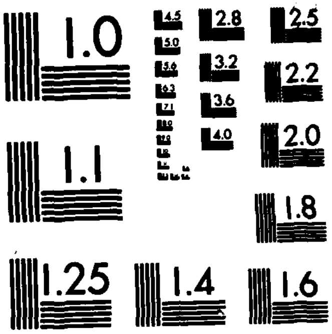
AD-A126 978 BEHAVIORAL INPUTS TO THE WEAPON SYSTEM ACQUISITION 1/1
PROCESS(U) NAVY PERSONNEL RESEARCH AND DEVELOPMENT
CENTER SAN DIEGO CA D MEISTER MAR 83 NPRDC-SR-83-21
UNCLASSIFIED F/G 5/9 NL

END



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

①

NPRDC SR 83-21

MARCH 1983

BEHAVIORAL INPUTS TO THE WEAPON SYSTEM ACQUISITION PROCESS

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED

NAVY PERSONNEL RESEARCH
AND
DEVELOPMENT CENTER
San Diego, California 92152

NPRDC Special Report 83-21 March 1983

BEHAVIORAL INPUTS TO THE WEAPON SYSTEM ACQUISITION PROCESS

D. Meister

Reviewed by
R. E. Blanchard

Released by
James F. Kelly, Jr.
Commanding Officer

Navy Personnel Research and Development Center
San Diego, California 92152

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
NPRDC SR 83-21

2. GOVT ACCESSION NO.
AD-A126 998

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
BEHAVIORAL INPUTS TO THE WEAPON SYSTEM ACQUISITION PROCESS

5. TYPE OF REPORT & PERIOD COVERED
Jan 1981 - Dec 1981

6. PERFORMING ORG. REPORT NUMBER
17-81-15

7. AUTHOR(s)
D. Meister

8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Navy Personnel Research and Development Center
San Diego, California 92152

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Navy Personnel Research and Development Center
San Diego, California 92152

12. REPORT DATE
March 1983

13. NUMBER OF PAGES
92

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

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

18. SUPPLEMENTARY NOTES

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

Human factors engineering
System development
Weapon system acquisition

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

The Navy's weapon systems are often designed with inadequate attention being paid to human factors considerations, with consequent reduction in system performance. To ensure that behavioral inputs are made to the weapon system acquisition process, they must be made to the various reviews conducted by the Defense System Acquisition Review Council (DSARC). This report describes the behavioral inputs required and the DSARC phases to which they are relevant. It describes the documents into which inputs must be inserted, the DSARC behavioral questions that should be asked, and the behavioral techniques that should be implemented.

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

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

FOREWORD

This report was written under project PO-53003 (Human Factors Considerations in Weapon Systems Acquisition) and was sponsored by the Chief of Naval Operations (OP-115). Its purpose was to advise those involved in the weapon system acquisition process (WSAP) of the behavioral inputs they should require of planners and system planners.

A number of reports have been produced over the years to familiarize and indoctrinate WSAP personnel about required behavioral inputs (c.f., Baker, Johnson, Malone, & Malone, 1979; Bureau of Naval Personnel, 1964; Condon, Hayes, Turner, & Walder, 1970; Greer, 1976, 1977; HARDMAN, 1979a, 1979b; Holshouser, 1975, 1977; Malone, Gloss, & Eberhard, 1967; Meister, 1971; Price, Fiorello, Lowry, Smith, & Kidd, 1980a, 1980b). The current report was produced because WSAP has changed somewhat over the years, particularly since the establishment of the Defense System Acquisition Review Council, which is responsible for making recommendations concerning program direction. Although this requirement has not changed the behavioral inputs to the acquisition process, the latter must be viewed in a different framework. Also, this report was written to familiarize planners and developers who have recently become involved in the acquisition process with behavioral requirements and inputs.

JAMES F. KELLY, JR.
Commanding Officer

JAMES W. TWEEDDALE
Technical Director

SUMMARY

Problem

The Navy's weapon systems are often designed with inadequate attention being paid to human factors considerations. As a result, system effectiveness is reduced. If behavioral inputs are to be included in the weapon system acquisition process (WSAP), they must be made to documents required for use in various phases of the process (e.g., decision coordinating papers). In particular, behavioral inputs must be made to the various WSAP reviews conducted by the Defense System Acquisition Review Council (DSARC) for Secretary of Defense approval.

