AD-A064 297 APPLIED PSYCHOLOGICAL SERVICES INC WAYNE PA SCIENCE --ETC F/G 5/10
OPERATIONAL DECISION AID EVALUATIONS.(U)
JUL 78 A I SIEGEL N00014-77-C-0448

UNCLASSIFIED NL

1 OF 1

AD
A064 297

END
DATE
FILMED
4 -79
DDC

LEVEL II

12

AD A064297

DDC FILE COPY

ANNUAL REPORT

Operational Decision Aid Evaluations

Arthur I. Siegel

prepared by

Applied Psychological Services, Inc.
Science Center
Wayne, Pennsylvania

for

Engineering Psychology Programs
Office of Naval Research
Washington, D. C.

under

Contract N00014-77-C-0448

July 1978

ANNUAL REPORT

Operational Decision Aid Evaluations

Arthur I. Siegel

prepared by

Applied Psychological Services, Inc.
Science Center
Wayne, Pennsylvania

for

Engineering Psychology Programs
Office of Naval Research
Washington, D. C.

under

Contract N00014-77-C-0448

July 1978

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (9)
4. TITLE (and Subtitle) (6) Operational Decision Aid Evaluations - Annual Report .		5. TYPE OF REPORT & PERIOD COVERED Annual rept. July 1977 - July 1978,
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Arthur I. Siegel		8. CONTRACT OR GRANT NUMBER(s) (15) N00014-77-C-0448
9. PERFORMING ORGANIZATION NAME AND ADDRESS Applied Psychological Services, Inc. 404 E. Lancaster Avenue Wayne, Pennsylvania 19087		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS (11) Engineering Psychology Programs Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE July 1978
		13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) (12) 18p.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This report has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Decision Aids	System Evaluation	Information Systems
Evaluative Research	Information Management	Performance Criteria
Man-Computer Interaction	Information Processing	Interface Design

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This Annual Report describes the work accomplished by Applied Psychological Services during the initial year of a program designed to evaluate a set of operational decision aids for Navy tactical planners. The work completed includes development of evaluative strategies, design of initial evaluative studies, and development of methods for testing specific interface features. Other support research is also discussed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

402 774

UNCLASSIFIED LB

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT

This Annual Report describes the work accomplished by Applied Psychological Services during the initial year of a program designed to evaluate a set of operational decision aids for Navy tactical planners. The work completed includes development of evaluative strategies, design of initial evaluative studies, and development of methods for testing specific interface features. Other support research is also discussed.

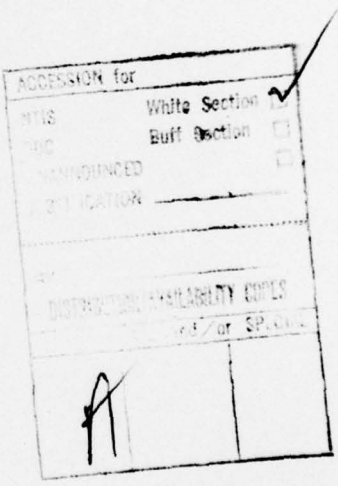

TABLE OF CONTENTS

I. INTRODUCTION AND BACKGROUND

This Annual Report describes the work accomplished during the initial year under contract N00014-77-C-0448 between the Applied Psychological Services and the Office of Naval Research. The purpose of the work undertaken by Applied Psychological Services is to provide methodological support to the operational decision aiding (ODA) program of the Office of Naval Research. The major task is to plan, develop, and implement procedures for evaluating ODA systems. The overall purpose is to collect systematically and analyze information about the usefulness of a variety of decision aids for Navy personnel.

ODA: The Operational Decision Aiding Program

The ODA program, as understood by Applied Psychological Services, represents a comprehensive effort to design, develop, and evaluate computer based decision aids for Naval operation purposes. There are two major emphases in the ODA program. The first entails actual development of demonstration decision aiding systems. Taken together, the developed and developing decision aids are man-computer interactive systems which take advantage of the most current advances in the behavioral, the management, the computer, and the information processing technologies/sciences.

