BQR-2C, and BQR-7 stations) at Charleston, South Carolina; and the 14E3, 14A2/C1, SQS-26CX, and 21B55 (OA1283 and BQR-2B stations) at Key West, Florida.

The procedures involved in quantifying these devices have been described at length in an earlier report (e.g., Wheaton, Mirabella, and Farina, 1971). Briefly, instructor personnel familiar with the operation of each device were asked to perform and describe in detail all of the primary and contingency actions comprising each of four trainee subtasks. These subtasks, found in most, but not all of the devices, included set-up, search or detection, localization, and classification. The task-descriptive data obtained for each subtask were then converted into flow-chart form for more convenient processing. An example of one of the types of flow charts generated is shown in Appendix A for the SQS-26CX set-up subtask.

Upon conversion of the task descriptive data to flow-chart form, they were analyzed in terms of a variety of quantitative indices. A reduced set of indices from the total compiled during Phase I was employed in the present research.

Exclusion of indices from this final set occurred for one of four reasons. Some, most notably a set of task characteristic rating scales, were excluded because: (1) they were often difficult to apply objectively, requiring a consensus among several analysts; and (2) they referred in many instances to characteristics which, although varying across very different types of devices, did not appear to reflect readily manipulable design features (e.g., the work load dimension). Still other indices were excluded either because they generated little variation for the present types of devices or because they had been found from past work to be correlated highly with other descriptors.

The set of descriptors finally adopted included 17 indices. A brief definition of each is given below, together with references when appropriate. Included were the following:

a. MAIN - defined as the number of responses comprising the main or dominant procedural sequence in an operations flow chart. In the flow chart, shown in Appendix A, there are 24 of these control and display actions (i.e., those connected by solid lines).

b. CNTG - defined as the number of responses comprising the auxiliary or contingency procedural sequences. The flow chart, shown in Appendix A, contains 24 responses of this type (i.e., those connected by dotted lines).

c. TA - defined as the total number of responses (actions) comprising the procedural sequence in an operations flow chart. It represents the sum of MAIN and CNTG.

d. CONT - defined as the total number of different controls manipulated during performance of a subtask.

e. DISP - defined as the total number of different displays referenced during performance of a subtask.

f. E - defined as the total number of different equipment elements interacted with, this index is given by the sum of CONT and DISP.

g. LV - the link value reflecting the relative strength of the sequence of use among the various controls and displays. As used here it is the sum of the products of the number of times a link is used, and the percentage of use of the link (Fowler, Williams, Fowler, & Young, 1968).

h. AA% - an index reflecting the percentage of alternative actions present in an operation. A score of, "...0% means that the highest number of alternative links are used, each with an equal frequency of use, and 100% score means there is only one link out of and into each control, with the same frequency used for all links." (Fowler et al., 1968).

i. F% - another index (Fowler et al., 1968) describing the extent to which all controls and displays are used an equal number of times (0%) or a theoretically defined optimum number of times (100%).

j. DEI - a measure of the effectiveness with which information flows from displays via the operator to corresponding controls. The index yields a dimensionless number representing a figure of merit for the total configuration of displays and controls (Siegel, Miehle, & Federman, 1962).

k - m. D%, C%, E% - defined respectively as the number of display, control, or combined equipment elements which the operator actually employs relative to the total number of such elements which are available for use.

n - q. CRPS, FBR, INFO, INST - refer to the frequency with which the operator makes various types of responses during performance of the task. Included are responses involving manipulation of controls (CRPS), securing of feedback (FBR), acquisition of information (INFO), as well as those primarily initiated by the instructor (INST).

The values actually obtained on each of these 17 indices for the 13 trainee stations previously listed are presented in Appendix B. Four separate tables are presented corresponding to each of the basic trainee subtasks. The index data for all four subtasks were used as predictors in the field validation effort. The index data obtained for the various set-up subtasks provided guidelines for the laboratory validation effort.

LABORATORY VALIDATION OF INDICES

The general approach to laboratory validation was to develop a modularized, synthetic sonar trainer, capable of being readily configured into a large number of sonar "trainers," varying in design characteristics, but with a common set of functions. The trainer was designed to evaluate set-up behavior alone. Other subtasks; i.e., detection, tracking, classification, were excluded because the instrumentation necessary was considered beyond the scope of available time and resources.

CONCEPTUALIZATION AND DEVELOPMENT OF THE SYNTHETIC TRAINER. Design of the trainer was preceded by an extensive examination and analysis of the task data collected during this and the previous phase of our research. Working from both the original task-analytic data and derivative flow charts, essential set-up functions were identified on a trainer-by-trainer basis. A relatively common set of functions; i.e., cutting across all the trainers studied, was generated (table 1). These functions are basic activities performed by the sonar trainee operator during set-up and are relatively common to all the sonar devices which have been explored in this program. Approximately 23 set-up functions were identified. Some of these were later combined to yield a reduced set of 19 functions. For each of these 19 functions, an equipment module was eventually designed.

On a second pass through the devices, displays and controls needed for each function were identified. These displays and controls were then collapsed across devices, and duplicate units eliminated to arrive at a final, non-redundant set for each function. These sets of equipment elements were the basis for designing a module for each of the 19 functions.

TABLE 1. SET-UP TASK FUNCTIONS
IDENTIFIED FROM TASK-ANALYTIC REVIEW

1. Energize the console
2. Check gyro status
3. Activate calibration mode
4. Select transducer operation modes, e.g., active/passive, ATF/MTB
5. Select range scale and adjust range cursor

Adjust PPI intensity/focus for:

6. Overall scope
7. Sweep
8. Cursor
9. Adjust audio for comfort level
10. Adjust console illumination for comfort level

Insert sonar parameters

Geo-references:

11. True/relative
12. Speed
13. Course
14. Ship centered display/target centered display

Other parameters:

15. Sound velocity
16. Pulse length/dwell time
17. Frequency
18. Sum/difference
19. Depression elevation angle

Calibrate the PPI re:

20. Range cursor
21. Bearing cursor
22. Sweep
23. Check signal meters for operation

Each module contained displays and controls which duplicated actual hardware found in the sonar devices, or which represented the essential functions of actual hardware. Representative displays and controls were used where the complexity of actual hardware was beyond the scope of the current effort. For example, simple meter movements arranged as voltmeters across a variable voltage source were used in place of the PPI. Manipulating this voltage source to effect a change in meter reading is somewhat analogous to manipulating a handwheel to effect changes in the position of a PPI range or bearing cursor. It was felt that the essential decision-making and perceptual-motor activity could be abstracted via this kind of substitution of hardware, even though the substituted version might appear rather different from the actual hardware. Where actual hardware consisted of such items as toggle switches, function switches, meters, and jeweled signal lights, actual hardware was used.

For most of the modules, a "simple" and a "complex" form was constructed to represent simple versus more complex hardware for discharging essentially the same function. Altogether, a total of 30 different modules was available for combination into a variety of trainer configurations.

SELECTION OF TRAINER CONFIGURATIONS. For purposes of the present research, an attempt was made to compile a set of configurations which would vary as much as possible along the 17 design indices selected for study. Toward this end two anchor configurations were initially selected representing extreme designs. There was a "complex" trainer consisting of all complex panels and a "simple" trainer consisting of all the simple panels which were available (i.e., simple panels were used at all those positions for which simple panels had been constructed). The complex and simple configurations are shown in figures 1 and 2. Given the two extreme configurations, an intermediate configuration was then generated by randomly selecting either a complex or a simple module for each function on the trainer console. This configuration, known as the medium-all trainer, is shown in figure 3.

In addition to these three primary trainers, nine additional trainers were selected to yield a range of design parameter values. These configurations essentially represented variations in the simple trainer or the medium trainer; i.e., the simple trainer embedded in the complex, medium trainer with feedback lights removed, simple trainer with additional contingency responses included in the training regimen. These manipulations were aimed at reducing correlations among the design parameters, in particular the correlation between number of displays or controls and other design characteristics.

For each trainer, a specific set of procedures or sequence of responses was developed. These served to define "trainee" tasks analogous to the trainee set-up subtasks associated with actual sonar training devices. To the extent that equipment elements were present on a panel, but not involved in task performance, the task was said to be embedded. If a reduced number of feedback lights was used, the task was labeled according to those indicator groups which were used (i.e., none, every third, all). The 12 tasks which were employed are listed in table 2, together with their values on the same set of task indices previously applied in the field.

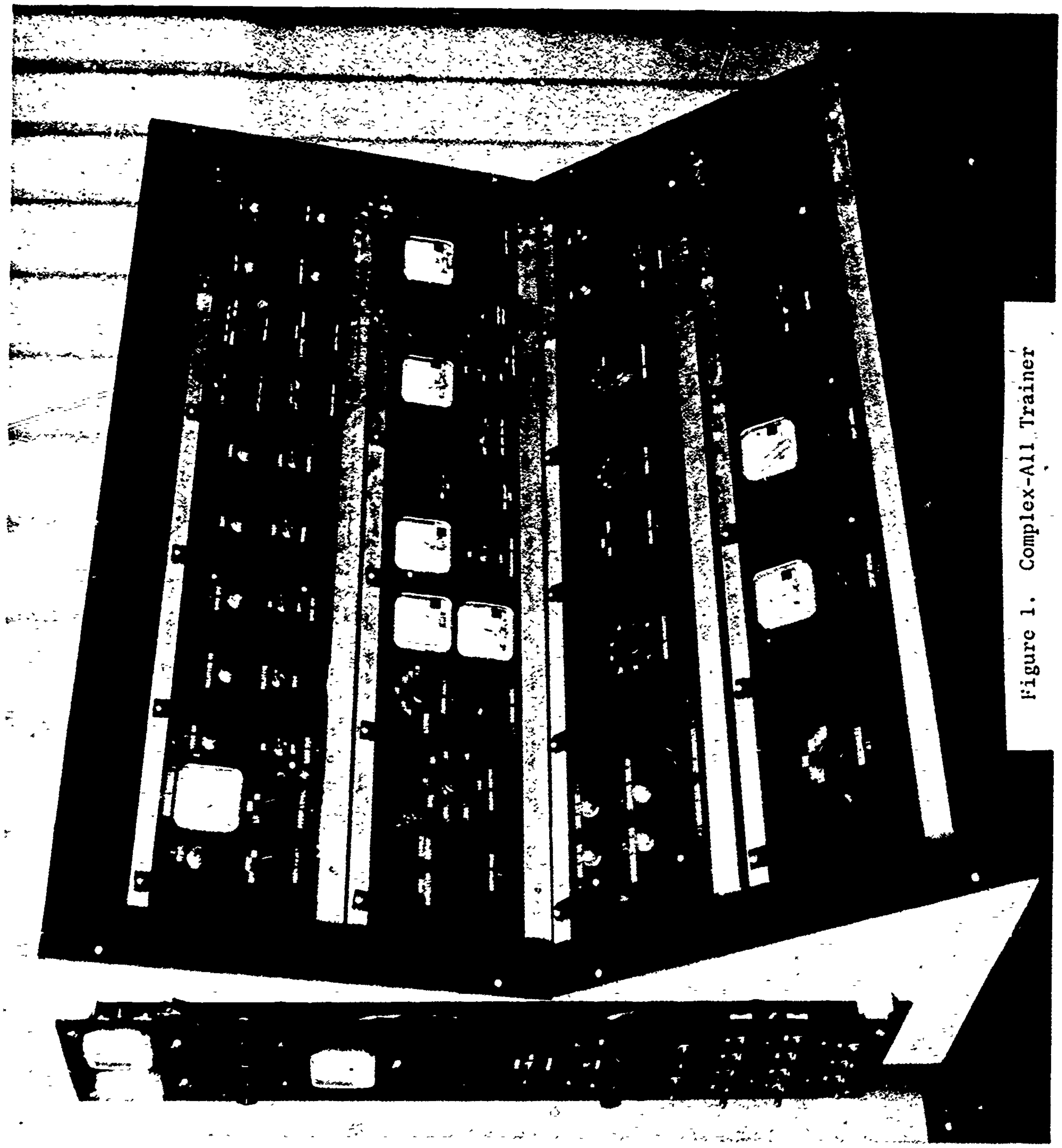


Figure 1. Complex-All Trainer

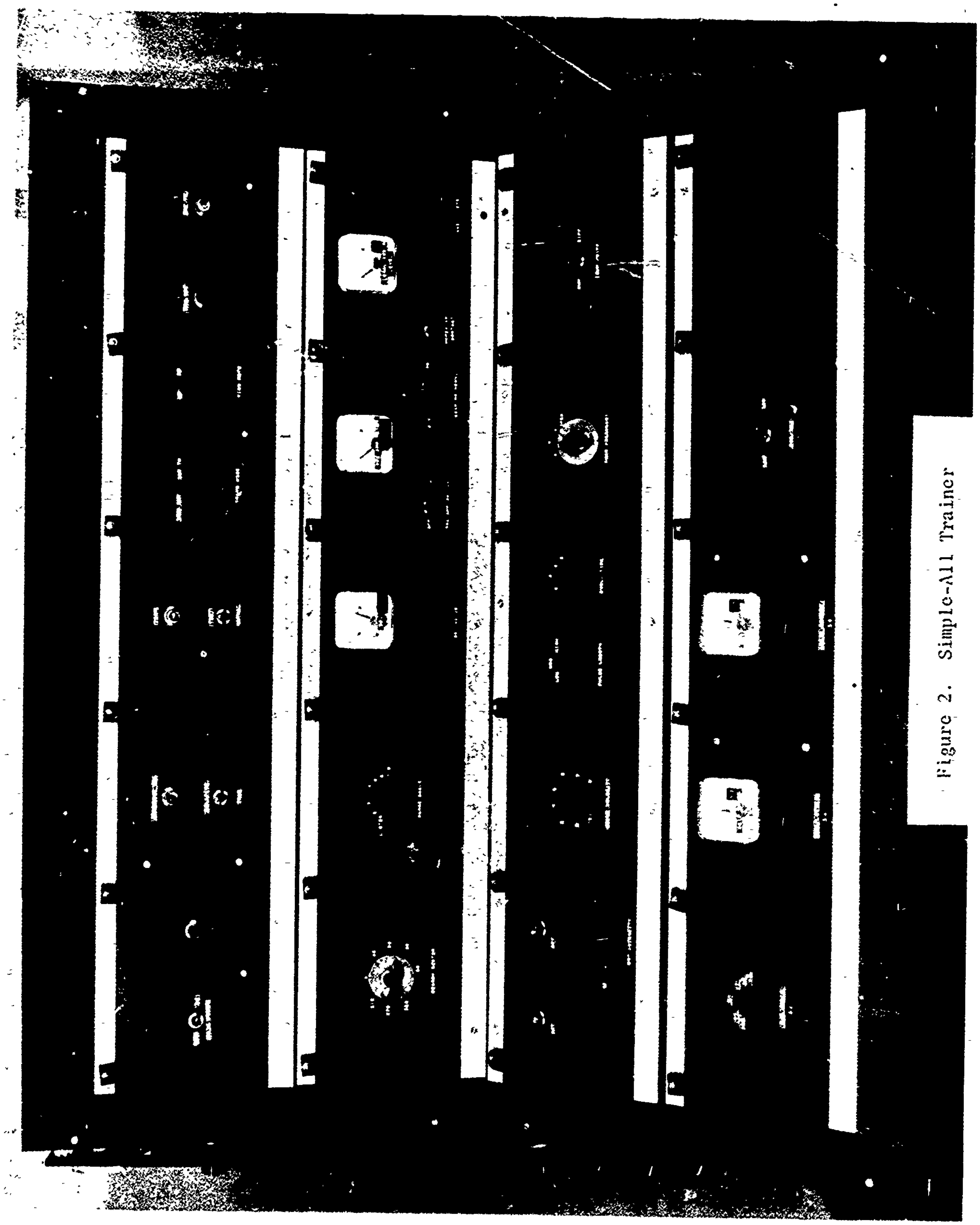


Figure 2. Simple-All Trainer

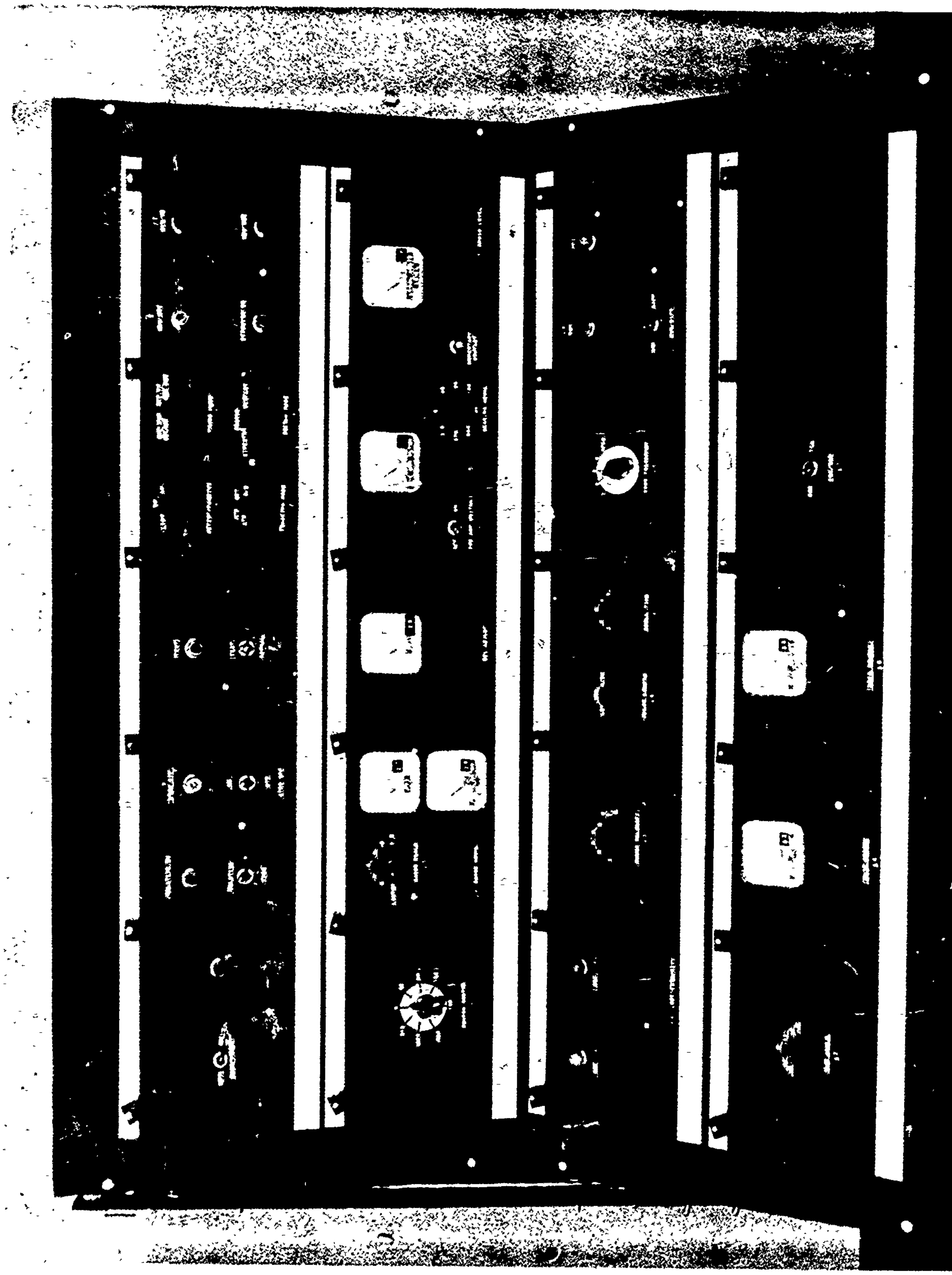


Figure 3. Medium-All Trainer

TABLE 2. TASK CHARACTERISTIC INDEX VALUES FOR SYNTHETIC TRAINER TASKS

Task	Task Indices								
	MAIN	CNTG	TA	CONT	DISP	E	LV	AA%	F%
1. Complex Task All Indicators	69	46	115	34	24	58	7591.2	65	66
2. Medium Task All Indicators	50	34	84	27	19	46	5788.8	68	73
3. Medium Task Third Indicator	47	23	70	27	12	39	4922.1	70	71
4. Medium Task + 2 Third Indicator Embedded in Complex	47	25	72	27	12	39	4922.1	68	68
5. Simple Task All Indicators	43	20	63	23	13	36	4516.7	71	83
6. Simple Task + 6 Third Indicator	41	20	61	23	9	32	4125.0	67	78
7. Simple Task + 6 Third Indicator Embedded in Complex	41	20	61	23	9	32	4125.0	67	78
8. Medium Task + 2 None	46	23	69	27	11	38	4722.1	68	70
9. Simple Task All Indicators Embedded in Complex	43	20	63	23	13	36	4516.7	71	83
10. Simple Task All Indicators Embedded in Medium	43	20	63	23	13	36	4516.7	71	83
11. Simple Task None Embedded in Medium	40	12	52	23	8	31	3728.7	71	89
12. Simple Task None	40	12	52	23	8	31	3728.7	71	89

TABLE 2. TASK CHARACTERISTIC INDEX VALUES FOR SYNTHETIC TRAINER TASKS (Cont)

Task	Task Indices							
	DEL $x10^{-4}$	D%	C%	E%	CRPS	FBR	INFO	INST
1. Complex Task All Indicators	5.3	100	100	100	60	26	29	7
2. Medium Task All Indicators	9.7	100	100	100	44	19	21	6
3. Medium Task Third Indicator	10.9	100	100	100	44	9	17	6
4. Medium Task + 2 Third Indicator Embedded in Complex	8.0	50	79	67	46	9	17	6
5. Simple Task All Indicators	16.3	100	100	100	35	12	16	6
6. Simple Task + 6 Third Indicator	10.8	100	100	100	39	8	14	6
7. Simple Task + 6 Third Indicator Embedded in Complex	8.2	37	68	55	39	8	14	6
8. Medium Task + 2 None	9.9	100	100	100	46	7	16	6
9. Simple Task All Indicators Embedded in Complex	12.7	54	68	62	35	12	16	6
10. Simple Task All Indicators Embedded in Medium	14.3	68	85	78	35	12	16	6
11. Simple Task None Embedded in Medium	17.3	42	68	67	34	5	13	6
12. Simple Task None	21.3	100	100	100	34	5	13	6

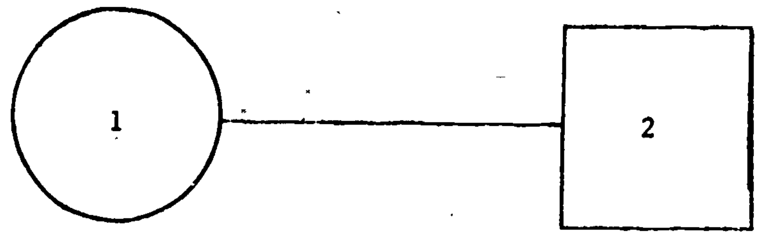
EXPERIMENTAL PROCEDURE. Following development of the synthetic trainer and selection of the specific tasks to be studied, the testing portion of the laboratory effort was initiated. Subjects who were to serve as trainees during this portion of the study were recruited from universities in the metropolitan Washington, D. C. area. The subjects were males who, on the average, were 22 years old, 71 inches tall, and weighed 159 pounds. Subjects were randomly assigned in groups of five to each of the 12 experimental tasks. The 60 subjects employed in this manner were paid for their services.