Purpose

The purpose of this report is to describe the behavioral questions that should be asked at DSARC reviews and the behavioral inputs that should be made during WSAP to address these questions.

Contents

1. WSAP was described in terms of acquisition directives, acquisition phases, and behavioral inputs to required WSAP documentation.

2. The questions that must be addressed at the various DSARC milestones are listed, together with their behavioral equivalents. Also, the activities required to supply the answer to the behavioral equivalent (i.e., testing analysis, etc.) are described in detail in Appendices A through N.

Recommendations

All personnel involved in the DSARC decision-making process should help ensure that the behavioral inputs described in this report are provided at the specified DSARC milestones.

CONTENTS

INTRODUCTION

Problem

The Navy's weapon systems are often designed with inadequate attention being paid to human factors considerations. As a result, system effectiveness is reduced. If behavioral inputs are to be included in the weapon system acquisition process (WSAP), they must be made to documents required for use in various phases in the process. In particular, behavioral inputs must be made to the various WSAP reviews conducted by the Defense System Acquisition Review Council (DSARC) for Secretary of Defense approval. The questions asked at these reviews are of an operational, logistical, financial, or engineering nature. They do not usually involve behavioral elements (e.g., number and skill level of personnel required, and their training), although it is obvious that the effectiveness or cost of any man-machine system (MMS) that is being acquired will depend heavily on those elements. For example, if the system's control operations are unduly complex, they will require personnel of exceptional skill. Personnel who have these skill requirements may not be available or, if they are available, will require an excessively long (and costly) training period.

Purpose

The purpose of this report is to describe the behavioral questions that should be asked at DSARC reviews and the behavioral inputs that should be made during WSAP to answer these questions. These behavioral inputs are of four types:

1. Those impacting on the selection and acquisition of system personnel:

    a. Determination of the number of personnel required by the system and available.

    b. Description of the jobs to be performed in the new system (usually in terms of standard Navy ratings).

    c. Description of the skills and skill levels required of operating and maintenance personnel.

2. Those impacting on personnel training:

    a. Specification of length of time (in a calendar year) required for training the jobs to be performed in the new system.

    b. Specification of number of students throughput during that time period.

    c. Specification of equipment facilities (e.g., trainers, simulators, plant) needed for training and available.

    d. Specification of number of instructors needed and available.

3. Those affecting the design of the system hardware, software, and procedures (human factors engineering (HFE)):

    a. Design and/or review/evaluation of man-machine interfaces (usually displays and control panels).

b. Design and/or review/evaluation of software in computers.

c. Design and/or review/evaluation of job procedures.

d. Specification of the required characteristics of the working environment (e.g., lighting, temperature, noise, etc.).

e. Prediction/measurement of personnel performance.

4. Those related to testing the personnel elements of the system and evaluating their operational effectiveness:

a. Specification of personnel performance criteria and measures.
b. Specification of appropriate statistical and experimental designs.
c. Design, review, and evaluation of test scenarios.
d. Conduct of personnel performance tests.
e. Analysis of personnel performance test data and resultant conclusions.

DESCRIPTION OF THE WEAPONS SYSTEM ACQUISITION PROCESS (WSAP)

Acquisition Directives

The documents described below direct the Department of Defense (DoD) and thus the Navy in conducting the WSAP. Since the behavioral inputs made to the DSARC reviews must be responsive to these directives, it is necessary to examine them in detail.

1. DoD Directive 5000.1 (1977a) provides basic policy for systems costing more than $75 million in research, development, testing, and evaluation (RDT&E) or over $300 million in full-scale procurement. That policy can be summarized as follows:

a. Acquisition is a sequence of phases that begins following approval of a mission need.

b. The individual services must analyze mission needs to develop systems that satisfy those needs.

c. The Secretary of Defense (SECDEF) makes decisions regarding program commitments (i.e., to initiate programs and provide funding) at four decision points (see Figure 1):

(1) Milestone I--Program initiation. Requires that a mission need be demonstrated in a document called the mission element need statement (MENS).

(2) Milestone II--Demonstration and validation. Depends on recommendations made in the decision coordinating paper (DCP).

(3) Milestone III--Full-scale engineering development. Based on updated versions of the DCP.

(4) Milestone IIIA--Production and deployment. Same as above.

d. Existing hardware and software are to be used as much as possible to satisfy mission needs.