The second emphasis of the ODA program, the one with which the Applied Psychological Services is concerned, is rigorous evaluation of the decision aids. To this end, the Applied Psychological Services program has taken steps in three directions. One direction was organization and implementation of a literature analysis and a set of working meetings to: (1) determine a unifying evaluation philosophy, (2) clarify criterion problems, (3) clarify evaluative procedures, and (4) organize properly the test bed. The second direction was to establish specific methods for test of one of the aids developed under the ODA program. The third thrust was the evaluation of the effectiveness and applicability of specific operator interface features including display characteristics, the use of the decision aiding procedures, and the techniques employed by the aids.

At this stage in the ODA evaluative program, the experimental and laboratory tests of the ODAs is entering an operational stage. The developers of the decision aids have conducted and reported several informal aid evaluative studies, but few formal experiments have been completed either by the decision aid developers themselves or independent researchers.

The Need

In an earlier review of the ODA program, Sinaiko (1977) referred to stages of decision aid development and clearly pointed out the importance of the experimental evaluation aspect in the ODA program. He wrote:

> The underlying rationale for all ODA work has been the assumption that all products of this program would be subjected to experimental test. Tests were to be undertaken at first by the various contractors in their own facilities, and later at a designated test bed that would serve the entire program. Ultimately, as various decision aids or other products moved toward fleet use, they would be tested in operational settings ...(p.1).

Applied Psychological Services is identified with the experimental evaluation aspect of the program. As an independent, private research organization with no proprietary interests in any of the ODA decision aids, Applied Psychological Services has designed, developed, and attempted to implement experimental test plans for the empirical evaluation of the decision aids.

As the first hurdle in the Applied Psychological Services' effort, the variables and conditions; that is, the dependent and independent factors to place under study were analyzed. Thus hurdle presented itself as no easy obstacle as the list of possible factors to investigate is lengthy and time and resources are limited. Dr. James H. Carlisle at the Annenberg School of Communications, University of Southern California, as a member of the ODA evaluation research team, provided a general framework for person-machine research and more specifically decision aiding system evaluation (Carlisle, 1978). His work listed, described, and provided operational definitions of the dependent and independent factors relevant to research interests in man-computer interaction (MCI) systems. Carlisle listed seven dependent variables and seven independent variables; he called them characteristics of performance and process and entities of man-computer interaction, respectively.

Characteristics of Performance and Process

1. time to perform the task
2. cost to perform the task
3. quantity and quality of the performance
4. errors committed
5. user's satisfaction
6. utilization of available resources
7. patterns of user and system behavior

Entities of Man-Computer Interaction

1. system
2. data base
3. user-system interface
4. user
5. task
6. training
7. setting

Against the backdrops of: (1) such a framework of variables, (2) the needs of the ODA program, (3) the interests and capabilities of the Applied Psychological Services, and (4) the time and resources available, a set of dependent and independent variables was selected for initial consideration. Generally speaking, the Applied Psychological Services' ODA evaluation program seeks to answer the following basic and applied research questions relative to each aid under consideration:

- Does use of the overall decision aid improve decisions and decision making effectiveness?
- Does one or more features of the decision aid enhance decision making more than other features of the aid?
- What features of the aid need to be changed and/or improved?
- What features of the aid should be deleted and what features should be added to the aid?
- How "valid" is the aid?
- Is the aid acceptable to users?
- Which features of the aid have the most value or usefulness to users?
- Are there individual differences in performance using the aid and in the acceptability of the aid?
- Does the type and complexity of the decision problem affect performance with the aid?

Other Associations

From the beginning, the ODA evaluation program has been a partnership endeavor. In selecting research variables and designing experimental plans, the Applied Psychological Services, along with working with the Office of Naval Research, has also worked closely with the contractors who developed the decision aids, and with members of the Department of Decision Sciences at the University of Pennsylvania, who have developed the actual test bed to be employed.

The University of Pennsylvania Test Bed

In order to provide an integrated residence for the various decisions aids and in order to allow interaction between the aids themselves and between the aids and various data banks, the various aids are programmed and installed at the Department of Decision Sciences of the Wharton School, University of Pennsylvania. The test bed has two major purposes in the ODA program: (1) to provide a central meeting and demonstration site for the ODA contractors and their decision aids, and (2) to provide a central experimental site for the standardized and objective evaluation of the various ODA products.

For evaluation purposes, the University of Pennsylvania system essentially provides a laboratory: it provides the hardware, software, and space for training and processing subjects through experiments using various aspects of the available decision aids, and for collecting and analyzing data resulting from the experiments. In designing test plans and training material, a major consideration of the Applied Psychological Services was to develop rigorous evaluative methods which are compatible with the capacity and equipment at the University of Pennsylvania decision laboratory.