Upon arrival at the American Institutes for Research (AIR), each subject was ushered into the laboratory and seated before the experimental console, configured according to the task group to which the subject had been assigned. The following standard instructions were then read:

> The experiment you are taking part in today is part of a research program to study how well and how quickly people learn to operate equipment, which is designed in a variety of different ways. Your task will be to learn to operate the equipment which is before you. I will go through the operation of the device step-by-step with you. I will do this twice, and then I will ask you to repeat the operations from memory a number of times. I will correct errors or omissions which you make, but please do your best to recall the operations. Accuracy and speed are both important for obtaining valid research data. Following each run-through, you will be asked to leave the room so that the equipment can be reset. You may wait in the lounge while this is being done. Are there any questions?

Following presentation of these instructions, the subject was given detailed information on how the task was to be performed. Using a specially prepared flow chart, similar to that presented in Appendix C for the complex-all task, the subject was instructed step-by-step in the procedure to be learned. An important aspect of these instructions concerned the standardized reporting language which the subject was to use when describing his task responses. For example, instructions for Panel 1 of the complex trainer included the following:

INSTRUCTION: Set main power Switch #1 to Standby Position — Check main power Indicator #2 for Green Indication



VERBAL RESPONSE: "1 to Standby" — "2 is Green"

Standardized responses were used to minimize the variability inherent in the time required for verbalization of behavior. The complete set-up procedure was described twice in this manner after which any final questions were answered.

Following this orientation session, 15 experimental trials were administered. Preliminary pilot work indicated that performance reached asymptote within this number of trials for a prototype trainer. Prior to each trial the subject left the testing area and the experimenter set all controls in randomized positions according to a predetermined scenario. Programming of the various trainer configurations was of the simplest kind. The experimenter preset switches and displays either on the trainer itself or on a peripheral control panel. Again, the present scope of effort limited the sophistication which could be applied to instrumentation.

Upon being recalled for each trial, the subject went through an entire set-up procedure, verbalizing each response which he made. Correct verbal responses were precoded on a trial-by-trial basis (for the randomized initial control settings) on the experimenter's response sheet. Therefore, measurement of performance consisted of simply checking off each response as it was emitted by the subject. Erroneous or omitted responses were so coded. Time to complete each run-through was measured with a stop watch. However, the watch was stopped while subject errors were being recorded and corrected. Thus, time, errors of omission, and errors of comission provided the dependent measures.

TRANSFER OF TRAINING PROCEDURE. The primary laboratory validation focused upon acquisition of set-up skills. However, as an adjunct to this effort, a pilot transfer study was also undertaken. In this effort additional training was provided for five of the 12 groups involved in the main study (groups 2, and 9 through 12 in table 2). These particular groups were chosen because they provided some interesting contrasts; i.e., effect of panel clutter or embedding on transfer (ratio of used to unused displays and controls). Following the regular acquisition trials, subjects in these groups were permitted to rest for one-half hour. They were then brought back to the laboratory and retrained on the medium-all task. This training regimen was identical to the acquisition regimen; i.e., two complete run-throughs. However, only 10 training trials were run rather than 15. One of the groups originally trained on "medium all" was not given any retraining, but merely tested for retention. Ten trials were also employed for this group.

FIELD VALIDATION OF INDICES

The second prong of the dual validation attempt involved a study of the effectiveness of the 13 sonar training devices which had been previously task analyzed. Ideally, such a study should involve carefully controlled measurement of actual training experiences by novice enlistees. Such a procedure, however, would require considerable interference with on-going training activity and normally is not feasible. Therefore, field validation was pursued via structured interviews with experienced sonar instructors. These instructors were asked to rate the tasks trained on their devices against a set of "synthesized" comparison tasks.

The data collection was undertaken at sites previously employed for training device analysis. These included the Fleet Sonar School at Key West, Florida, the Fleet Ballistic Missile Submarine Training Center at Charleston, South Carolina, and the Fleet Training Center and Fleet Airborne Training Unit at Norfolk, Virginia, and the Quonset Point Naval Air Station in Rhode Island.

At each sonar training device installation visited, a group of four or five instructors was convened who were qualified on the device under examination. These instructors had the following average experience profile:

Experience category	Mean number of years
Total Navy	10.9
Sonarman at sea	5.9
Sonar instructor	1.9
Experience on device being rated	1.3

Instructors were assembled in groups in a classroom setting and were given a series of instructions. These introduced the background of the project, stated the purpose of the current visit, and explained the method which was to be employed in making judgments about the particular training device under examination. This method required the instructors to compare the set-up, detection, localization, and classification subtasks performed on their device against a similar set of subtasks associated with a hypothetical sonar trainer. This same set of hypothetical subtasks was used as a common frame of reference for all groups of instructors. The hypothetical trainer actually represented a disguised amalgam of several of the devices being studied.

Following this general orientation, instructors were given detailed instructions about four specific ratio judgments which they were to make. These instructions, included in Appendix D, concerned how estimates were to be made about: (1) training time; (2) proficiency level; (3) degree of transfer of training; and (4) level of task difficulty.

Upon completion of the instructions and, after answering any questions, instructors were provided with flow charts designed to facilitate their judgments. Two types of flow charts were used. One set described the subtasks to be evaluated and were similar, for instance, to the set-up flow charts included in Appendix A. The other set consisted of the standard flow charts which were to be used as the frame of reference. These flow charts appeared in Appendix E.

One subtask was dealt with at a time, starting with set-up and finishing with classification. For a given subtask, the standardized flow chart was distributed first, and reviewed step-by-step with the instructors. Next, the flow chart, representing the same subtask in the device to be evaluated, was distributed and reviewed in similar fashion. Based upon a comparison of their own subtask with the standard, instructors were then asked to provide ratio estimates on each of the four criterion dimensions, using the response blank shown in Appendix F.

When evaluations of all four subtasks were completed, a group discussion was held to try to arrive at consensus judgments. No attempt was made to force consensus, but instructors were encouraged to discuss any disagreements among their ratings. Misunderstandings about evaluation procedures were also taken up at this time. On the basis of the group discussion, each instructor provided a final judgment. That judgment was accepted, no matter how disparate it was from any other judgments.

Following evaluation of all of the subtasks for the actual device, instructors were finally asked to make a last series of judgments concerning the relative difficulty of the standard subtasks. This time they were to evaluate the standard detection, localization, and classification subtasks, using the standard set-up subtask as a basis for comparison. Such judgments were designed to provide a means for expressing the ratio estimates in terms of a common metric, thus permitting direct comparisons across subtasks.

SECTION III

RESULTS

Three distinct sets of results are presented in this section. The first concerns the acquisition data obtained on the synthetic set-up trainer. The second set, also based on laboratory research, stems from the pilot transfer of training study. Final portions of the results section deal with findings from the field validation exercise.

In the major sections which follow, the same general format is used. The basic layout of the data is given first, followed by a brief description of general findings. The more specific analyses are then presented. These are primarily in correlational form, attempting to describe the relationship between task index variables and a variety of criterion measures.

LABORATORY FINDINGS

Results of the acquisition and transfer portions of the laboratory study are presented in figures 4-11 and tables 3-5. They describe variations in performance speed and accuracy as a function of synthetic trainer task configurations.

ACQUISITION. The basic performance data for acquisition training are shown in figures 4-9. In each case either mean performance time (figures 4-6) or mean number of errors (figures 7-9) is plotted as a function of trial block with task configuration as the parameter. The 15 acquisition trials originally administered were collapsed into seven blocks in order to improve stability of the data. Thus, each point in these figures represents an average value for ten scores (five subjects per trial over two trials). An exception is the final block (T_{13-15}) which spans three trials and represents, therefore, 15 scores.

Figures 4-6 and 7-9 have essentially been broken out from two larger time and error composites in order to improve clarity of presentation. The simple-third and simple-none configurations provide one grouping (figures 4 and 7). The simple-all configurations provide a second grouping (figures 5 and 8), and the medium and complex configurations yield a third grouping (figures 6 and 9). These pairs of figures describe mean performance time and mean number of errors respectively.

Viewed in their entirety, all six figures reveal substantial variance in performance across task configurations. This variance is shown most clearly for the mean performance times of the simple-none, simple-all, and simple-third groups (figures 4 and 5). The medium groups, while contributing to overall variance, are fairly homogeneous, especially when compared to the complex-all configuration (figure 6). Variation across tasks in terms of error scores, though somewhat less dramatic, is still marked (figures 7-9). This is again particularly true for the simple-third + 6 and simple-none tasks (figure 7). Demonstrable variance in both the time and error criterion measures was, of course, a prerequisite for the anticipated correlational analyses.

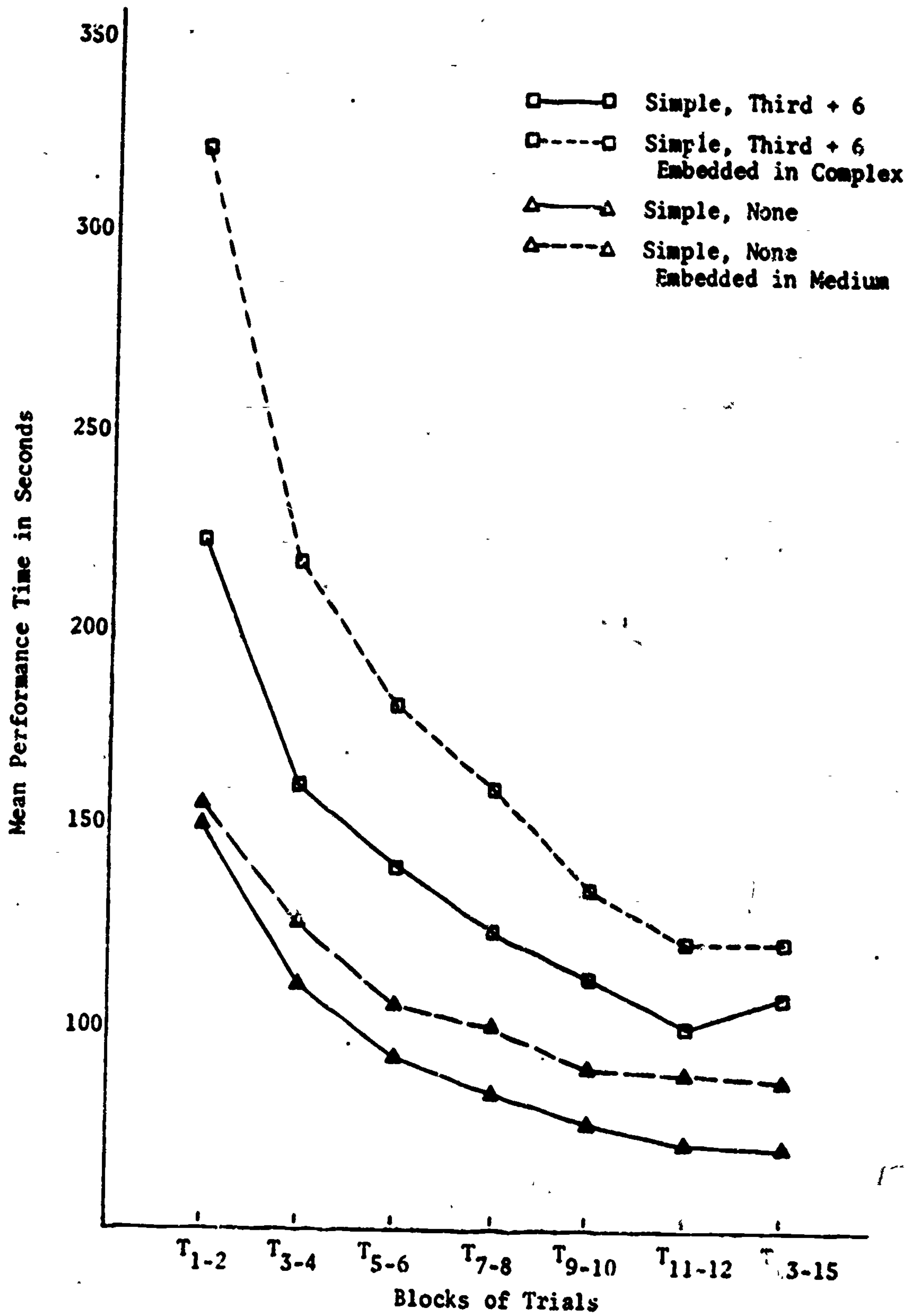


Figure 4. Mean performance time as a function of trial block during acquisition training for simple-third and simple-none tasks

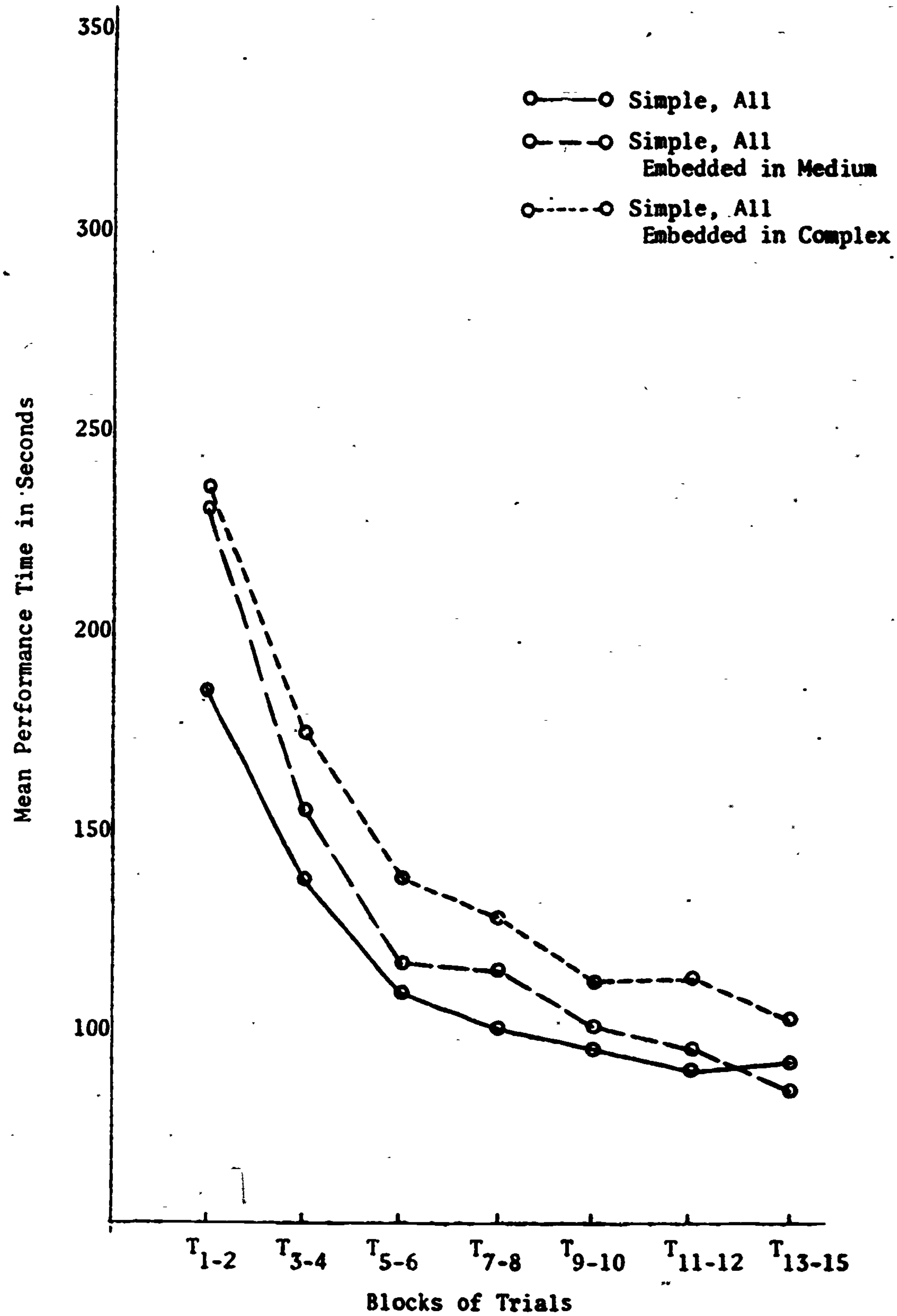


Figure 5. Mean performance time as a function of trial block during acquisition training for simple-all tasks

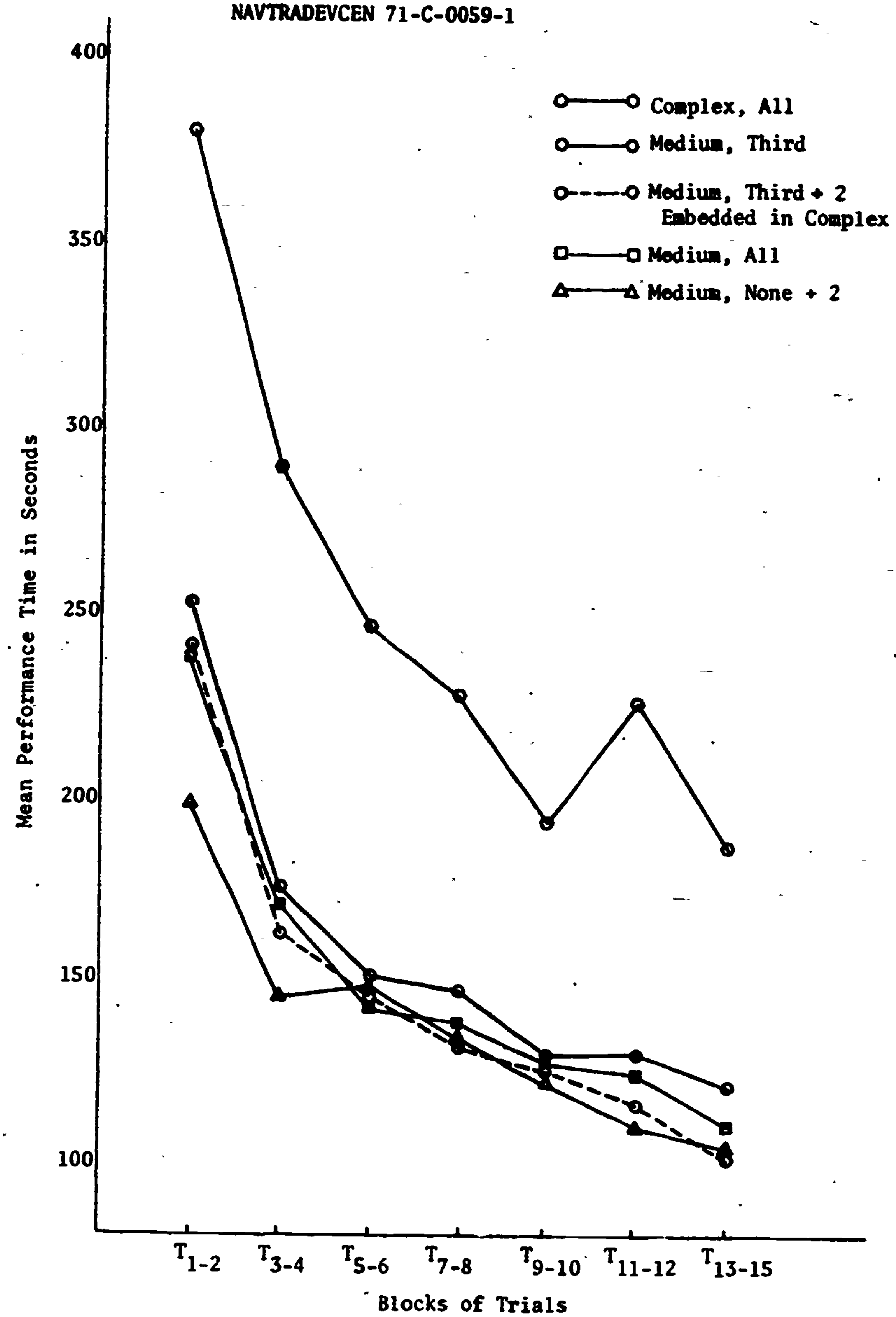


Figure 6. Mean performance time as a function of trial block during acquisition training for medium and complex tasks