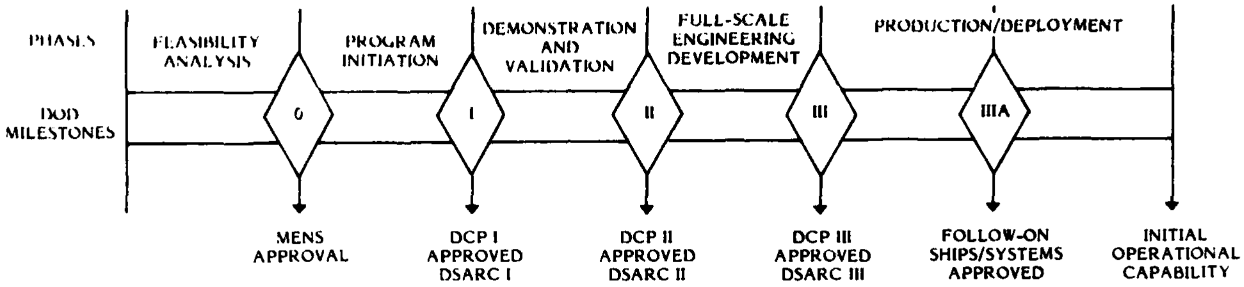


Figure 1. Major DoD milestones and phases in the WASP.

e. Test and evaluation are to begin as early as possible.

f. Human engineering factors are to be included as constraints on system design and as a system element, starting with initial concept studies and continuing throughout the development process. These factors are to form the basis for personnel selection, training, training devices, simulators, and planning of the system as it is affected by these factors.

2. DoD Directive 5000.2 (1977b) establishes the process by which major systems are procured. SECDEF controls acquisition programs through the four milestone decisions noted above. Directive 5000.2 also establishes the DoD DSARC and its subordinate, the Department of the Navy System Acquisition Review Council (DNSARC), which are responsible for reviewing DCPs and for making recommendations concerning program direction.

3. DoD Directive 5000.3 (1977c) requires that:

a. All systems will be subject to test and evaluation (T&E).

b. T&E shall begin as early as possible in the acquisition cycle and shall be conducted throughout the development process.

c. Acquisition schedules will be based on T&E milestones.

d. T&E of existing or modified equipment may be performed before a new system is developed and shall consider environmental issues.

e. The DCP at milestone I must identify critical questions and areas of risk to be resolved by T&E

f. The DCP at milestone II must provide the results of T&E to date.

g. DSARC will review T&E results before making recommendations at milestone III.

Acquisition Phases

As shown in Figure 1, there are five acquisition phases, each leading to a program milestone: (1) feasibility/analysis, (2) program initiation, (3) demonstration and validation, (4) full-scale engineering development, and (5) production and deployment. The last four phases correspond to the milestones noted in DoD Directives 5000.1 and 5000.2. Since the studies, analyses, and development occurring in these four phases are iterative, there is considerable reiteration of the various documents that must be developed during these phases. For example, DCPs and Navy DCPs (NDCPs) are continually updated and amplified in each of the phases. Hence, the behavioral inputs made to these documents must also be repeatedly updated.

Since the organizational relationships involved in reviewing the various documents leading to milestone decisions are fairly complex, no attempt has been made in this report to describe them fully. Table 1, which was modified from HARDMAN 79-0 (1979a), lists the organizations and the activities they perform for each phase in the WSAP. It should be noted that none of the organizations listed has a staff of behavioral specialists, with the possible exception of the system development program manager (PM). Therefore, any behavioral inputs that do reach DSARC, if they do at all, do so through the PM. They are not (with a few exceptions, such as proposed manning) routinely provided to DSARC decision makers. The five acquisition phases are described below.

Feasibility/Analysis Phase

The major activity in this phase is the identification and definition of a mission need. Such a need is either a technological development that counters a known threat or the recognition of a strategic or tactical threat that requires the development of a new weapon system. The mission need is analyzed in either MENS or the operational requirement (OR), which are described below.

1. MENS indicates the following:

   a. Mission area and need in terms of mission tasks to be performed.
   b. Projected threat assessment.
   c. Existing capabilities to accomplish the mission.
   d. Need in terms of deficiency in capability.
   e. Known constraints to solutions (e.g., cost, time frame, etc.).
   f. Effect of lack of capability.
   g. A plan for identifying and exploring alternative systems.