II. SPECIFIC WORK ACCOMPLISHED

Conceptual Development

The initial work under the present program focused on organizing and conceptualizing an ODA evaluation philosophy. This work was accomplished in coordination with members of the staff of the Department of Decision Sciences, Wharton School, University of Pennsylvania. Dr. James Carlisle of the University of California also participated in three working meetings. As the end result, Applied Psychological Services produced a working paper which attempted to place into perspective various evaluative research concepts and considerations (Applied Psychological Services, 1978).

In a review of related efforts, the working paper pointed out that:

> ...Rees (1967), King (1968), and Katter (1969), have attempted to integrate the literatures of various orientations. Katter (1969) suggested that design and evaluation activities are necessarily related, even though they are often performed by different groups of people and at different times with respect to any one system. Ideally, these activities can be integrated into a continuous system development process. Similarly, Martin and Parker (1971) contended that systematic experimentation and evaluation is valuable at all stages of design of any man-computer interactive system. Based on extensive experience in both design and evaluation of a large-scale library system at Stanford University, they argued that many design questions cannot be answered properly without an interactive process of design and testing. That is to say, user and task characteristics are important in addition to software and hardware variables in the determination of user and system behavior.
>
> In seeming contradiction to the potential benefits of studying the use of a man-computer interactive system as an integral part of the design process, relatively little systematic research has been carried out on the process and effectiveness of presently operating systems. The Spires/Ballots Project at Stanford, the TIP and INTREX projects at MIT, the Psych Abstracts Project at Syracuse, the Mead Data Central Project with the Ohio Bar Association, and the Index Medicus Project with SDC and the National Library of Medicine are notable exceptions (cf., Parker & Paisley, 1966, Marcus, Benenfield, and Kugel, 1971; Cook, 1970; Carlisle, 1970; and Katter and McCarn, 1971).

> Even in these research efforts, thus far, progress toward understanding the relationships between system, task and user behavior has been hampered, in part, by three problems. First, there has been no integrative framework from which potentially important variables could be defined. Because variables have been contrived *ad hoc* in many studies, few standard measures have emerged and been widely used. A second problem has been the strong reliance on either the contrived and rigidly controlled experiment or the questionnaire as means of collecting data. Little use has been made to date of the process of monitoring and utilization ststistics of actual system operation. Parker (1966), Cook (1970) and Gerrity (1971) emphasize the potential value of utilization monitoring, but this technique has not been widely used. A third problem is that research on the system use is often regarded by designers and programmers as threatening and contrary to their design goals. As evaluative research makes positive and major contributions to the on-going design process of MCI systems, this third problem should be greatly reduced.

The working paper attempted to decompose and clarify the evaluative problem by visualizing the operational decision aids to involve a complex set of interactive procedures for augmenting decision effectiveness. The purpose of an evaluation in this dynamic, machine-human interactive context is to state the effects of operator, interface, and system variables on a variety of system output measures which reflect decision quality. This conceptualization is shown schematically in Figure 1.

The variables shown in Figure 1 include items such as operator experience and intelligence level.

The interface variables include characteristics of the display and the information input/output subsystem.

The conditions of use variables shown in Figure 1 are thought to be fundamental to demonstrating a wide and realistic range of information extraction and decision making performance effects by such measures as response time, error rate, and quality of work completed.

Operating procedures include the use of aids along with the manipulation of perceptional and cognitive factors embedded within the aids. These factors are also manipulated relative to both information extraction and decision making conditions.