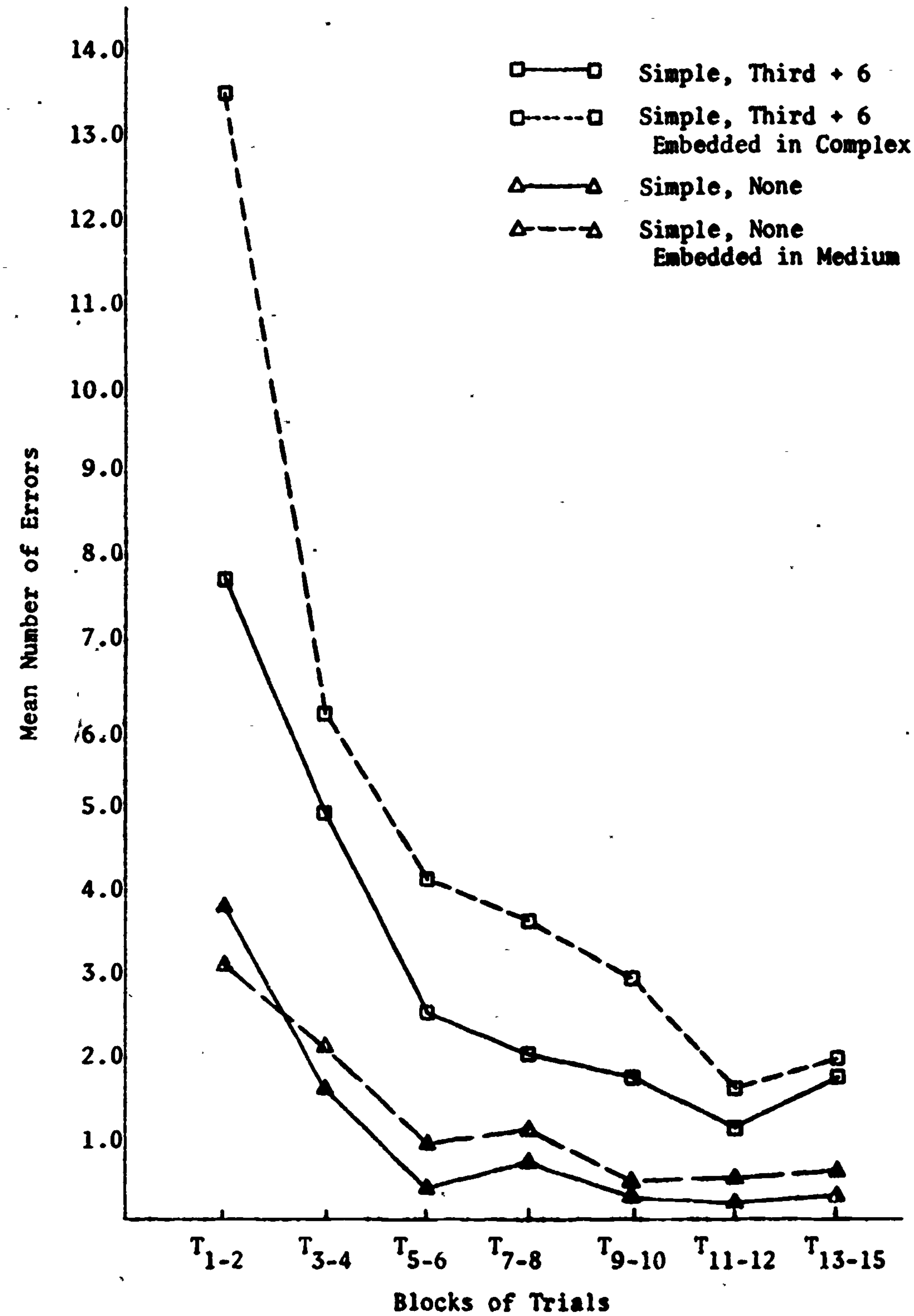


Figure 7. Mean errors as a function of trial block during acquisition training for simple-third and simple-none tasks

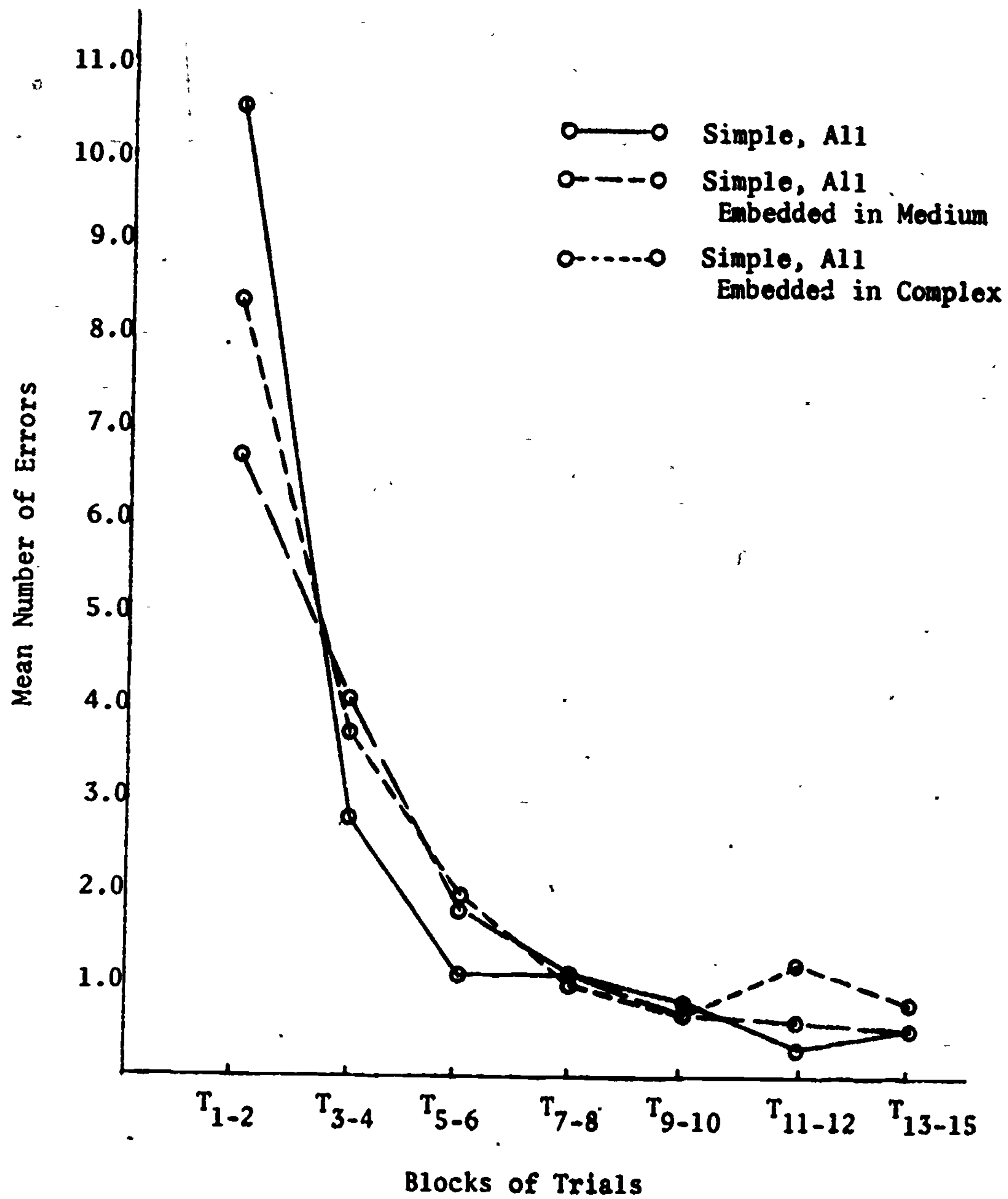


Figure 8. Mean errors as a function of trial block during acquisition training for simple-all tasks

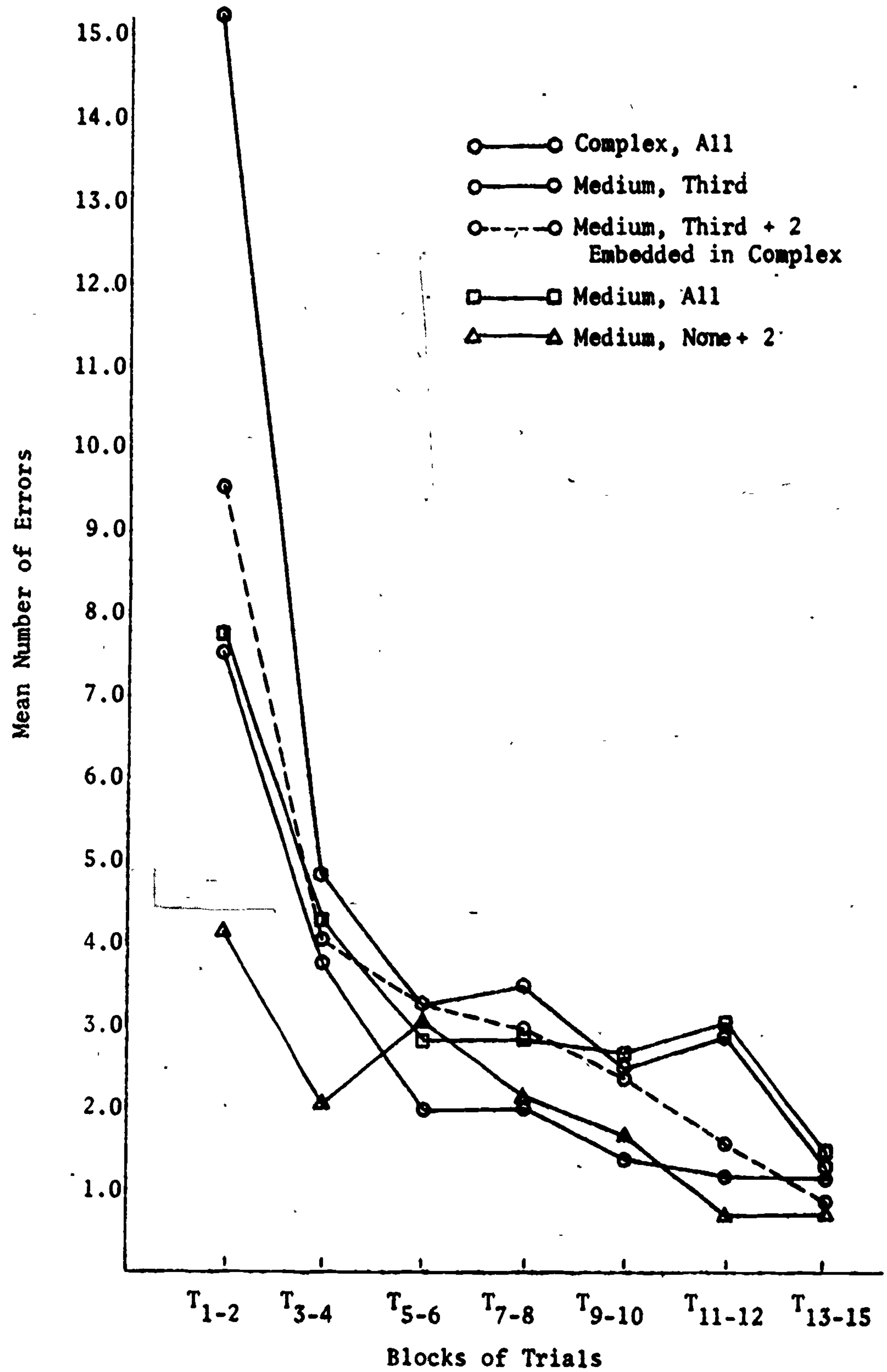


Figure 9. Mean errors as a function of trial block during acquisition training for medium and complex tasks

Closer inspection of both sets of data shows that learning occurred on all tasks. The training regimen brought about a consistent reduction in the time required to perform each task as well as in the number of errors made. In the case of the "simpler" tasks, time and error scores appear to be reaching asymptotic levels (figures 4 and 5, 7 and 8). On the medium and complex tasks, however, continued improvement is still noticeable (figures 6 and 9).

It is of interest that in both the time and error data there are two apparent sources for the observed differences among the various plots. The first is related to type of task, while the second involves task embedding. In this connection task refers to a specific set of procedural responses performed in a prescribed sequence. Embedding refers to the degree to which all of the displays and controls available for use are indeed used during task performance.

Variation in performance time due to type of task is clearly seen when the simple-none, simple-all, and simple-third + 6 plots are compared (figures 4 and 5). The consistent ordering in performance time throughout acquisition holds up for all task types with the single exception of the medium-none + 2 task (figure 6). With respect to error scores, the clearest consistent difference is seen between the simple-none and simple-third + 6 tasks (figure 7).

Particularly noteworthy are the different levels of performance associated with task embedding. For example, time (figure 4) and errors (figure 7) are both greater for the embedded versions of the simple-third + 6 and simple-none tasks. For simple-all tasks, this relationship holds only with respect to the time measures which increase as a function of degree of embedding (figure 5). With only two reversals, the differences in performance associated with task embedding are maintained throughout acquisition. The amount of training provided, although reducing the initial spread among these groups, is insufficient to eliminate the effects of extraneous displays and controls. This finding is made all the more interesting by the fact that performance for these simple task groups appears to be reaching an asymptote (figures 4, 5, and 8). The relationship is not as clear in the case of the medium-third task, which behaves as the simple embedded tasks do with respect to error (figure 9), but shows the opposite relationship for time (figure 6).

In much of the criterion data just described, relationships are strongly implied between performance during acquisition and the type of task to which subjects are exposed. The fairly consistent ordering of tasks with respect to performance level directly raises an issue of basic concern in the present research. To what extent are the indices, descriptive of the various trainer configurations, related to criterion performance? The Pearson product-moment correlation coefficients shown in tables 3 and 4 bear on this issue.

As shown in table 3, correlations of task indices with mean performance time at each trial block are, in general, highly consistent. With the exception of three variables (D%, C%, and E%), all reported coefficients are significant ($p < .05$). The three exceptions are in themselves interesting because of the consistently small correlations which they exhibit across all seven trial blocks. The same general pattern of relationships is also found in the mean error data reported in table 4. D%, C%, and E% fail to correlate substantially with mean error at any of the trial blocks. All other indices do exhibit substantial correlations with the error criterion. With the exception of the AA% and DEI indices, however, the correlations with error are neither as strong nor as consistent as they were with the performance time criterion.

TABLE 3. INTERCORRELATIONS OF TASK INDEX VALUES AND MEAN PERFORMANCE TIMES ACROSS TRIAL BLOCKS FOR THE LABORATORY TASKS†

Task Indices	Trial Blocks						
	1	2	3	4	5	6	7
MAIN	73	81	83	86	88	94	88
CNTG	78	82	84	86	90	91	86
TA	77	82	85	87	91	94	88
CONT	66	72	80	83	87	90	83
DISP	65	71	68	72	76	83	74
N	69	75	77	81	84	90	82
LV	74	80	81	84	88	92	85
AA%	-75	-75	-85	-80	-83	-73	-79
F%	-65	-62	-75	-76	-83	-72	-71
DEI x10-4	-83	-79	-86	-85	-89	-77	-79
D%	-06	-01	06	08	13	17	18
C%	-04	-01	08	10	15	17	17
E%	-12	-06	03	05	10	14	15
CRPS	73	77	87	88	92	90	87
FBR	70	76	69	73	75	82	76
INFO	72	79	79	83	85	92	84
INST	71	81	80	81	79	90	86

†Decimal points have been omitted from coefficients for clarity.

With 10 degrees of freedom: $r \geq .708,\ p \leq .01$
$r \geq .576,\ p \leq .05$

TABLE 4. INTERCORRELATIONS OF TASK INDEX VALUES AND MEAN ERRORS ACROSS TRIAL BLOCKS FOR THE LABORATORY TASKS†

Task Indices	Trial Blocks 1	2	3	4	5	6	7
MAIN	59	28	41	57	48	73	18
CNTG	65	46	58	69	66	86	36
TA	63	39	51	64	59	81	28
CONT	46	19	46	61	53	69	17
DISP	58	32	34	46	43	78	14
N	55	28	41	55	50	78	16
LV	61	35	45	59	54	80	22
AA%	-61	-62	-83	-89	-88	-73	-73
F%	-49	-41	-75	-76	-76	-66	-43
DEI $x10^{-4}$	-67	-72	-93	-88	-88	-77	-68
D%	-07	-24	-19	-10	-04	07	-07
C%	-04	-20	-12	-01	05	09	-06
E%	-13	-30	-23	-11	-06	05	-11
CRPS	54	33	62	73	67	72	34
FBR	65	42	35	47	44	80	22
INFO	61	33	40	55	49	79	19
INST	60	27	28	45	33	57	16

†Decimal points have been omitted from coefficients for clarity.

With 10 degrees of freedom: $r \geq .708$, $p \leq .01$
$r \geq .576$, $p \leq .05$

Of particular concern in both tables 3 and 4 are the generally large coefficients associated with the TA index. TA, representing the total actions or total number of responses comprising a task, correlates positively and highly significantly ($p < .01$) with all time scores. Although the coefficients are generally smaller, TA also exhibits a strong relationship with error scores (table 4). By themselves, these relationships are of trivial interest. They simply reflect the fact that the longer a task is, the more time will be required for its performance and the more potential errors there will be. What is disturbing, however, is that the relationships between the other indices and the performance criteria may arise because of dependencies between the remaining indices and TA.

During construction of the various trainers, concern arose over this very point. As previously mentioned, it was extremely difficult to manipulate many of the indices completely independently of TA. Examination of the task index intercorrelation matrix (not shown) confirms this impression. TA correlates significantly with all other task indices ($p < .01$), with the exception of D%, C%, and E%. With respect to the basic criterion data, therefore, it is unclear to what extent the other indices themselves relate to the criteria or simply mirror TA's relationships.

In an attempt to minimize potential contamination due to TA's influence, acquisition time and error scores were transformed prior to further analysis. The data selected for treatment were from the first, fourth, and seventh trial blocks, these points being chosen to represent performance at early, intermediate, and later stages of acquisition. Time and error data sets for each of the three trial blocks were treated separately. For each data set, single variable regression analyses were conducted using TA as the independent or predictor variable. This procedure resulted in sets of residual criterion scores from which all variance related to TA had been removed. The residual scores were simply the difference between the observed raw score values and the values predicted by the TA variable.

Evidence that the residualizing procedure had its intended effect comes from two sources. First, correlations between TA and the residual scores are zero. Second, correlations between the other (16) task indices and the residual criteria are greatly reduced. The only significant correlation is between E% and performance time at the first block ($r = -.58$, $p < .05$). Relationships among the predictor task index variables are, of course, undisturbed by the adjustment procedure. TA is no longer included in this set and appears in none of the regression analyses described below.

Six separate regression analyses were performed, one for each of the three time and three error criterion data sets. A step-wise (step-up) regression procedure was employed with a maximum of four predictor variables being fitted. Standard values were employed for the F-level criteria for predictor variable inclusion or deletion. The results of the six analyses are summarized in table 5. For each analysis, denoted by type of criterion, the multiple correlation coefficient (R) is reported together with the percentage of variance in the criterion accounted for (R^2). Also provided are the degrees of freedom (df) used in testing the significance of R and the resultant F-value. Finally, the specific indices included in each regression solution are listed. They appear from left to right in the order in which they were entered by the step-wise

TABLE 5. SUMMARY OF MULTIPLE REGRESSION ANALYSES OF PERFORMANCE TIME AND NUMBER OF ERRORS FOR FIRST, MIDDLE, AND LAST BLOCK OF ACQUISITION TRIALS

Criterion	R	R^2	df†	F	Indices in order of selection by step-wise regression program
			Time Scores		
T_{1-2}	.780	.608	3, 8	4.69*	E%, AA%, D%
T_{7-8}	.744	.553	3, 8	3.30	E%, AA%, DISP
T_{13-15}	.626	.392	3, 8	1.72	AA%, C%, DISP
			Error Scores		
T_{1-2}	.651	.423	3, 8	1.96	E%, C%, D%
T_{7-8}	.896	.802	3, 8	10.80**	AA%, MAIN, D%
T_{13-15}	.875	.766	3, 8	8.73**	AA%, CONT, DEI

*$p < .05$.

**$p < .01$.

†Sample size (N) = $df_1 + df_2 + 1$.

procedure. Only three indices are shown even though in all cases four were fitted. The small sample size (N = 12) suggested a conservative approach to description of the predictor indices.

As shown in table 5, when the effects upon performance time due to (TA) number of responses are removed, a significant multiple correlation between task indices and time is found only during the very early stages of acquisition ($R = .780$, $p < .05$). The relationship is between mean performance time and E%, AA%, and D%. The first and last of these indices reflect the extent to which superfluous equipment elements, especially displays, are encountered during task performance. One interpretation is that extraneous equipment has a distracting value which initially retards performance time, but whose impact decreases as the trainee masters the figure-ground (task-configuration) distinction. In line with this hypothesis, only E% is entered into the solution at T_{7-8}, while neither E% nor D% is entered at T_{13-15}. Also consistent with this same idea the zero-order correlations of E% and D% with residual time scores are negative and decrease over trial blocks. [For E%, $r = -.58$, $-.49$, and $-.30$; for D%, $r = -.52$, $-.45$, and $-.29$.]

As shown in table 5, a complementary situation exists with respect to relationships between task indices and error scores. That is, no relationship exists early during acquisition, but strong relationships emerge toward the end of training. By the middle of training, AA%, MAIN, and D% are significantly correlated with the mean number of errors being made ($R = .896$, $p < .01$). AA%, MAIN, and D% individually, however, have non-significant zero-order correlations with residual error scores at this time point (i.e., $r = -.57$, $-.08$, $-.42$). During the final block of trials the relation between indices and error scores is still significant ($R = .875$, $p < .01$). The mixture of related indices has changed, however. MAIN and D% have been replaced by CONT and DEI, while AA% is still present, as it is in five of the six analyses. The zero-order correlations of AA%, CONT, and DEI with residual errors are $r = -.55$, $-.10$, and $-.50$ respectively.

More generally, both sets of data show that task indices of the type employed in the present study can be related to learning or performance criteria. The strength of the obtained relationships suggests that it may be possible to use task index information to predict training criterion levels.

TRANSFER. The basic criterion data for the pilot transfer study are shown in figures 10 and 11. In each case either mean performance time (figure 10) or mean number of errors (figure 11) is plotted as a function of trial block with task configuration used during acquisition as the parameter. The ten transfer trials actually administered have been collapsed into five blocks. Therefore, each point in these figures represents an average value for ten scores.

In both figures the results are expressed in terms of performance on the medium-all task. In each case six different plots are shown. Two of these are used as frames of reference. The first portrays performance of the medium-all group during the first portion (trials 1 to 10) of the acquisition session. The second plot shows the performance of this same group during the later, transfer session. All groups rested for one-half hour between acquisition and transfer sessions. The remaining four plots portray performance on the medium-all task during the transfer session, after practice was given on interpolated tasks during acquisition.

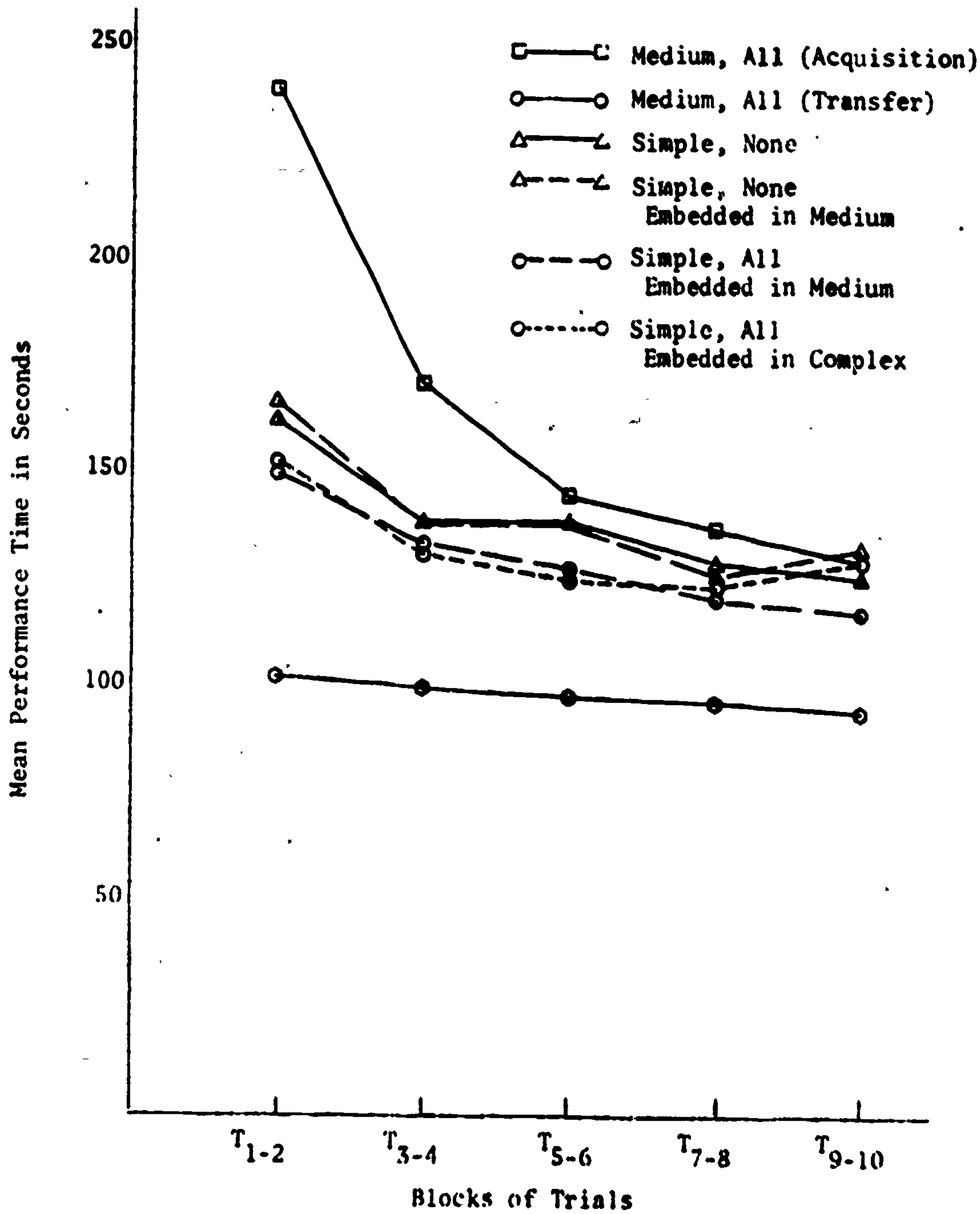


Figure 10. Mean performance time as a function of trial block during transfer to medium-all task.

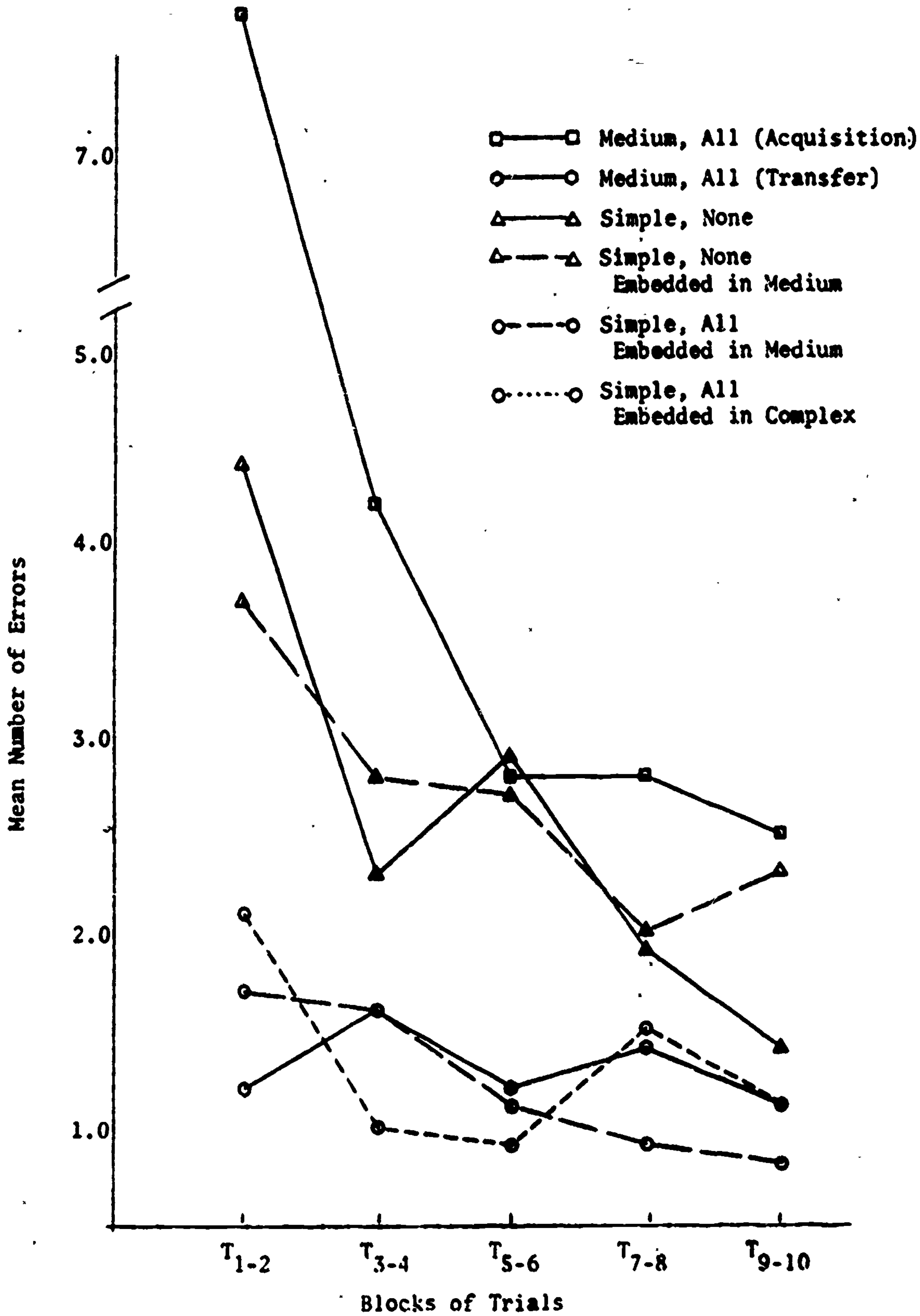


Figure 11. Mean errors as a function of trial block during transfer to medium-all task

In figure 10, the medium-all subjects provide an extremely clean baseline in performance time against which the other functions may be viewed. Performance time for this group is apparently at asymptote and clearly represents an improvement over the times achieved during acquisition. The interpolated task groups show a slight reduction in performance time during transfer, but across all blocks are slower than the medium-all (transfer) group ($p < .05$). Even more interesting, perhaps, is the fact that the interpolated groups are significantly faster than the medium-all acquisition group only at the first two blocks ($p < .05$). Thereafter the interpolated task and medium-all acquisition data are indistinguishable. This is in spite of the fact that the interplated groups have, by the third block, had 3.5 times as much practice on set-up consoles.

The breakout due to embedding which occurs during acquisition is not obtained in the transfer time data. Furthermore, there is only the barest hint of a difference in performance time during transfer due to interpolated task type.

The error data shown in figure 11 show a slightly different set of relationships. The baseline mean number of errors for the medium-all group is somewhat variable, though approaching what appears to be an asymptote. Again, there clearly are lower numbers of errors made by this group during transfer than during acquisition. As in figure 10, there is no suggestion of an effect on errors made due to task embedding.

Particularly noteworthy, however, is the evidence for a task-type effect upon error scores which was not so clearly seen in the time data. The simple-none tasks have significantly fewer mean errors than the medium-all acquisition group only at the first block ($p < .05$). Significantly fewer mean errors are associated with simple-all tasks, relative to the medium-all acquisition group, on all but the last block of trials ($p < .05$). Conversely, the simple-all groups have significantly fewer errors than the simple-none task across the first three blocks of trials ($p < .05$).

Considered jointly, the pilot data presented in both figures suggest that the simple-all subjects can perform well during transfer with respect to accuracy but that they pay a price in terms of speed. On the other hand, groups which were trained on more dissimilar trainers (simple-none groups) pay a price in terms of both speed and accuracy.

FIELD FINDINGS

The basic ratio estimation data obtained during the field study are shown in Appendix G. In each of four tables, representing the set-up, detection, localization, and classification subtasks, four criterion estimates are shown across training devices. Each datum represents the mean of instructors' consensus magnitude estimates relative to the values assigned to the standards for comparison. These standard values were arbitrarily set at 100, 50, 50, and 100 for the four types of criteria.

In any of the tables comprising Appendix G the first striking feature of the data is the difference in values across columns. This is, of course, primarily due to the use of different standards of comparison (i.e., 100, 50, 50, 100). The estimation data within any column, however, do show appreciable variability. On the set-up task (table G-1), for example, the first and fourth scales have ranges of 50-390 and 50-260, respectively. Although not as extreme, the second and third scales also show good variance. Finally, on all scales, mean estimates are obtained which lie both above and below the respective standard values. These aspects of the data suggest that the ratio estimation procedure which was employed apparently succeeded in spreading out estimates across devices. As in the laboratory, reasonable variance in the criteria was a necessary condition for achieving any predictability.

Two additional types of variation are of interest in these data. First, consider the amount of variation, within any subtask and on any specific scale, for similar devices found at different locations. In many cases agreement is extremely good. In others it is not. On the training time scale for the set-up task (table G-1), for example, a fairly large difference between OA1283 stacks exists. The BQR-2B and 2C stacks, however, lead to amazingly similar judgments. A more thorough examination of these issues is underway, the details of which are beyond the present level of analysis.

Another interesting variation is seen when one focuses on a specific device and scale, and then looks across subtasks. But before subtasks can be compared, any differences between the <u>standard task examples</u> have to be removed. Toward this end, instructors <u>in the present study</u> scaled the detection, localization, and classification <u>standards</u> relative to the set-up standards. Based upon these data, averaged <u>across all</u> instructors, a set of weights was derived for each subtask. The weights for the first two criteria are shown at the bottom of the tables in Appendix G for each subtask. Using these weights, for example, one would conclude that classification training time on the 14E3 is almost seven times longer (212 x 1.81 = 384) than localization training (49 x 1.13 = 55). Since comparisons of this type were of interest in the present study, weighted consensus scores were used in all subsequent analyses. Use of these transformed estimates also made a number of combinatory analyses possible.

In Appendix H, zero-order, product-moment correlation coefficients are shown in separate tables for each of the four subtasks. The coefficients describe the relation between task indices and criterion estimates. Two features of the data are of interest. First, significant relationships between individual criteria and indices are obtained and cut across all four subtasks. Second, for the most part, when a task index exhibits a significant correlation with one criterion, its correlations with the remaining criteria also tend to be strong if not always significant. The redundancy among criteria implied by this observation is confirmed when the intercorrelations among criteria are examined. In all four subtasks, the correlations between estimated training time and task difficulty range between $r = .96$ and $r = .84$. Those for proficiency level and transfer lie between $r = .92$ and $r = .96$. The correlations between training time and proficiency level estimates, while still significant, tend to be somewhat lower (i.e., $r = -.67$ to $r = -.89$). Because of this smaller redundancy, and because these two estimates were in a sense analogous to criteria employed in the laboratory, they alone were chosen for analysis. In the following analyses (C_1) denotes the training time estimate, and (C_2) stands for the proficiency level judgment.

Finally in Appendix H, significant correlations are shown between the TA variable and the two criteria selected for analysis. TA represents the number of actions or responses comprising a task. In the flow charts examined by the instructors it was possible to convert TA rather directly and perhaps superficially into a concept of task length or difficulty. To reduce the impact of flow-chart length upon instructor estimates and to use data analogous to those analyzed in the laboratory, the regression adjustment procedure was used again. The C_1 and C_2 data were transformed into residual scores for analysis, thereby reducing that portion of criterion variance associated with TA. Resultant correlations between the remaining 16 task indices and the residual criterion scores were greatly reduced.

Results of the seven distinct regression analyses performed on the training time (C_1) and proficiency level (C_2) residual data are summarized in table 6. The column headings are the same as those previously used in reporting the laboratory data (table 5). Four of the seven analyses are at the basic subtask level. The remaining three are combinatory and examine different poolings of the subtasks. Set-up and localization are pooled because they seem to represent cases in which the trainee interacts most directly with his stack, particularly in making control settings and adjustments. The detection and classification tasks are pooled because of their perceptual, signal processing flavor. At the highest level of analysis, all four subtasks are examined simultaneously.

In table 6 significant relationships are shown between selected task indices and the instructor ratio estimate criteria. These relationships are obtained in spite of the highly conservative procedure of using residual scores, a procedure which greatly reduced the zero-order correlations between predictors and criteria. Significant relationships are established in all but two of the analyses. The multiple correlations associated with the classification and set-up tasks are not significant by conventional standards ($p < .05$). However, the fact that more than half of the variance is accounted for in the set-up (C_1) analysis cannot be ignored ($p < .10$).

One of the most interesting features of the data shown in table 6 is that the patterns of indices which contribute to significance change from subtask to subtask and from individual subtasks to pooled subtasks. The DEI index, for instance, while related to both criteria in the overall analysis, does not fall out in the intermediate poolings. It does appear, however, at the single task level. Similarly, AA%, which is one of the primary indices at the intermediate level, disappears from the overall analyses. These shifting patterns imply that different index factors may be required, depending upon the subtask under examination.

TABLE 6. SUMMARY OF MULTIPLE REGRESSION ANALYSES OF INSTRUCTORS' RATIO ESTIMATES: INDIVIDUAL SUBTASKS AND POOLED SUBTASKS

	R	R^2	df	F	Indices in order of selection by the step-wise regression program
					All Tasks
C_1	.597	.356	7, 37	2.92*	D%, INFO, CNTG, F%, DEI, E%, CONT
C_2	.658	.433	7, 37	4.04**	INFO, MAIN, LV, D%, E%, CONT, DEI
					Set-up + Localization
C_1	.628	.395	4, 20	3.25*	AA%, D%, INFO, DISP
C_2	.644	.415	4, 20	3.54*	DISP, INFO, MAIN, C%
					Detection + Classification
C_1	.765	.584	4, 15	5.28**	F%, AA%, CNTG, E
C_2	.358	.128	4, 15	0.55	CONT, FBR, CNTG, MAIN
					Set-up
C_1	.741	.550	3, 9	3.66	E, DEI, LV
C_2	.615	.379	3, 9	1.83	D%, E, DEI
					Detection
C_1	.892	.796	2, 7	13.64**	INST, CONT
C_2	.811	.658	2, 7	6.73*	E, DEI
					Localization
C_1	.848	.719	3, 8	6.84*	D%, CRPS, DEI
C_2	.629	.395	3, 8	1.74	DISP, MAIN, FBR
					Classification
C_1	.569	.324	2, 7	1.67	F%, D%
C_2	.448	.201	2, 7	0.88	F%, DEI

1. C_1 = Training time needed to achieve instructor proficiency.
2. C_2 = Proficiency level after 2 hours of practice on the device.

*$p < .05$
**$p < .01$

SECTION IV

DISCUSSION

In this section the results which have been detailed in Section III are summarized separately for the laboratory and the field. The significance of these results for task quantification and performance prediction is then discussed. Finally, major conclusions and implications for future research are drawn.

PREDICTION OF SET-UP TASK SKILL ACQUISITION

The results of the laboratory acquisition study generally showed wide variation in performance as a function of task/trainer configuration, variations which were at least intuitively systematic. Furthermore, the systematic spreads in performance, established early in training, were generally maintained throughout acquisition. This is particularly significant because performance tended to reach stable, asymptotic levels toward the end of acquisition. Finally, regression analysis demonstrated a substantial amount of significant correlation between the task indices and performance.

The predictability which was obtained is all the more significant because the prepotent effects of total actions (TA) were statistically eliminated. This predictability was also obtained in spite of a number of sources of error variation which were not dealt with to our complete satisfaction. These included variations due to subjects, variations due to the use of two experimenters, and restrictions in the ranges of some of the index values. For example, DEI for the field devices ranged from 10 to 500 X 10^{-4}. In the laboratory we obtained a range of 5 to 21 X 10^{-4}. This restriction may have accounted in part for the somewhat different patterns of predictors which emerged from the step-wise regressions for laboratory and field. More comparable ranges of index values may have increased the correspondence among the predictors.

The predictability obtained gains further significance because of its presence (in some sense) throughout acquisition; i.e., ability to predict performance from task indices was more than a Block 1 phenomenon. Moreover, there was some, though not perfect, consistency in the patterns of predictors which emerged over time: E%, AA%, and D%, for example, were selected by the stepwise program at more than one block.

But, while predictability was possible throughout acquisition the relationship between type of predictability and phase of training was not a simple one. A significant multiple R was obtained early in training using the time criterion, but later in acquisition, significance was obtained with the error criterion. A possible explanation for this pattern of modes of predictability is that all the devices were equally error prone on Block 1 (i.e., T_{1-2}), but that differential elimination of errors occurred by Block 7 (i.e., T_{13-15}). Differential elimination of time effects is also possible, of course, but appears less likely. It was apparent to the experimenters during data collection that on more complex devices, subjects tended to rush through long sequences of calibration type responses with attendant carelessness in setting controls or reading displays.

The results of the acquisition study have a number of implications. First, they support the feasibility of differentiating set-up performance on sonar type stacks by manipulating panel design. Such differentiation is critical if any predictability from task indices is going to be possible. They suggest further that it is, in fact, possible to relate such performance explicitly to design parameters, even when those parameters are purged of effects of variables which are prepotent, but of trivial interest.

The implication of removing TA, eliminating most of the zero-order correlation, and still obtaining significant multiple correlations is that the multivariate approach is essential; i.e., individual task indices did not appear capable of predicting performance on our training devices. Rather, collections of indices, with perhaps specific, but as yet, unidentified patterns of features, are crucial. Moreover, there is some hint in the results that these patterns may depend upon training stage, though some indices did appear to occur rather often.

In addition to implying that predictor patterns may vary with stage of training, the results also imply that criterion patterns may be similarly influenced. Thus, the designer may have to ask—not whether indices relate to training effectiveness, but what patterns of indices relate to what criterion of effectiveness at what stage of training. This is a question which the present research cannot answer.

An interesting sidelight is provided by the ordering effect due to embeddedness. This effect implies that there may be value in using overlays to train the set-up task, much in the same way that overlays are used to teach anatomy or to facilitate the performance of an assembly line, electronics inspector. Through such a device, small sets of related details can be presented, while other immediately unrelated details are held back temporarily. Given that embeddedness does, in fact, substantially retard training of set-up tasks, it would be of interest to determine whether the use of successive overlays can improve this training.

PREDICTION OF SET-UP TASK TRANSFER OF TRAINING

The transfer study was a pilot effort to relate some specific design variations to transfer of training from a "simple" device to a more complex device. This might have a very approximate parallelism to training on a synthetic device in the school setting and then going to a specific stack in the field.

The particular configurations which were used reflected increasing values of embeddedness, as reflected in DEI, and increasing numbers of total actions needed to complete the task, as reflected in TA and DEI. Here TA was not a trivial variable because, unlike the case for acquisition, TA (on the acquisition task) did not directly affect performance on the task being measured (the transfer task).

The results were encouraging because there was an intuitively systematic ordering of the configurations on Block T_{1-2} of the transfer session. While true for both speed and accuracy of performance, it was particularly striking in the latter case. Error was proportional to the distance (along a similarity dimension) between interpolated and transfer tasks. This ordering was supported by correlational analysis of DEI and TA which showed significance at $p < .05$. These results suggest that it may very well be feasible to predict transfer effects from quantitative indices. Emphasis, of course, has to be placed on the word suggest, since we intended here only to obtain some pilot information. The small number of cases can, at best, only provide encouragement for pursuing this line of investigation in a more rigorous fashion.