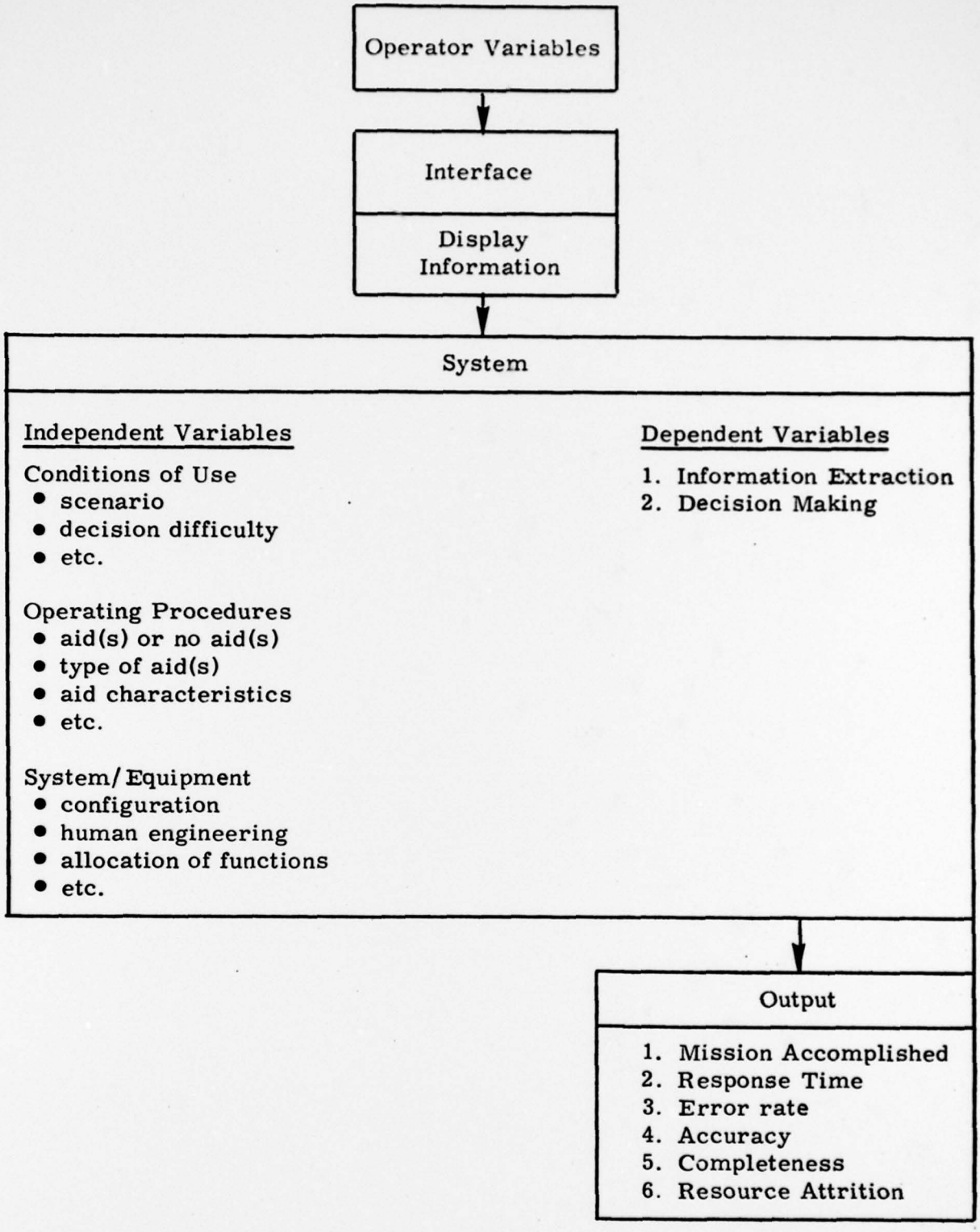

Figure 1. Conceptualization and decompositon of evaluative process.

The system/equipment independent variable class includes manipulations of configuration, human engineering, and functional allocations. Similarly, these factors will be reflected by information extraction and decision making dependent variables.

Restricting the Number of Treatment Conditions

The question arises as to the range over which the independent variables should be presented during any single aid evaluation. Variables fall into two obvious categories: (1) those which are continuous in nature, such as information load, and (2) those which are discrete, such as interface configuration. Unfortunately, the expense of system evaluation experiments precludes testing systems in a controlled fashion using all possible or even a large number of treatments of a specific independent variable.

ASTDA Test Plan

For the initial tests of an ODA system, the Applied Psychological Services developed and presented an experimental test plan for evaluating Analytics, Inc. Strike Timing Decision Aid (ASTDA). The test plan was formally presented in May of 1978 (Siegel, 1978), and represented a synthesis of earlier test plans developed both by Analytics and Applied Psychological Services. The purpose of the test plan was to describe the methods and procedures for test of five major hypotheses relative to the ASTDA:

Hypothesis 1. More effective strike timing decisions can be made using the ASTDA than without the aid.