With respect to the time criterion, results showed that the interpolated groups (i.e., the "simple" groups) never caught up with the group originally trained on the medium-all device. They lagged in performance speed during the entire transfer session. This suggests that operators trained on synthetic devices and then transferred to field stacks might pay a price in speed, which is not readily mitigated, though conceivably they could attain a satisfactory level of accuracy. Herein lies another very interesting and pragmatic line of investigation.

PREDICTION OF JUDGED TRAINING EFFECTIVENESS

The extent of correspondence between results obtained in the laboratory and results obtained in the field was substantial and very encouraging. This is particularly so, given the expected softness of the field estimation data.

First, the success obtained in the laboratory in generating performance variation across devices was continued in the field. Mean estimates of trainer "effectiveness" showed wide dispersion across the 13 field stacks which were studied. In particular, the set-up and localization tasks generated wide variance. While performance variability was not as great for detection and classification, it was nonetheless substantial.

When the ratio estimate data were scaled for effects of standard task difficulty, it was seen that there was considerable variation across subtasks, within devices as well as across devices. This finding, coupled with results of the regression analysis, supports the contention that device effectiveness may depend very heavily on the manner in which the device is used. That is, it may not be feasible to talk about the effectiveness of a training device in generic (i.e., figure of merit) terms, but only in terms of the use to which the device is being put (i.e., in training specific subtasks). Thus, we can extend a prescription stated earlier—the designer may have to ask what patterns of indices relate to what criteria of effectiveness at what stage of training for which subtask.

Second, correlational support for the indices was obtained. Relationships between task indices and judged device "effectiveness" were demonstrated, though not for all the subtasks. Unfortunately, one of the subtasks not included among the significant correlations—set-up—was of primary importance here, since it provided the only generalization test of successful prediction in the laboratory. Therefore, comparisons between the field and laboratory must be made with due caution.

Some correspondence was obtained between the patterns of indices which were selected by the regression program for the field data and for the laboratory data. DEI and D% were common to and prominent in both field and laboratory set-up. The E index was selected by the regression program for field set-up, while E% was selected for the laboratory set-up. The AA% index was prominent in the laboratory, but did not appear at all for field set-up. The discrepancies may have in part occurred because many of the indices are difficult to reflect in the types of flow chart used in the field. The E% index, for example, directly measures the amount of "clutter on a panel"; i.e., displays and controls which are not used for a particular subtask. Though the instructors were thoroughly familiar with their devices, they did work from flow charts rather than from the actual device and may not have considered such factors as E%. The AA% index reflects looping behavior, and similarly may not have been fully appreciable from the flow charts. In view of these methodological problems, it was gratifying to find as much correspondence as did appear.

While it was possible to demonstrate validity for patterns of predictors, those patterns were not uniform across subtasks just as the patterns were not entirely uniform across time in the laboratory work. For example, while DEI entered into the regression for both localization and detection, it entered for different criteria. And other than DEI there was no index which was common to both localization and detection. When either of these subtasks was pooled, an almost completely different pattern of predictors emerged. Somewhat consistent with the laboratory data and the statement made earlier concerning variability across subtasks, these facts would appear to indicate that different patterns of predictors; i.e., different "factors" are needed depending upon the particular use to which the trainer will be put. They also suggest the possible fruitfulness of a factor analysis of the predictor and criterion data.

CONCLUSIONS AND IMPLICATIONS FOR RESEARCH

The current research effort has supported the feasibility of relating quantitative indices of equipment design to performance, at least for the restricted set of indices and trainer stacks examined in the present study. The effort has also supported the feasibility of predicting transfer effects from equipment design indices. Moreover, this predictability has been demonstrated for sets of indices corrected for the effects of a prepotent but trivial factor and has been shown to be more than a transient event.

Clarification of the specific meaning of the data and development of a practical methodology require further research. In general, the following efforts appear warranted:

a. Results of the present laboratory and field work need to be generalized to other classes of trainers. This includes determining the applicability of the current set of indices to other devices. An attempt is also required, if possible, to validate the laboratory results in the field based on trainee performance data.

b. Judgments obtained in the field via opinion sampling should be validated against actual performance measurement.

c. Relationships between quantitative indices and transfer of training require more rigorous investigation—at least in the laboratory, but preferably under field conditions also.

d. Effects of individual differences and conditions of training need to be interwoven with effects of training task variation.

Subsequent phases of this program will deal with one or more of the issues raised above.

REFERENCES

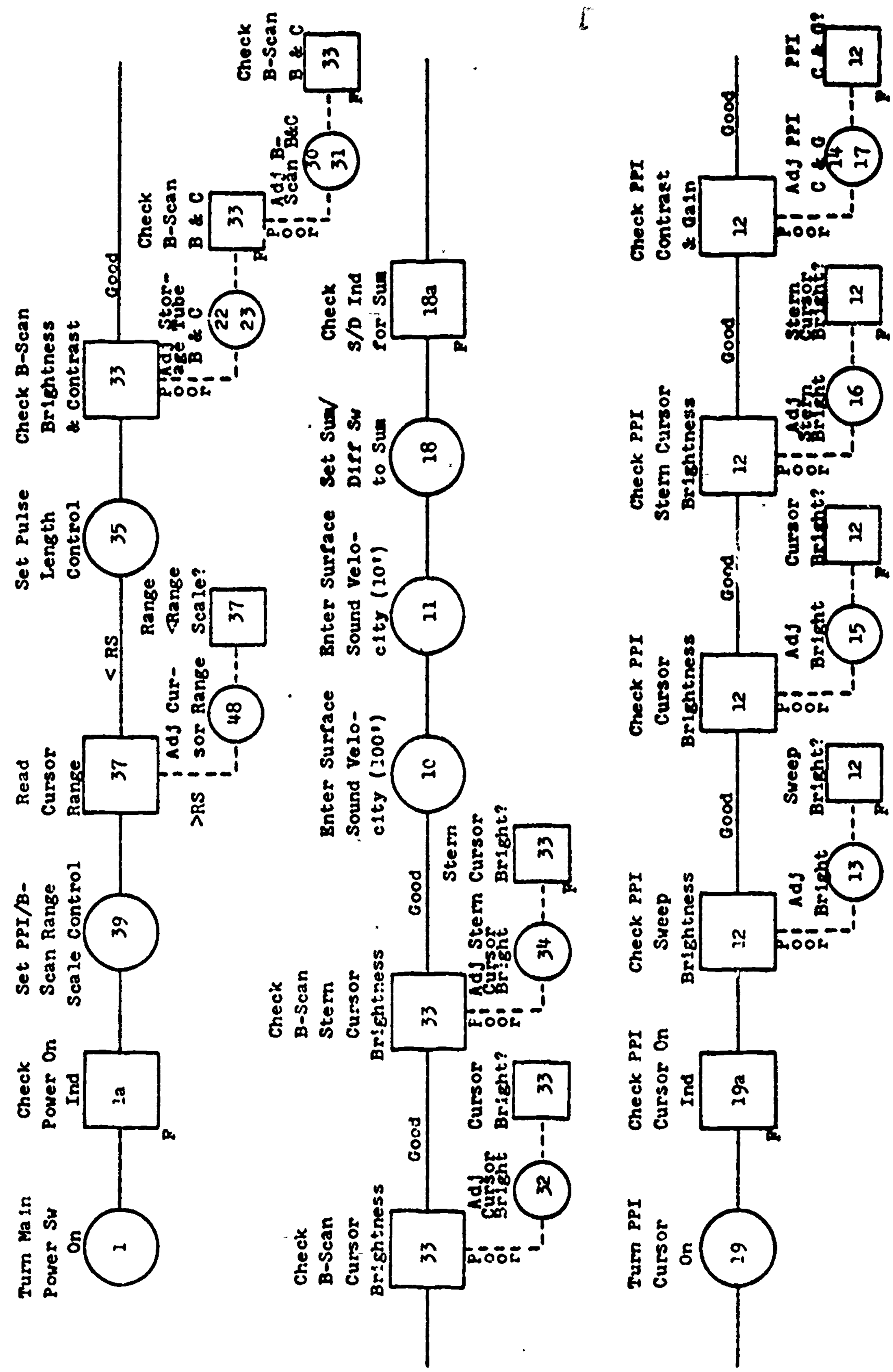
Fowler, R. L., Williams, W. E., Fowler, M. G., & Young, D. D. An investigation of the relationship between operator performance and operator panel layout for continuous tasks. Report No. AMRL-TR-68-170, December 1968. Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

Siegel, A. I., Miehle, W., & Federman, P. Information transfer in display-control systems. VI. A manual for the use and application of the display evaluative index (Sixth Quarterly Report). Applied Psychological Services, 1962, Wayne, Pennsylvania.

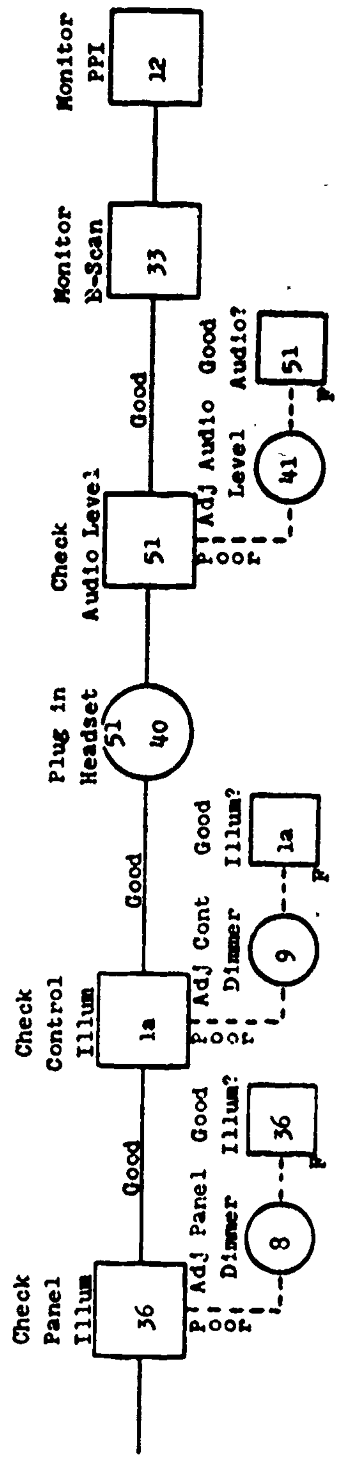
Wheaton, G. R., Mirabella, A., & Farina, A. J., Jr. Trainee and instructor task quantification: Development of quantitative indices and a predictive methodology. Report No. NAVTRADEVCEN 69-C-0278-1, January 1971. Naval Training Device Center, Orlando, Florida.

APPENDIX A

SQS-26CX Set-up Subtask Operations Flow Chart



Turn Main Power Sw On
1
Check Power On Ind
1a
Set PPI/B-Scan Range Scale Control
39
Read Cursor Range
37
>RS
Adj Cursor Range
48
Range < Range Scale?
37
< RS
Set Pulse Length Control
35
Check B-Scan Brightness & Contrast
33
Adj Storage Tube B & C
22
23
Check B-Scan B & C
33
Adj B-Scan B&C
30
31
Check B-Scan B & C
33
Good
Check B-Scan Cursor Brightness
33
Adj Cursor Bright
32
Cursor Bright?
33
Good
Check B-Scan Stern Cursor Brightness
33
Adj Stern Cursor Bright
34
Stern Cursor Bright?
33
Good
Enter Surface Sound Velocity (100')
10
Enter Surface Sound Velocity (10')
11
Set Sum/Diff Sw to Sum
18
Check S/D Ind for Sum
18a
Turn PPI Cursor On
19
Check PPI Cursor On Ind
19a
Check PPI Sweep Brightness
12
Adj Bright
13
Sweep Bright?
12
Good
Check PPI Cursor Brightness
12
Adj Bright
15
Cursor Bright?
12
Good
Check PPI Stern Cursor Brightness
12
Adj Stern Bright
16
Stern Cursor Bright?
12
Good
Check PPI Contrast & Gain
12
Adj PPI C & G
14
17
PPI C & G?
12
Good


Check Panel Illum
36
Poor
Adj Panel Dimmer
8
Good Illum?
36
Good
Check Control Illum
1a
Poor
Adj Cont Dimmer
9
Good Illum?
1a
Good
Plug in Headset
51
40
Check Audio Level
51
Poor
Adj Audio Level
41
Good Audio?
51
Good
Monitor B-Scan
33
Monitor PPI
12

APPENDIX B

Task Characteristic Index Values for Trainee Subtasks Evaluated in the Field

TABLE B-1. TASK CHARACTERISTIC INDEX VALUES FOR SONAR SET-UP TASKS

Device	Task indices								
	MAIN	CNTG	TA	CONT	DISP	E	LV	AA%	F%
21B55 OA1283	24	15	39	16	6	22	2500.2	62	94
21A39/2 OA1283	48	12	60	17	6	23	2927.8	46	86
21B55 BQR-2B	27	6	33	15	10	25	2616.7	78	95
21A39/2 BQR-2C	26	9	35	11	11	22	2399.9	67	82
21A39/2 BQR-7B	30	17	47	14	13	27	3149.9	66	62
14A2C1 SQS-23B	25	4	29	17	7	24	2466.7	84	100
X14A2 SQS-23 (TRAM)	17	12	29	15	5	20	2040.0	69	92
14E3 SQS-4	23	17	40	15	10	25	2740.0	67	95
14E14 SQS-4	26	14	40	15	11	26	2933.6	72	97
SQS-26CX	24	24	48	20	8	28	2799.8	57	95
14B10 AQS-13	41	5	46	19	6	25	2731.6	58	80
14B31B AQA-1	51	38	89	18	5	23	3680.3	39	93
14B31B ASA-20	112	7	119	22	5	27	4369.9	34	67

TABLE B-1. TASK CHARACTERISTIC INDEX VALUES FOR SONAR SET-UP TASKS (Cont)

Device	DEI x10^-4	D%	C%	E%	CRPS	FBR	INFO	INST
		Task indices						
21B55 OA1283	23.20	67	84	79	19	10	10	0
21A39/2 OA1283	26.84	67	85	79	30	24	6	1
21B55 BQR-2B	25.07	62	65	64	17	10	6	0
21A39/2 BQR-2C	65.53	85	61	71	14	10	11	0
21A39/2 BQR-7B	25.23	81	58	68	19	10	18	0
14A2/C1 SQS-23B	32.23	33	65	51	18	6	5	4
X14A2 SQS-23 (TRAM)	16.51	28	54	44	15	6	8	4
14E3 SQS-4	35.98	83	75	78	18	12	10	4
14E14 SQS-4	37.57	85	58	67	17	11	12	2
SQS-26CX	03.78	29	63	47	20	15	13	2
14E10 AQS-13	14.43	67	83	78	27	9	10	5
14B31B AQA-1	10.04	83	62	66	42	31	16	2
14B31B ASA-20	25.20	45	100	82	72	32	15	1

TABLE B-2. TASK CHARACTERISTIC INDEX VALUES FOR DETECTION TASKS

Device	Task indices								
	MAIN	CNTG	TA	CONT	DISP	E	LV	AA%	F%
21B55 OA1283	21	4	25	3	3	6	1883.3	69	0
21B55 BQR-2B	28	2	30	2	5	7	2000.0	60	0
21A39/2 BQR-2C	15	3	18	2	5	7	983.3	46	0
21A39/2 BQR-7B	19	3	22	2	5	7	1216.7	46	0
14A2/C1 SQS-23B	11	9	20	3	4	7	1053.4	43	0
X14A2 SQS-23 (TRAM)	21	4	25	4	2	6	1354.1	43	0
14E3 SQS-4	3	4	7	3	2	5	399.9	43	0
14E14 SQS-4	21	4	25	3	3	6	998.2	24	0
SQS-26CX	5	2	7	2	3	5	399.9	43	0
14E10 AQS-13	10	10	20	7	2	9	1057.3	46	0

TABLE B-2. TASK CHARACTERISTIC INDEX
VALUES FOR DETECTION TASKS (Cont)

Device	DEI x10^{-4}	Task indices D%	C%	E%	CRPS	FBR	INFO	INST
21B55 OA1283	234.04	33	15	27	12	3	10	6
21B55 BQR-2B	62.26	31	8	17	7	5	18	1
21A39/2 BQR-2C	48.21	38	11	23	4	3	11	0
21A39/2 BQR-7B	61.34	31	8	17	5	3	14	0
14A2/C1 SQS-23B	28.68	20	12	15	6	4	10	0
X14A2 SQS-23 (TRAM)	55.46	22	7	13	7	5	13	0
14E3 SQS-4	61.28	17	15	16	3	2	2	0
14E14 SQS-4	29.73	23	12	15	7	5	13	0
SQS-26CX	594.23	11	6	8	2	0	5	0
14E10 AQS-13	20.66	22	29	27	9	3	8	0

TABLE B-3. TASK CHARACTERISTIC INDEX VALUES FOR LOCALIZATION TASKS

Device	Task indices								
	MAIN	CNTG	TA	CONT	DISP	E	LV	AA%	P%
21B55 OA1283	25	10	35	5	5	10	2116.2	56	51
21B55 BQR-2B	36	8	44	8	8	16	2155.7	45	60
21A39/2 BQR-2C	19	12	31	4	5	9	1263.4	31	32
21A39/2 BQR-7B	6	9	15	2	5	7	600.0	28	29
14A2/C1 SQS-23B	11	20	31	6	4	10	1395.5	38	67
X14A2 SQS-23 (TRAM)	11	0	11	4	4	8	700.0	58	91
14E3 SQS-4	5	2	7	3	2	5	399.9	43	50
14E14 SQS-4	14	4	18	5	4	9	1000.0	49	100
SQS-26CX	14	9	23	5	6	11	1414.4	57	92
14E10 AQS-13	17	7	24	6	4	10	1334.4	50	75
14B31B AQA-1	36	21	57	7	6	13	2481.3	38	60
14B31B ASA-20	30	14	44	11	6	17	2224.5	47	58

TABLE B-3. TASK CHARACTERISTIC INDEX VALUES FOR LOCALIZATION TASKS (Cont)

Device	DEI x10-4	Task indices						
		D%	C%	E%	CRPS	FBR	INFO	INST
21B55 OA1283	111.69	56	26	36	14	8	13	5
21B55 BQR-2B	37.28	50	33	40	18	7	19	0
21A39/2 BQR-2C	25.85	38	22	29	10	4	17	0
21A39/2 BQR-7B	43.80	31	8	17	3	3	9	0
14A2/C1 SQS-23B	35.23	20	23	22	14	6	11	0
X14A2 SQS-23 (TRAM)	38.44	22	14	17	4	2	5	0
14E3 SQS-4	48.10	17	15	17	3	2	2	0
14E14 SQS-4	27.38	31	19	23	6	4	8	0
SQS-26CX	37.50	21	16	19	9	5	9	0
14E10 AQS-13	26.60	44	26	31	8	7	9	0
14B31B AQA-1	15.56	100	24	37	21	9	27	0
14B31B ASA-20	17.73	55	50	52	19	10	15	0

TABLE B-4. TASK CHARACTERISTIC INDEX VALUES FOR CLASSIFICATION TASKS

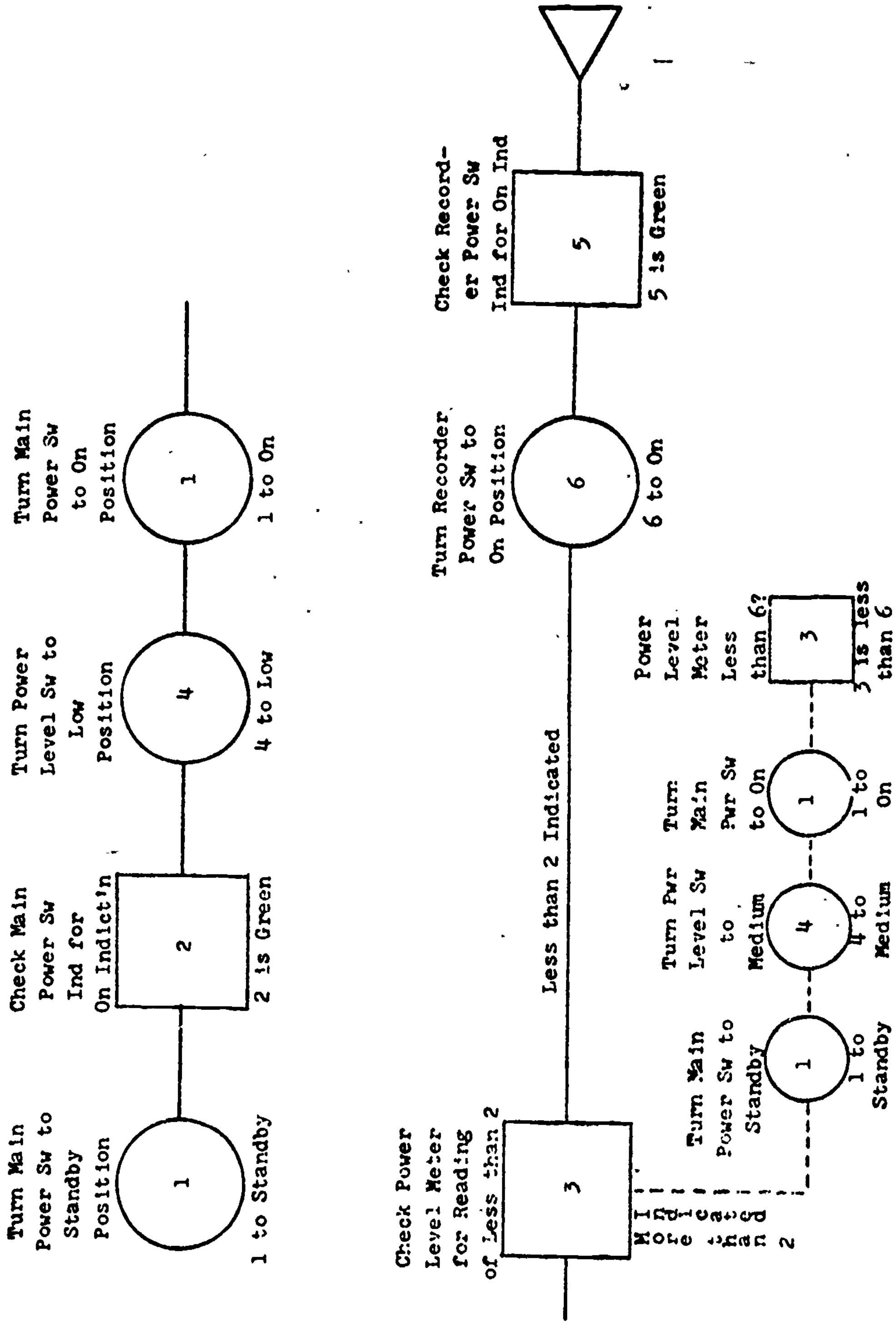
Device	Task indices								
	MAIN	CNTG	TA	CONT	DISP	E	LV	AA%	F%
21B55 0A1283	6	0	6	2	2	4	400.0	50	75
21B55 BQR-2B	6	0	6	2	2	4	300.0	25	75
21A39/2 BQR-2C	8	0	8	2	4	6	500.0	53	100
21A39/2 BQR-7B	12	0	12	3	3	6	850.0	64	50
14A2/C1 SQS-23B	12	3	15	2	2	4	1120.0	62	61
X14A2 SQS 23 (TRAM)	10	11	21	6	4	10	1133.3	48	52
14E3 SQS-4	13	6	19	4	4	8	1016.8	46	91
14E14 SQS-4	5	0	5	1	2	3	400.0	60	26
SQS-26CX	15	5	20	2	4	6	1311.3	57	68
14E10 AQS-13	8	1	9	4	2	6	466.7	40	100

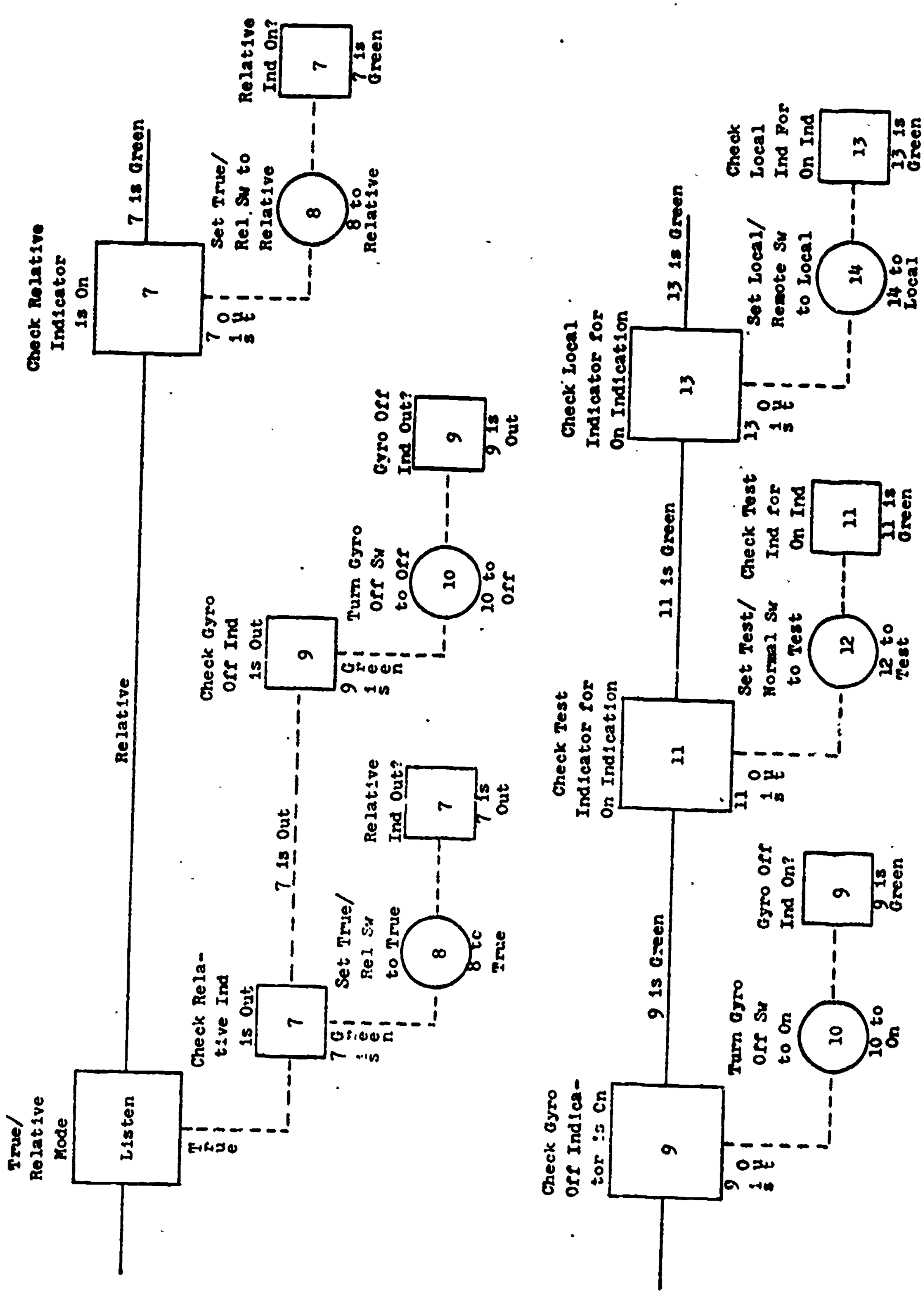
TABLE B-4. TASK CHARACTERISTIC INDEX VALUES FOR CLASSIFICATION TASKS (Cont)

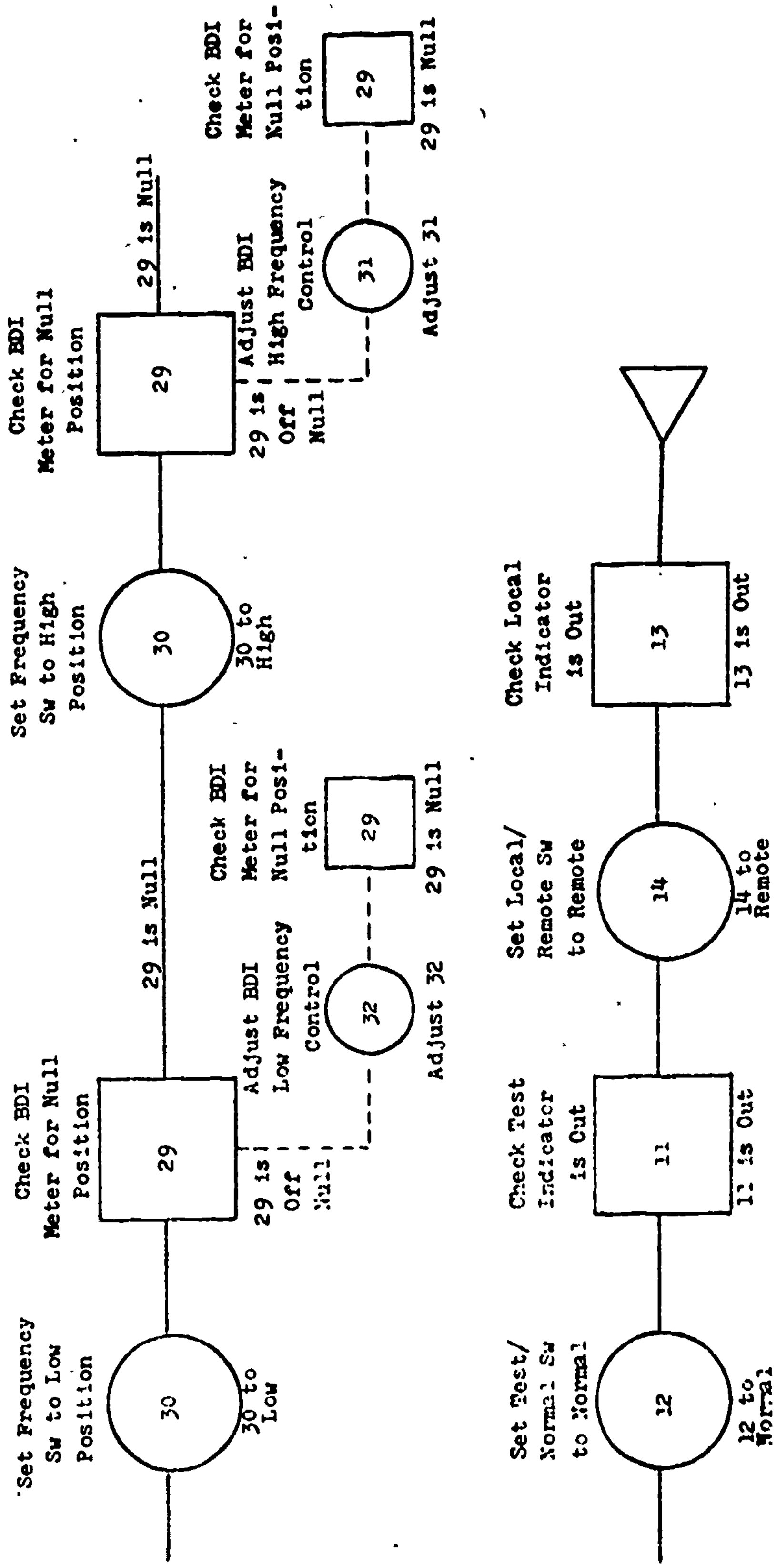
Device	DEI x10⁻⁴	Task indices D%	C%	E%	CRPS	FBR	INFO	INST
21B55 OA1283	273.86	22	10	14	2	0	4	0
21B55 BQR-2B	49.93	12	8	10	2	0	4	0
21A39/2 BQR-2C	48.04	31	11	19	2	0	6	0
21A39/2 BQR-7B	123.40	19	12	15	4	0	8	0
14A2/C1 SQS-23B	91.51	10	8	9	7	0	8	0
X14A2 SQS-23 (TRAM)	18.75	22	21	22	7	5	9	0
14E3 SQS-4	3.60	33	20	25	7	3	9	0
14E14 SQS-4	342.33	15	4	8	1	0	4	0
SQS-26CX	16.50	14	6	10	4	3	13	0
14E10 AQS-13	26.70	22	17	18	4	1	4	0

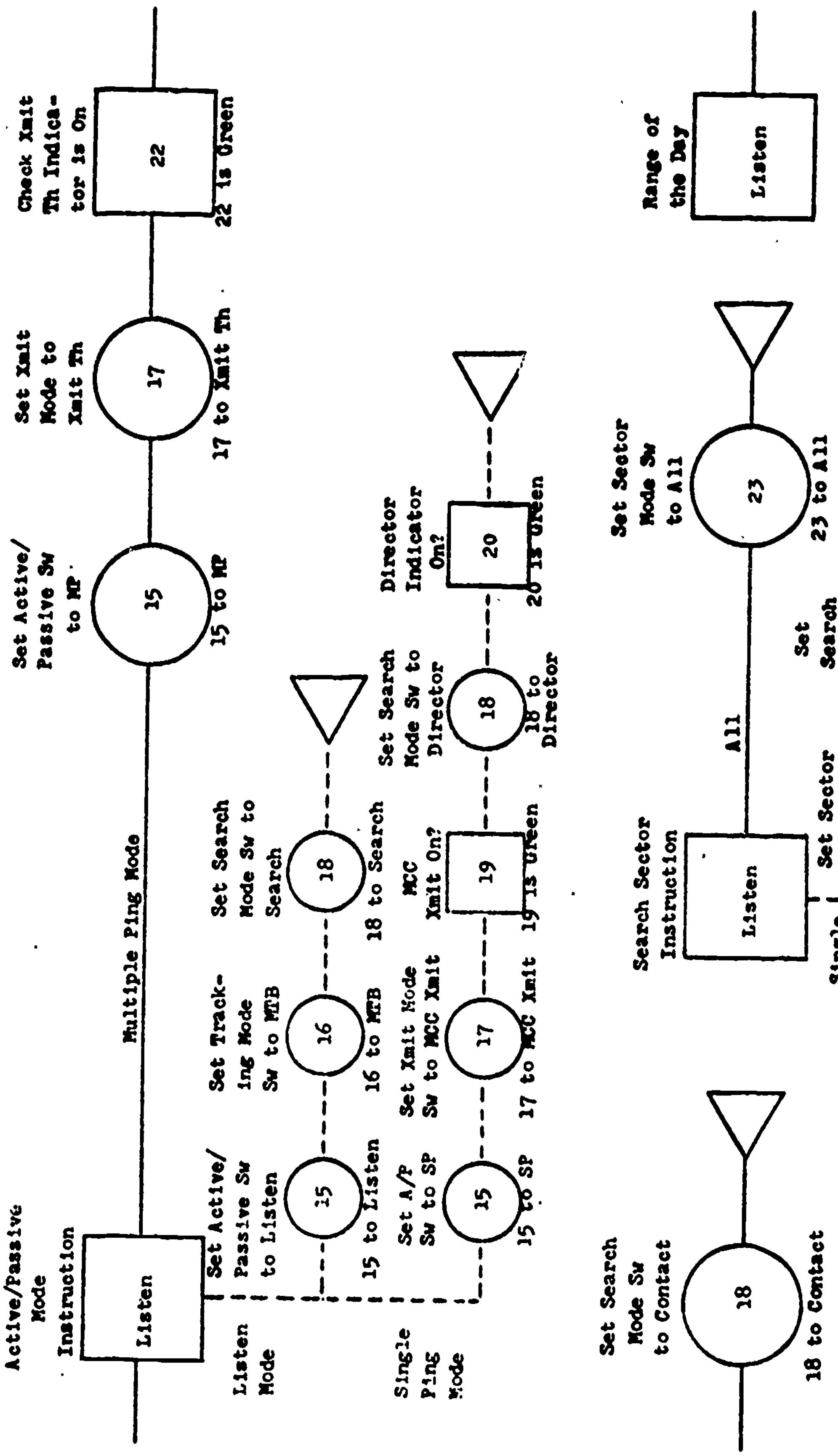
APPENDIX C

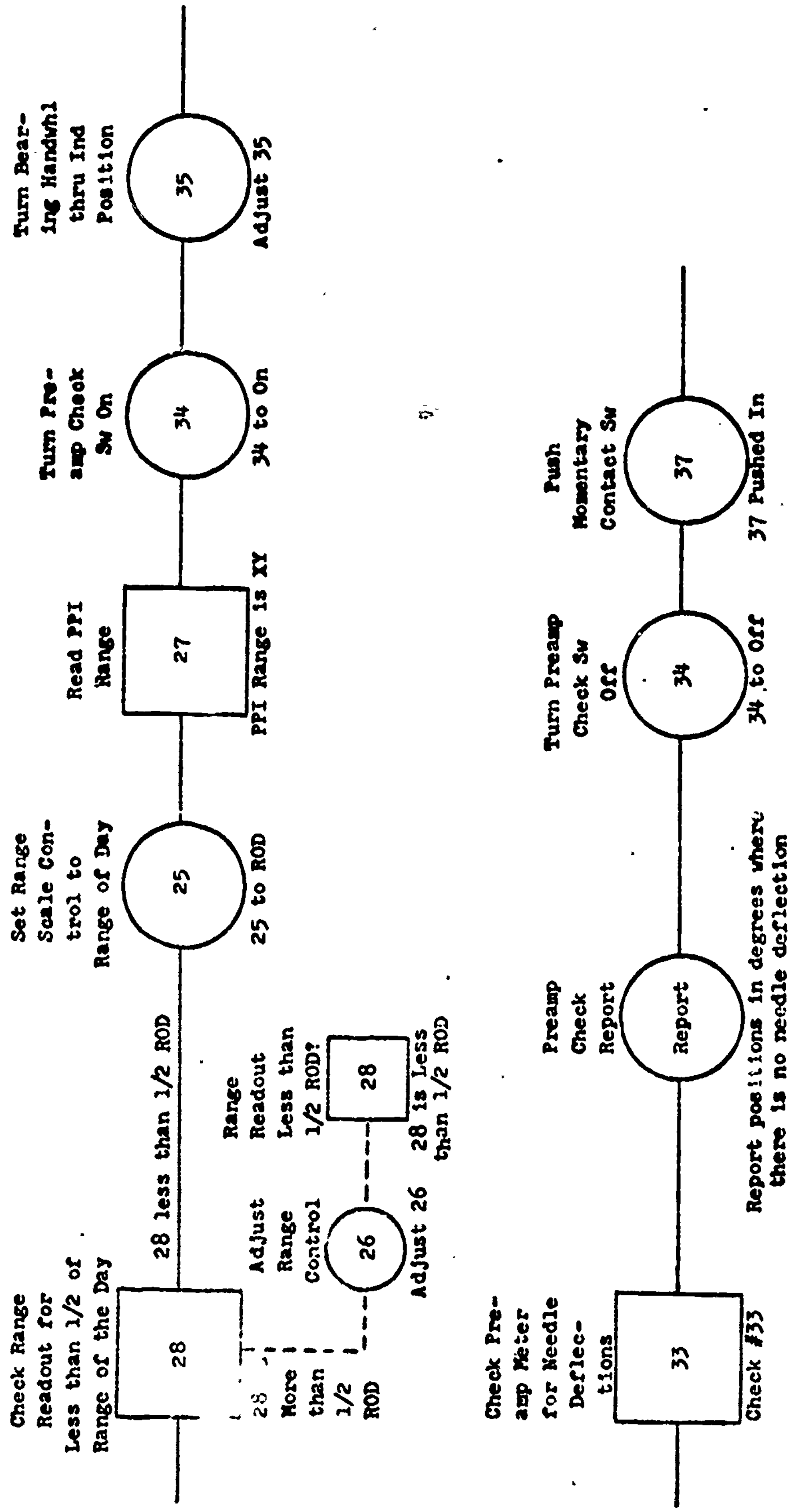
Operations Flow Chart for the Complex-All Synthetic Set-up Task

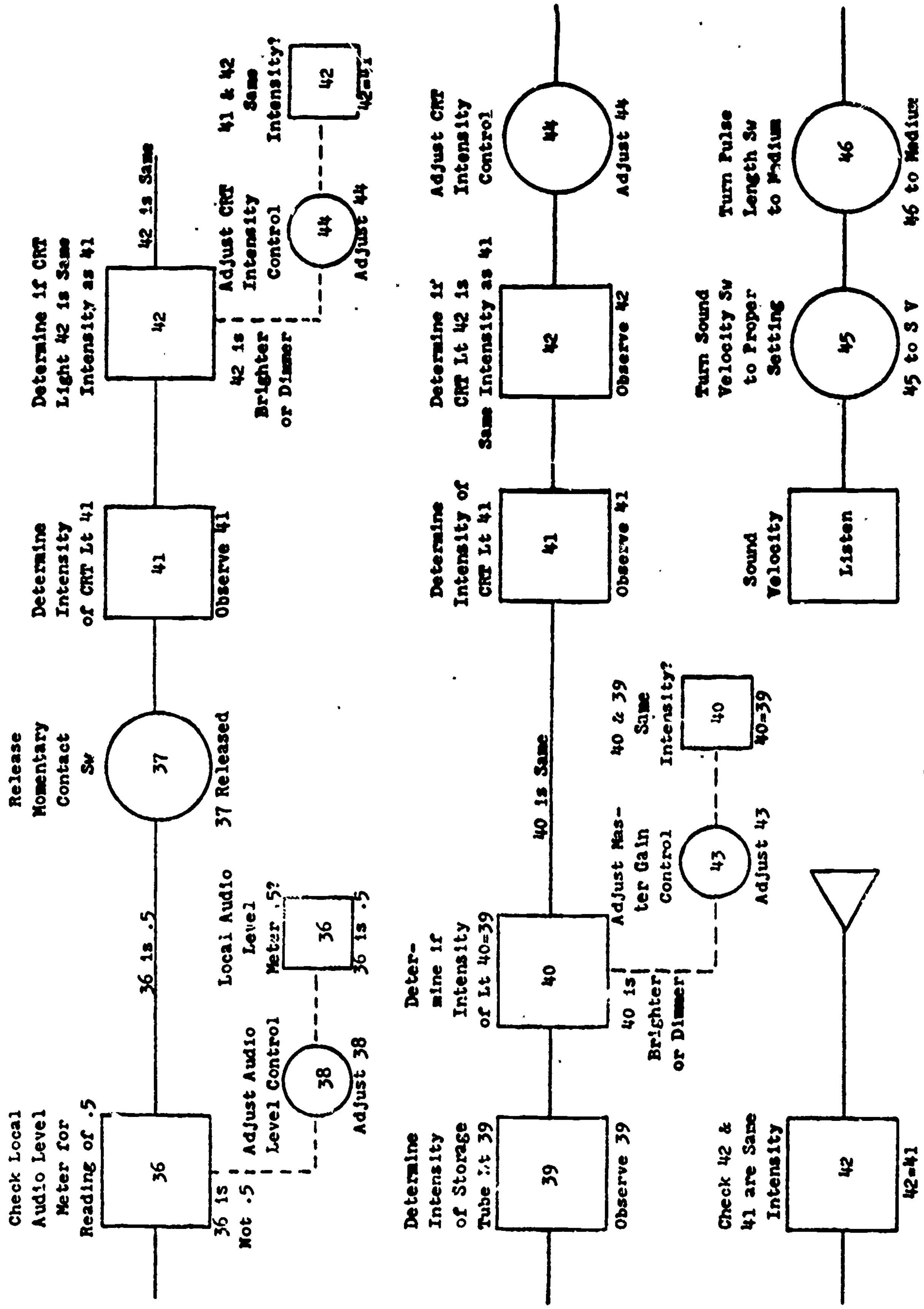

Turn Main Power Sw to Standby Position
1
1 to Standby
Check Main Power Sw Ind for On Indict'n
2
2 is Green
Turn Power Level Sw to Low Position
4
4 to Low
Turn Main Power Sw to On Position
1
1 to On
Check Power Level Meter for Reading of Less than 2
3
Less than 2 Indicated
More than 2 Indicated
Turn Main Power Sw to Standby
1
1 to Standby
Turn Pwr Level Sw to Medium
4
4 to Medium
Turn Main Pwr Sw to On
1
1 to On
Power Level Meter Less than 6?
3
3 is less than 6
Turn Recorder Power Sw to On Position
6
6 to On
Check Recorder Power Sw Ind for On Ind
5
5 is Green