Hypothesis 2. Users will perceive the ASTDA to possess value.

Hypothesis 3. The effectiveness and perceived value of the ASTDA will not vary as a function of user experience level or decision problem difficulty.

Hypothesis 4. The ASTDA possesses criterion related validity where the criterion is best strike time judgments of experienced strike planning Navy officers.

Hypothesis 5. Decision effectiveness will systematically vary as ASTDA features are varied. Three features of the ASTDA aid are to be varied: (1) display of expected own and enemy losses, (2) display of expected utility of overall strike mission outcomes, and (3) display information showing the uncertainty of actual air strike conditions (e. g., own force readiness, enemy force strength, weather) and outcomes.

The research design is a 5 x 2 x 2 mixed factor design. The three factors are: (1) ASTDA features, (2) decision problem difficulty, and (3) experience level of subjects in Naval strike timing planning. Five different levels of ASTDA features including a no aid control conditions are varied. These are shown in Table 1.

Table 1

Summary of ASTDA Levels

Treatment (Level)	Information Provided			
	Input	Utility	Outcome	Uncertainty Bands
1	✓	✓	✓	✓
2	✓	✓		✓
3	✓		✓	✓
4	✓	✓	✓	
5 (unaided)	✓			

In the unaided condition (control), the experimental subjects will receive their information by way of a telephone link with an actor (or actors) who will provide such information as would normally be available from the operations officer, maintenance officer, aerology, and the like.

The total situation represents a fully controlled laboratory evaluation which allows the collection and subsequent analysis of both quantitative and qualitative data including the process of monitoring/utilization.

By the close of the annual reporting period, the evaluation was fully designed, the test bed was developed, the actual problems to be employed were written, and plans were established for determining problem difficulty along a graded scale. The test bed development was completed by the Department of Decision Sciences, University of Pennsylvania, and the problem development was completed by Analytics, Inc.

Plans were also completed, and methods were developed by the end of the reporting period for collecting data relative to the difficulty level of each problem and relative to the decision making heuristic employed by experienced Naval officers, when making strike timing decisions while employing the problems along with the information provided by the Analytics' Strike Timing Aid.

Limitations to ASTDA Evaluation

Human-interactive systems studies can be conducted at various levels of complexity, i. e., component, subsystem, and system levels. However, in the present work, the component and subsystem levels will largely be embedded in the system. Fragmentary studies may be indicated, but the emphasis in the current evaluation program is on the total system--human operator, software, hardware, and displays. The number of criteria against which the system can be evaluated is potentially quite large. However, the actual magnitude of this problem is practically reduced because all criteria are not available for test.

The selection among system criteria must include an understanding of the user's requirements or a methodology by which some criteria can be traded off for others. The need to tradeoff arises because individual criteria often conflict with one another. For example, it may not be possible to accomplish a mission without suffering some losses. Similarly, response time and quality may come into conflict. While it may be extremely important to process given items of information in the shortest possible time, it may also be necessary to sacrifice some performance quality to do so.

Support Research

While the evaluative research constitutes the principal thrust of the present Applied Psychological Services' program, a collateral effort involves the development of basic data important to the design of any man-computer interactive interface. To this end, an experimental study was designed and the stimuli were developed for an investigation of the advantages and disadvantages of color and type of display (tabular or graphic), in interaction with the type of use of information (information extraction or decision making). This type of work is viewed as an adjunct to the primary evaluative research of the present program. However, the collateral work will yield important user interface information pertinent to the structural design of any computer based operational decision aid.

Interview Techniques

A major consideration in system experimentation is gaining insight into system problems from the point of view of the system user. Specific user problems which differentiate between effective and ineffective performance are often difficult to identify. They are often nonobvious to the evaluator. In order to identify such problems, an interview is included in the evaluative techniques. During each evaluation, an observer will observe and then debrief (interview) test personnel. Responses to questions relative to critical aspects of performance, training, existing display formats, and the man-machine interface will be elicited. For each of the interview items, the information will be classified into preestablished categories and summarized. The summary will then be employed diagnostically and proscriptively to improve performance, training, and human engineering of the man-machine interface for later evaluative experiments.

REFERENCES

Applied Psychological Services, *DAISY Evaluative concepts and initial experiment.* Working paper. Wayne: Applied Psychological Services, 1978.

Carlisle, J.H. *An experimental framework for man-computer interaction research.* Unpublished report under contract with the Office of Naval Research, 1978.