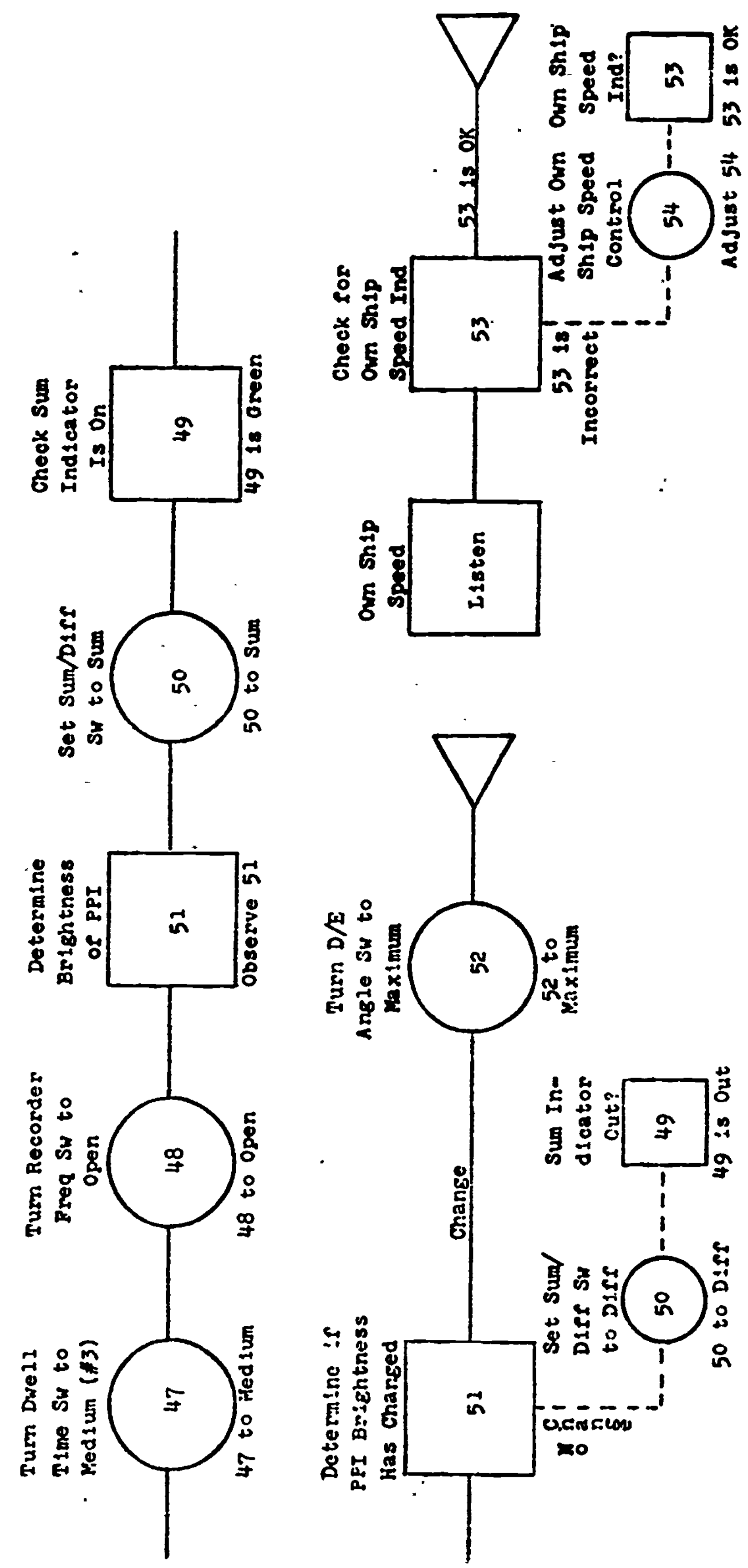

True/ Relative Mode
Listen
True
Relative
Check Relative Indicator is On
7
7 is Green
7 Out
Set True/ Rel Sw to Relative
8
8 to Relative
Relative Ind On?
7 is Green
Check Relative Ind is Out
7 Green
7 is Out
Set True/ Rel Sw to True
8 to True
Relative Ind Out?
7 is Out
Check Gyro Off Ind is Out
9
9 Green
Turn Gyro Off Sw to Off
10
10 to Off
Gyro Off Ind Out?
9 is Out
Check Gyro Off Indicator is On
9 Out
9 is Green
Turn Gyro Off Sw to On
10 to On
Gyro Off Ind On?
9 is Green
Check Test Indicator for On Indication
11
11 Out
11 is Green
Set Test/ Normal Sw to Test
12
12 to Test
Check Test Ind for On Ind
11 is Green
Check Local Indicator for On Indication
13
13 Out
13 is Green
Set Local/ Remote Sw to Local
14
14 to Local
Check Local Ind For On Ind
13 is Green

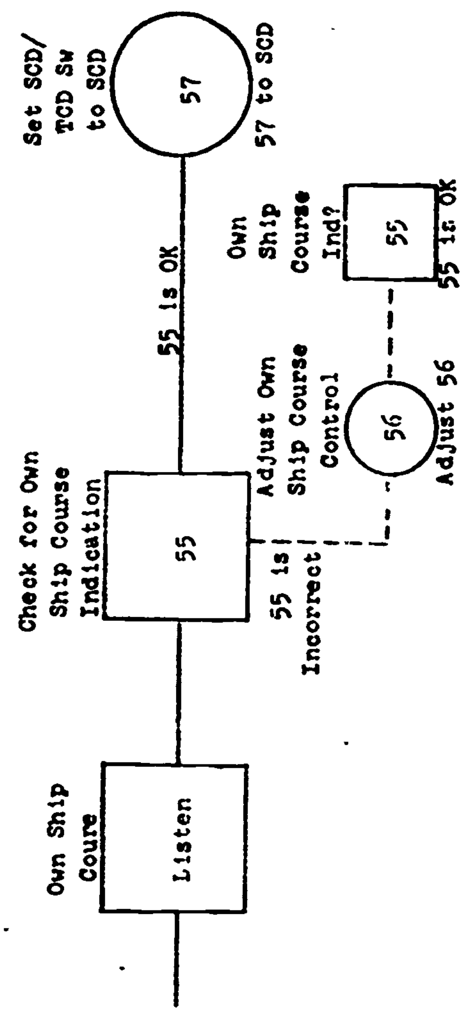

Set Frequency Sw to Low Position
30
30 to Low
Check BDI Meter for Null Position
29
29 is Off Null
Adjust BDI Low Frequency Control
32
Adjust 32
Check BDI Meter for Null Position
29
29 is Null
29 is Null
Set Frequency Sw to High Position
30
30 to High
Check BDI Meter for Null Position
29
29 is Off Null
Adjust BDI High Frequency Control
31
Adjust 31
Check BDI Meter for Null Position
29
29 is Null
29 is Null
Set Test/Normal Sw to Normal
12
12 to Normal
Check Test Indicator is Out
11
11 is Out
Set Local/Remote Sw to Remote
14
14 to Remote
Check Local Indicator is Out
13
13 is Out


Active/Passive Mode Instruction
Listen
Multiple Ping Mode
Set Active/Passive Sw to MP
15
15 to MP
Set Xmit Mode to Xmit Th
17
17 to Xmit Th
Check Xmit Th Indicator is On
22
22 is Green
Listen Mode
Set Active/Passive Sw to Listen
15
15 to Listen
Set Tracking Mode Sw to MTB
16
16 to MTB
Set Search Mode Sw to Search
18
18 to Search
Single Ping Mode
Set A/P Sw to SP
15
15 to SP
Set Xmit Mode Sw to MCC Xmit
17
17 to MCC Xmit
MCC Xmit On?
19
19 is Green
Set Search Mode Sw to Director
18
18 to Director
Director Indicator On?
20
20 is Green
Set Search Mode Sw to Contact
18
18 to Contact
Search Sector Instruction
Listen
All
Set Sector Mode Sw to All
23
23 to All
Range of the Day
Listen
Single Sector of X°
Set Sector Mode Sw to Single Sector
23
23 to Single Sector
Set Search Sector Sw to X°
24
24 to X°


Check Range Readout for Less than 1/2 of Range of the Day
28
28 less than 1/2 ROD
28 More than 1/2 ROD
Adjust Range Control
26
Adjust 26
Range Readout Less than 1/2 ROD?
28
28 is Less than 1/2 ROD
Set Range Scale Control to Range of Day
25
25 to ROD
Read PPI Range
27
PPI Range is XY
Turn Preamp Check Sw On
34
34 to On
Turn Bearing Handwhl thru Ind Position
35
Adjust 35
Check Preamp Meter for Needle Deflections
33
Check #33
Preamp Check Report
Report
Report positions in degrees where there is no needle deflection
Turn Preamp Check Sw Off
34
34 to Off
Push Momentary Contact Sw
37
37 Pushed In


Check Local Audio Level Meter for Reading of .5
36
36 is Not .5
Adjust Audio Level Control
38
Adjust 38
Local Audio Level Meter .5?
36
36 is .5
36 is .5
Release Momentary Contact Sw
37
37 Released
Determine Intensity of CRT Lt 41
41
Observe 41
Determine if CRT Light 42 is Same Intensity as 41
42
42 is Brighter or Dimmer
Adjust CRT Intensity Control
44
Adjust 44
41 & 42 Same Intensity?
42
42 is Same
Determine Intensity of Storage Tube Lt 39
39
Observe 39
Determine if Intensity of Lt 40=39
40
40 is Brighter or Dimmer
Adjust Master Gain Control
43
Adjust 43
40 & 39 Same Intensity?
40
40=39
40 is Same
Determine Intensity of CRT Lt 41
41
Observe 41
Determine if CRT Lt 42 is Same Intensity as 41
42
Observe 42
Adjust CRT Intensity Control
44
Adjust 44
Check 42 & 41 are Same Intensity
42
42=41
Sound Velocity
Listen
Turn Sound Velocity Sw to Proper Setting
45
45 to S V
Turn Pulse Length Sw to Medium
46
46 to Medium


Turn Dwell Time Sw to Medium (#3)
47
47 to Medium
Turn Recorder Freq Sw to Open
48
48 to Open
Determine Brightness of PPI
51
Observe 51
Set Sum/Diff Sw to Sum
50
50 to Sum
Check Sum Indicator Is On
49
49 is Green
Determine if PPI Brightness Has Changed
51
No Change
Change
Set Sum/Diff Sw to Diff
50
50 to Diff
Sum Indicator Out?
49
49 is Out
Turn D/E Angle Sw to Maximum
52
52 to Maximum
Own Ship Speed
Listen
Check for Own Ship Speed Ind
53
53 is Incorrect
53 is OK
Adjust Own Ship Speed Control
54
Adjust 54
Own Ship Speed Ind?
53
53 is OK


Own Ship Coure
Listen
Check for Own Ship Course Indication
55
55 is Incorrect
Adjust Own Ship Course Control
56
Adjust 56
Own Ship Course Ind?
55
55 is OK
55 is OK
Set SCD/TCD Sw to SCD
57
57 to SCD

APPENDIX D

Instructions for Magnitude Estimates

MAGNITUDE ESTIMATES

What we would like you to do today is to make four judgments about each sonar subtask (the comparison task) which is performed on your device. The four types of judgments are described below, together with some practice examples.

1. Relative to the case of the standard sonar subtask, how many more or how many fewer units of practice would the average A-school trainee need on the comparison task in order to perform it as quickly and as accurately as the typical instructor?

[In order to do this in the case of the standard task, 100 units of practice are required.]

------

2. Relative to the case of the standard sonar subtask on which 2 hours of practice was given, how much better or worse would the average A-school trainee perform the comparison task after the same amount of practice?

[Performance on the standard task after 2 hours of practice is at the 50 level.]

------

3. If the degree of transfer of training from the training situation to the operational situation is 50 on the standard task, how much greater or less than 50 is it on the comparison task?

[The degree of transfer of training on the standard task is 50.]

------

4. Relative to the case of the standard sonar subtask, how much more or less difficult would it be for the average A-school trainee to learn to perform the comparison task?

[The difficulty in learning to perform the standard subtask is 100.]

Below are four practice examples. Please complete them now.

1. With respect to the first type of estimate described above, if you thought your task required 2.5 times as much practice as the standard, what value would you assign?

__________.

2. With respect to the second type of estimate, if you thought trainees would perform only one-third as well, what value would you assign?

__________.

3. In the third type of estimate, what value would you assign if you thought the degree of transfer of training on the comparison task was:

Twice as great relative to the standard ______ ?

Half as great relative to the standard ______ ?

At the same level as in the standard case ______?

4. In the final type of estimate, what value would you assign if you thought the comparison task was 1-1/2 times more difficult than the standard?

__________ .

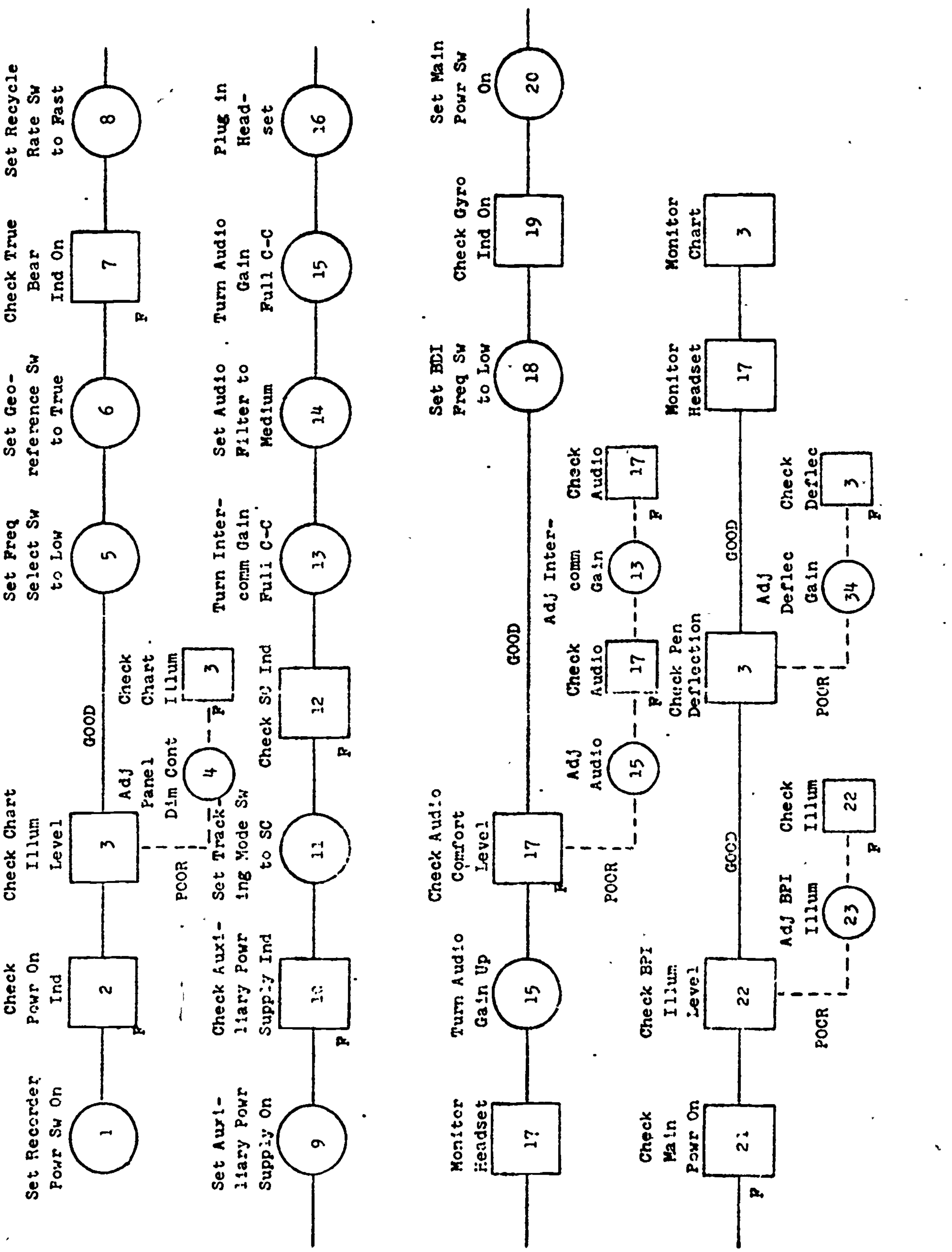
In making the estimates, remember:

- Think in terms of our task descriptions rather than in terms of how you do or teach the task.
- Make your judgments with respect to the overall task.
- Remember to assign a value to your judgment which is some fraction or multiple "times" the value of the standard.

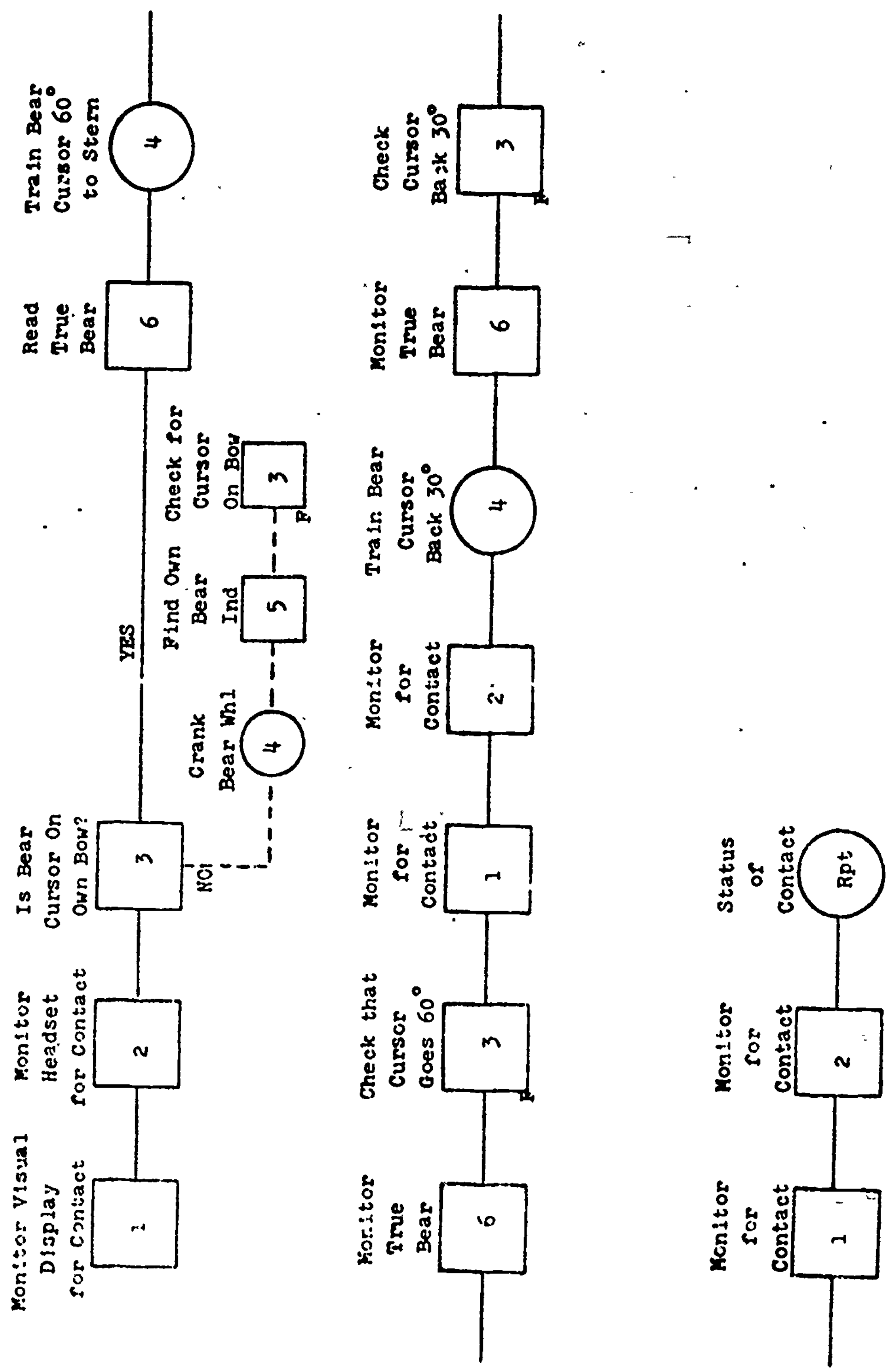
APPENDIX E

Subtask Standard Flow Charts

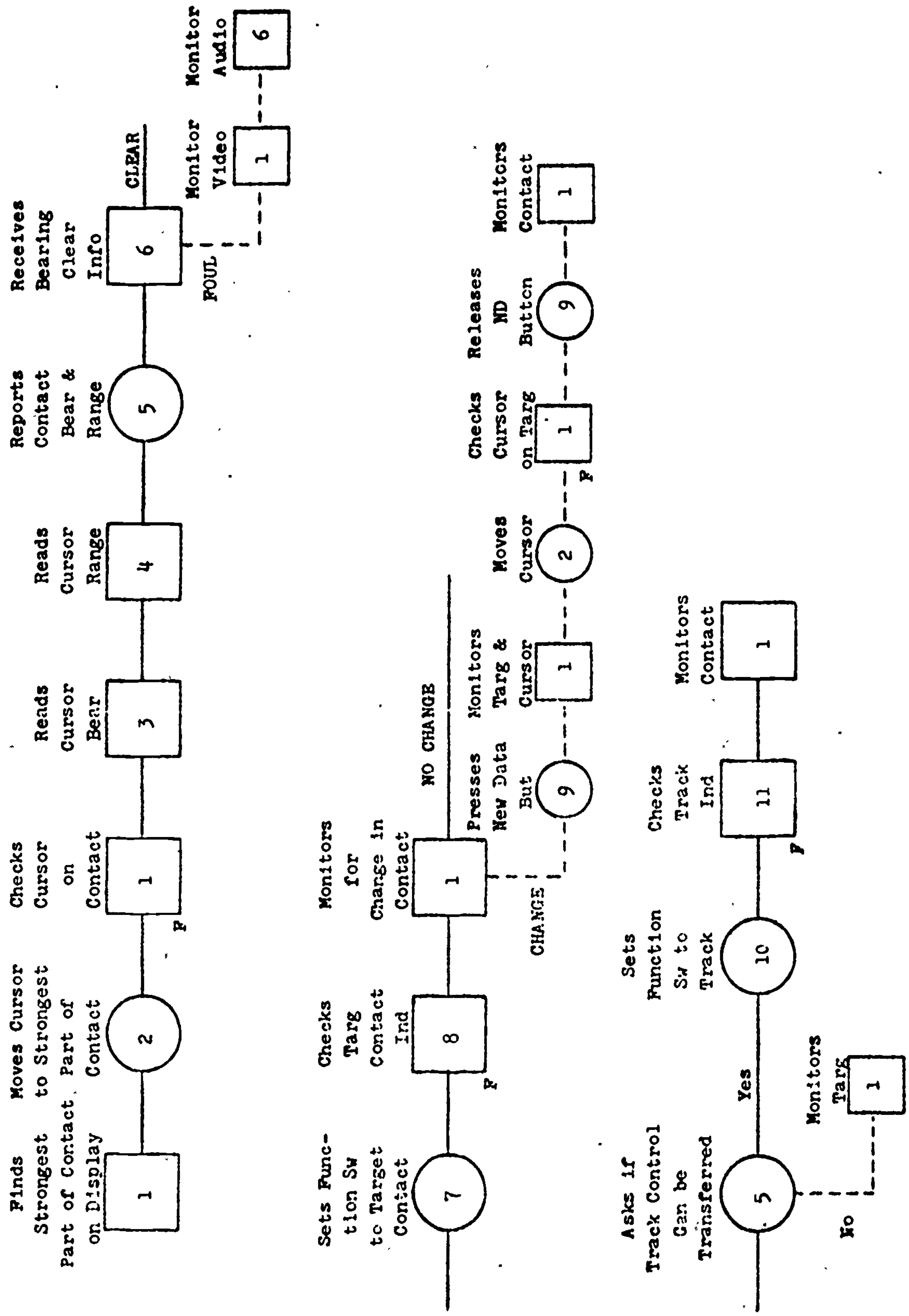
Sonar Set-Up Task Standard



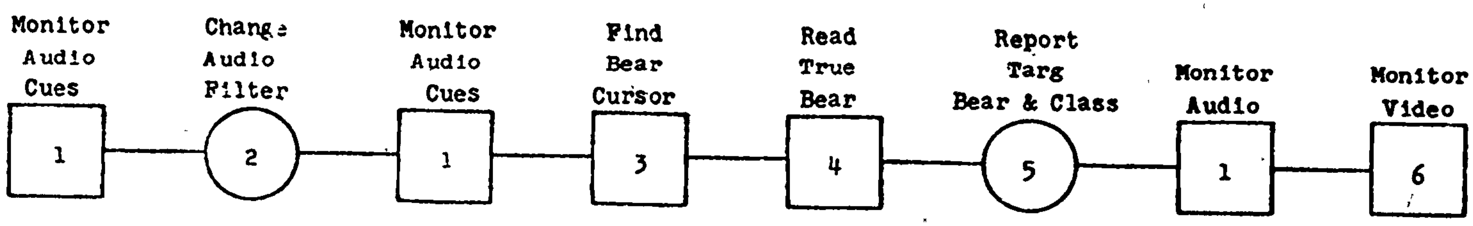
Sonar Detection Task Standard



Sonar Localization Task Standard



Sonar Classification Task Standard



APPENDIX F

Ratio Estimate Answer Sheet

Name________________________________ Task__________________

Type of Estimate	Standard Value	Your Estimate
1. Relative to the case of the standard sonar subtask, how many more or how many fewer units of practice would the average A-school trainee need on the comparison task in order to perform it as quickly and as accurately as the typical instructor?	100	
2. Relative to the case of the standard sonar subtask on which 2 hours of practice was given, how much better or worse would the average A-school trainee perform the comparison task after the same amount of practice?	50	
3. If the degree of transfer of training from the training situation to the operational situation is 50 on the standard task, how much greater or less than 50 is it on the comparison task?	50	
4. Relative to the case of the standard sonar subtask, how much more or less difficult would it be for the average A-school trainee to learn to perform the comparison task?	100	

APPENDIX G

Mean Instructor Ratio Estimates for the Four Subtasks

TABLE G-1. MEAN INSTRUCTOR RATIO ESTIMATES FOR THE SET-UP TASK

Device	Criteria			
	Training Time	Proficiency Level	Transfer	Task Difficulty
21B55 OA1283	136	42	43	126
21A39/2 CA1283	225	22	22	225
21B55 BQR-2B	87	57	60	91
21A39/2 BQR-2C	90	55	54	93
21A39/2 BQR-7B	195	20	20	170
14A2/C1 SQS-23B	50	94	75	50
X14A2 SQS-23 (TRAM)	79	63	61	79
14E3 SQS-4	143	31	40	136
14E14 SQS-4	125	38	44	113
SQS-26CX	390	25	27	260
14E10 AQS-13	172	46	46	152
14B31B AQA-1	231	33	27	156
14B31B ASA-20	333	20	23	233
Composite weights:	1.00	1.00		

TABLE G-2. MEAN INSTRUCTOR RATIO ESTIMATES FOR THE DETECTION TASK

Device	Criteria			
	Training Time	Proficiency Level	Transfer	Task Difficulty
21B55 0A1283	220	32	32	158
21B55 BQR-2B	144	[illegible]	35	163
21A39/2 BQR-2C	100	50	50	100
21A39/2 BQR-7B	100	50	50	100
14A2/C1 SQS-23B	125	45	50	119
X14A2 SQS-23 (TRAM)	129	34	29	129
14E3 SQS-4	44	125	90	63
14E14 SQS-4	125	45	45	125
SQS-26CX	70	90	60	90
14E10 AQS-13	170	31	43	140
Composite weights:	1.49	1.32		

TABLE G-3. MEAN INSTRUCTOR RATIO ESTIMATES FOR THE LOCALIZATION TASK

Device	Criteria			
	Training Time	Proficiency Level	Transfer	Task Difficulty
21B55 OA1283	200	36	36	180
21B55 BQR-2B	410	24	31	299
21A39/2 BQR-2C	133	35	39	130
21A39/2 BQR-7B	76	61	80	71
14A2/C1 SQS-23B	363	19	35	438
X14A2 SQS-23 (TRAM)	85	70	61	83
14E3 SQS-4	49	150	119	45
14E14 SQS-4	86	70	88	81
SQS-26CX	110	55	55	105
14E10 AQS-13	92	52	50	94
14B31B AQA-1	156	29	29	150
14B31B ASA-20	267	22	23	217
Composite weights:	1.13	1.22		

TABLE G-4. MEAN INSTRUCTOR RATIO ESTIMATES FOR THE CLASSIFICATION TASK

Device	Criteria			
	Training Time	Proficiency Level	Transfer	Task Difficulty
21B55 OA1283	59	91	100	66
21B55 BQR-2B	87	52	55	90
21A39/2 BQR-2C	100	50	50	100
21A39/2 BQR-7B	100	50	50	100
14A2/C1 SQS-23B	191	31	34	166
X14A2 SQS-23 (TRAM)	225	21	21	194
14E3 SQS-4	212	28	29	300
14E14 SQS-4	200	33	54	175
SQS-26CX	170	41	27	115
14E10 AQS-13	110	50	50	100
Composite weights:	1.81	2.01		

APPENDIX H

Intercorrelations of Task Index Values and Adjusted Mean Instructor Ratio Estimates

TABLE H-1. INTERCORRELATIONS OF TASK INDEX VALUES AND MEAN INSTRUCTOR RATIO ESTIMATES—SET-UP†

Task Indices	Criteria: Training Time	Proficiency Level	Transfer	Task Difficulty
MAIN	56*	-43	-50	54
CNTG	43	-46	-53	34
TA	69**	-57*	-66*	64*
CONT	72**	-30	-39	64*
DISP	-24	-09	01	-19
E	60*	-47	-46	56*
LV	67*	-62*	-69**	-61*
AA%	-76**	70**	78**	-77**
F%	-35	50	54	-42
DEI x10^{-4}	-52	25	32	-45
D%	-19	-36	-29	-08
C%	39	-35	-35	48
E%	12	-50	-44	28
CRPS	61*	-44	-52	55*
FBR	68*	-60*	-69**	66*
INFO	56*	-65*	-70**	47
INST	-16	36	35	-21

*$p < .05$
**$p < .01$

†Decimal points have been omitted for clarity.

TABLE H-2. INTERCORRELATIONS OF TASK INDEX VALUES AND MEAN INSTRUCTOR RATIO ESTIMATES—DETECTION†

Task Indices	Criteria			
	Training Time	Proficiency Level	Transfer	Task Difficulty
MAIN	57	-65*	-76**	36
CNTG	52	-69*	-68*	63*
TA	63*	-77**	-84**	53
CONT	66*	-60*	-69*	52
DISP	59*	-72**	-71*	37
E	71*	-73**	-78**	52
IV	63*	-75**	-84**	50
AA%	-07	10	-03	-14
F%	-10	06	09	-06
DEI x10^{-4}	01	11	08	-00
D%	21	-49	-58*	10
C%	63*	-54	-66*	48
E%	56	-61*	-72**	37
CRPS	74**	-74**	-84**	64*
FBR	56	-70*	-79**	48
INFO	48	-72**	-75**	38
INST	08	-14	-19	06

*$p < .05$
**$p < .01$

†Decimal points have been omitted for clarity

TABLE H-3. INTERCORRELATIONS OF TASK INDEX VALUES AND MEAN INSTRUCTOR RATIO ESTIMATES—LOCALIZATION†

Task Indices	Criteria			
	Training Time	Proficiency Level	Transfer	Task Difficulty
MAIN	57	-76*	-79**	75*
CNTG	33	-26	-06	18
TA	72*	-89**	-86**	85**
CONT	40	-28	-18	27
DISP	-04	-28	-18	11
E	48	-65*	-42	47
LV	79**	-79**	-80**	89**
AA%	54	-27	-32	49
[illegible]	00	00	00	00
DE1 x10-4	-16	36	13	-19
D%	45	-56	-47	42
C%	39	-13	-05	19
E%	67*	-44	-28	44
CRPS	97**	-74*	-71*	85**
FBR	43	-67*	-60	60
INFO	43	-78**	-76*	66*
INST	72*	-29	-38	53

*p .05
**p .01

†Decimal points have been omitted for clarity.

TABLE H-4. INTERCORRELATIONS OF TASK INDEX VALUES AND MEAN INSTRUCTOR RATIO ESTIMATES—CLASSIFICATION†

Task Indices	Criteria			
	Training Time	Proficiency Level	Transfer	Task Difficulty
MAIN	43	-44	-68*	38
CNTG	74*	-63*	-71*	62
TA	68*	-62	-79**	58
CONT	35	-39	-46	41
DISP	38	-41	-59	40
E	42	-46	-60	47
LV	65*	-58	-75*	48
AA%	30	-16	-12	10
F%	-40	29	08	-10
DEI $\times 10^{-4}$	-16	40	67*	-21
D%	-02	04	00	35
C%	24	-27	-33	47
E%	17	-19	-26	47
CRPS	63*	-60	-69*	65*
FBR	65*	-54	-66*	55
INFO	52	-48	-71*	36
INST	00	00	00	00

*p .05
**p .01

†Decimal points have been omitted for clarity.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
American Institutes for Research
8555 Sixteenth Street
Silver Spring, Maryland 20910

2a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2b. GROUP

3. REPORT TITLE

EFFECTS OF TASK INDEX VARIATIONS ON TRAINING EFFECTIVENESS CRITERIA

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Final Report

5. AUTHOR(S) *(First name, middle initial, last name)*

George R. Wheaton and Angelo Mirabella

6. REPORT DATE
October 1972

7a. TOTAL NO. OF PAGES
87

7b. NO. OF REFS
3

8a. CONTRACT OR GRANT NO.
N61339-71-C-0059

b. PROJECT NO.
NAVTRAEQUIPCEN 1752-02

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

Approved for public release; distribution is unlimited.

11. SUPPLEMENTARY NOTES

None

12. SPONSORING MILITARY ACTIVITY

Naval Training Equipment Center
Orlando, Florida

13. ABSTRACT

A feasibility study was undertaken as part of a program to develop quantitative techniques for prescribing the design and use of training systems. As the second step in this program, the present study attempted to: (1) refine quantitative indices employed during earlier research; (2) conduct laboratory research on the effects which task index variations have on training criteria; and (3) support the laboratory results with data gathered in the field. Two laboratory investigations and a field study were conducted. In the first laboratory study, effects of variations in task indices on skill acquisition of a set-up task were examined. In a companion effort, preliminary data were collected on relationships between task index variations and performance during transfer of training. In the field study quantitative task index data, descriptive of a variety of sonar trainers and sonar trainee tasks, were related to ratio estimates provided by instructors on four training effectiveness criteria. Significant multiple correlations were obtained between task indices and speed and accuracy of performance during skill acquisition. Predictor patterns changed over time and between criteria. Set-up task speed was predicted early in training, while errors made were predicted later during acquisition. Similar but more provisional relationships were found during transfer of training. Speed and, in particular, accuracy of performance during transfer bore consistent relationships to task index values. Support for these general findings was obtained in the field. Significant relationships were established between instructors' judgments of training criteria and trainee subtask index values.

DD FORM 1 NOV 65 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

UNCLASSIFIED
Security Classification

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Task Analysis Quantitative Task Indices Training Effectiveness						

UNCLASSIFIED
Security Classification

DISTRIBUTION LIST

NOTE

Mailing labels are prepared, when needed, as a computer listing, the source of which is updated on a weekly basis. It is not practical to prepare distribution lists each time labels are prepared. Therefore, such lists are prepared semiannually, and a slight discrepancy may exist between the addresses on this list and those appearing on the labels used to distribute this publication.

USAF HUMAN RES LAB
AFHRL-DO OPERATION OFC
BROOKS AFB TX 78235
1

USAF HUMAN RES LAB
AFHRL-TT
TECH TRNG DIV
LOWRY AFB CO 80230
1

USAF HUMAN RES LAB
PERSONNEL RSCH DIV
LACKLAND AFB TX 78236
1

USAF HUMAN RES LAB
AFHRL-TR
TRNG RES DIV
WRIGHT-PATTERSON AFB
OH 45433
1

CHIEF NAVAL RESEARCH
PSYCHOLOGICAL SCIENCES
CODE 450 NAVY DEPT
ARLINGTON VA 22217
1

DR RALPH R CANTER
DIR MIL MANPWR RSCH
OASD M-RA MR-U
PENTAGON RM 3-960
WASHINGTON DC 20301
1

NATL SCIENCE FDN
ATTN DR HENRY S ODBERT
1800 G ST NW
WASHINGTON DC 20550
1

CHIEF NAVAL OPERATIONS
ATTN DR J J COLLINS
OP-07T16 NAVY DEPT
WASHINGTON DC 20350
1

DR JOHN MEYER
HDQTRS AIR TRNG CMD XPT
RANDOLPH AFB TX 78148
1

USAF HUMAN RES LAB
AFHRL-FT
FLYING TRNG DIV
WILLIAMS AFB AZ 85224
1

CHIEF NAVAL MATERIAL
MAT 031M
WASHINGTON DC 20360
1

JOSEPH J COWAN
CH PERS RSCH BRANCH
USCG HQ PO-1 STA 3-12
400 SEVENTH ST S W
WASHINGTON DC 20226
1

BUREAU NAVAL PERSONNEL
ATTN PERS A3
ARLINGTON ANNEX
WASHINGTON DC 20370
1

DEPT OF PSYCHOLOGY
UNIV OF IL AT
URBANA-CHAMPAIGN
PROF PSYCH J A ADAMS
CHAMPAIGN IL 61820
1

CHIEF NAVAL OPERATIONS
ATTN CDR H J CONNERY
MSC OP-701E2 NAVY DEPT
WASHINGTON DC 20350
1

INSTITUTE OF AVIATION
UNIV OF IL
URBANA-CHAMPAIGN
SAVOY IL 61874
1

CHIEF NAV OPERATIONS 1
ATTN M K MALEHORN
OP-14C NAVY DEPT
WASHINGTON DC 20350

COMMANDER 1
NAVELECTSYSCMD
CODE 03
WASHINGTON DC 20360

NAV AEROSPACE MED INST 1
NAV AEROSPACE MED CTR
ATTN CH AVIA PSYCH DIV
PENSACOLA FL 32512

COMMANDER 1
NAVSUPSYSCMD
CODE 03
WASHINGTON DC 20360

COMMANDER 1
NAVORDSYSCMD
CODE 03
WASHINGTON DC 20360

HUMAN RES RSCH ORG 1
DIV NO 1 SYS OPER
300 N WASHINGTON ST
ALEXANDRIA VA 22314

CO NAV MISSILE CTR 1
ATTN HD HUMAN FACTORS
ENGRG BR
POINT MUGU CA 93042

CH RSCH AND DEVLPMT 1
CH BEHAVIORAL SCI DIV
DEPT OF ARMY
WASHINGTON DC 20310

COMMANDING OFFICER 1
NAVY MEDICAL
NEUROPSYCHIATRIC RESCH
UNIT
SAN DIEGO CA 92152

DIR DEF RSCH - ENGRG 1
ARPA
BEHAVIORAL SCI DIV
ATTN LCOL A W KIBLER
WASHINGTON DC 20301

COMMANDER 1
NAVSHIPSSYSCMD
CODE 03
WASHINGTON DC 20360

CDR NAVSHIPS SYS CMD 1
NAVSHIPS SYS CMD HQS
SHIPS 03H GEO N GRAINE
WASHINGTON DC 20360

COMMANDER 1
NAVAIRSYSCMD
CODE 03
WASHINGTON DC 20360

CO NAVMED RSCH INST 1
NATL NAVMED CTR
ATTN TECH REF LIB
BETHESDA MD 20014

SCIENTIFIC TECHNICAL 1
INFORMATION OFFICE
NASA
WASHINGTON DC 20546

BEHAVIORAL SCI DIV 1
OFC CHIEF RSCH - DEVEL
DEPT OF ARMY
WASHINGTON DC 20310

HUMAN RESOURCES RESCH 1
ORGZTN DIV 6 AVIATION
P O BOX 428
FORT RUCKER AL 36360

HQS AF SYS CMD DLSL 1
OFC SCIENTIFIC RSCH
ANDREWS AFB
WASHINGTON DC 20331

COMMANDER 1
NAVFAC ENGR CMD
NAVFAC ENGR CMD HDQ
ASST DIR R&D DIV(CODE 03)
WASHINGTON DC 20390

DIRECTOR 1
US ARMY MOTIVATION & TNG LAB
COMMONWEALTH BLDG
1300 WILSON BLVD
ARLINGTON VA 22209

COMDT OF MARINE CORPS 1
CODE AO3C
WASHINGTON DC 20380

CHIEF NAVAL RESEARCH 1
ATTN CODE 458 DEPT NAV
ARLINGTON VA 22217

COMMANDING OFFICER 1
PERS-TRNG RSCH DEV LAB
SAN DIEGO CA 92152

US ARMY BEHAVIORAL SCI 1
RESEARCH LAB
COMMONWEALTH BLDG
RM 239
1320 WILSON BLVD
ARLINGTON VA 22209

EXECUTIVE EDITOR 1
PSYCHO ABSTRACTS
AMERICAN PSYCH ASSOC
1200 17TH ST NW
WASHINGTON DC 20036

NAV PERSONNEL RSCH & 1
DEVELOPMENT LAB
ATTN LIBRARY
BLDG 200 RM 3307 WNY
WASHINGTON DC 20390

CHIEF OF NAVAL TECH TRNG 1
NAVAL AIR STATION, MEMPHIS (75)
MILLINGTON, TN 38054
(ATTN: DR. LARRY HARDING,
CODE 34)

Commanding Officer 77
Naval Training Equipment Center
Orlando, Florida 32813

Defense Documentation Center 12
Cameron Station
Alexandria, Va 22314

DIR HUMAN RES RSCH ORG 1
300 WASHINGTON ST.
ALEXANDRIA VA 22314

CHIEF OF NAVAL TRNG 1
CODE 017 NAS
PENSACOLA FL 32508

CDR NAVAIR DEVEL CTR 1
ATTN HUMAN ENGRG BR
WARMINISTER PA 18974

DR JOHN W WEISZ 1
DIR HUMAN ENGRG LAB
USA ABERDEEN RSCH
DEVEL CTR
ABERDEEN PROV GROUNDS
MD 21005

CO NAVAIR TECH TRNG 1
NAS MEMPHIS
ATTN DR G D MAYO
HD RESEARCH BR
MILLINGTON TN 38054

ERIC CLEARINGHOUSE ON 1
EDUCATION MEDIA--TECH
STANFORD UNIV
STANFORD CA 94305