Cook, K.H. An experimental on-line system for psychological abstracts. *ASIS Proceedings,* 1970, *7*, 111-114.

Gerrity, Jr., T.P. Design of man-machine decision systems: An application to Portfolio Management. *Sloan Management Review.* Winter, 1971. 59-75.

Katter, R.V., & McCarn, D.B. *AIM-TWX: An experimental on-line bibliographic retrieval system.* In D. Walker (Ed.), Proceedings of the first workshop on User Interface. New York: AFIPS Press, 1971.

Katter, R.V. Design and evaluation of information systems. *Annual Review of Information Science and Technology,* Vol. 4. Chicago: Encyclopedia Britannica, 1969.

King, D.W. *Design and Evaluation of Informative Systems.* In C. Cuadice (Ed.) *Annual Review of Information Science and Technology,* Vol 3. New York: Interscience, 1968.

Marcus, R.S., Benenfeld, A.R.B., Kugel, P. The user interface for the intrex retrieval system. Proceedings of a workshop on *The User Interface for Interactive Search of Bibliographic Data Bases.* Palo Alto, Calif., 1971. In press, Montvale, N.J.: AFIPS Press, 1971.

Martin, T.H., & Parker, E.B. Designing for user acceptance of an interactive bibliographic facility. Proceedings of a workshop on *The User Interface for Interactive Search of Bibliographic Data Bases.* Palo Alto, Calif., 1971 . In press, Montvale, N.J.: AFIPS Press, 1971.

Rees, A.M. *Evaluation of Information Systems and Services.* In C. Caudice (Ed.), *Annual Review of Information Science and Technology.* Vol. 2, Interscience, New York: 1967.

Siegel, A.I. *An experimental design for evaluating the "Strike Timing Aid."* Wayne: Applied Psychological Services, 1978.

DISTRIBUTION LIST

Contract N00014-77-C-0448

Director, Engineering Psychology Programs
(Code 455)
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217 (5 copies)

Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, VA 22314 (12 copies)

CD Paul Chatelier
Office of the Deputy Under Secretary of Defense
OUSDRE (E&LS)
Pentagon, Room 30129
Washington, D.C. 20301

Dr. H. W. Sinaiko
Smithsonian Institution
801 N. Pitt Street
Alexandria, VA 22314

Dr. Stephen Andriole
Director, Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

Office of the Chief of Naval Operations
OP987P10
Personnel Logistics Plans
Washington, D.C. 20350

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, VA 22202

Dr. Amos Freedy
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Commanding Officer
Office of Naval Research Branch Office
ATTN: Dr. J. Lester
Bldg. 114, Section D
Boston, MA 02210

Commanding Officer
Office of Naval Research Branch Office
ATTN: Dr. Charles Davis
536 South Clark Street
Chicago, IL 60605

Commanding Officer
Office of Naval Research Branch Office
ATTN: Dr. E. Gloye
1030 East Green Street
Pasadena, CA 91106

Commanding Officer
Office of Naval Research Branch Office
ATTN: Mr. R. Lawson
1030 East Green Street
Pasadena, CA 91106

Information Systems Program
Code 437
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Naval Analysis Programs
Code 431
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Operations Research Program
Code 434
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Mr. Robert Garnero
Stanford Research Institute
Naval Warfare Research Center
Menlo Park, CA 94025

Dr. G. Hurst
University of Pennsylvania
Wharton School
Philadelphia, PA 19174

Dr. Miley Merkhofer
Stanford Research Institute
Decision Analysis Group
Menlo Park, CA 94025

Dr. C. Peterson
Decisions and Designs, Inc.
8400 Westpark Drive, Suite 600
P. O. Box 907
McLean, VA 22101

Mr. George Pugh
Decision Science Applications, Inc.
1401 Wilson Blvd.
Arlington, VA 22209

Mr. David Walsh
Integrated Sciences Corporation
1640 Fifth Street
Santa Monica, CA 90401

Mr. Merlin Malehorn
Office of the Chief of Naval Operations
(OP 102X)
Washington, D.C. 20350

Mr. Harold Crane
CTEC, Inc.
7777 Leesburg Pike
Falls Church, VA 22043

Dr. S. D. Epstein
Analytics, Inc.
2500 Maryland Road
Willow Grove, PA 19090