It would be incorrect to assume that MENS does not require any behavioral inputs. In particular, one of the known contraints to solutions (e. above) may be lack of manpower, personnel, and training (MP&T) resources. MENS must also be reviewed to ensure that at least one of the alternatives identified in the document keeps manpower requirements within current mission area levels. The stated need must be assessed in terms of a manpower deficiency in the existing functional capability.

2. OR is a statement of the operational need. The need may deal with an MP&T deficiency (in the case of behavioral research designed to find an answer to the deficiency). In the case of a system development requirement (as in MENS), any MP&T deficiency constraining solution of the problem must be explored. It is limited to three pages.

Table 1

Summary of WSAP Procedures

Step	Organization	Activity
		Feasibility/Analysis Phase
1	OP-03	Prepares MENS to document the mission need.
2	OP-01	Reviews MENS for manpower, personnel, and training implications.
3	OP-090	Reviews and concurs with MENS.
4	CNO Executive Board	Reviews and concurs with MENS.
5	CNO	Forwards MENS to the Navy Acquisition Executive (NAE) with recommendation for approval.
6	NAE	Reviews MENS and prepares a position paper for SECNAV.
7	SECNAV	Forwards draft MENS to the Defense Acquisition Executive (DAE).
8	DAE	Obtains comments on draft MENS from Office of the Secretary of Defense (OSD) staff and the Office of Joint Chief of Staff (OJCS).
9	SECNAV	Forwards revised MENS to the DAE with recommendation for approval.
10	DAE	Prepares position paper and proposed action memorandum for SECDEF.
11	SECDEF	Authorizes establishment of a SECDEF-designated program.
		Program Initiation Phase
1	OP-03	Appoints a program coordinator (PC).
2	CNM	Charters a project manager (PM) and establishes a project manager, ship (PMS) officer.
3	PM	Solicits contractor and in-house conceptual responses to MENS.
4	OP-03	Issues the Top Level Requirements Document (TLR) (OP-01 provides the necessary manning limitations).
5	PM	Issues the Top Level Specifications (TLS) to establish the functional baseline.
6	PM	Prepares the DP for the review of alternative concepts by the Office of the Chief of Naval Operations (OPNAV).
7	OP-01	Reviews the DP for MP&T implications of alternative selections.
8	OP-03	Submits a proposed DCP outline to the DAE for OSD approval.
9	OSD and SECNAV Staff	Reviews and approves the DCP outline and schedules the DNSARC review and the DSARC I review.
10	CNM	Conducts review of program through Acquisition Review Board (ARB) and the Logistics Review Group (LRG).
11	OP-03	Prepares a draft "For Comment" DCP and distributes for review.
12	OPNAV Staff	Returns comments on the draft DCP to OP-03.
13	PM and PC	Make DCP presentation to the CEB.
14	CNO Executive Board	Reviews the DCP and formulates a CNO recommendation.
15	CNO	Approves the DCP and forwards recommendation to the NAE.
16	PM and PC	Make DCP presentation to the DNSARC.
17	DNSARC	Reviews the DCP and formulates a SECNAV recommendation.
18	SECNAV	Approves the DCP and forwards recommendation to the DAE.
19	DAE	Obtains comments from the OSD Staff and the OJCS and returns the DCP to OP-03 with proposed revisions (repeat steps 11 to 19 if major issues need to be resolved internally as a result of the OSD revisions).
20	OP-03	Prepares a draft "For Coordination" DCP for OSD review and approval.
21	OP-01	Review and verify "For Coordination" draft DCP.
22	PM and PC	Make pre-DSARC presentation to the CEB and DNSARC for final CNO/SECNAV coordination.
23	PM and PC	Make DCP presentation to the DSARC (DSARC I).
24	DSARC	Reviews the DCP and formulates recommendations for SECDEF decision.
25	SECDEF	Approves the DCP and authorizes the initiation of the demonstration and validation phase.
		Demonstration and Validation Phase
1	PM	Develops preliminary ILS plan including specific MP&T requirements.
2	SEA-04/MAT-04	Reviews ILS plan.
3	OP-04	Reviews ILS plan.
4	PM	Prepares draft preliminary ship manpower document (PSMD) and submits it to OP-01 for review and evaluation.
5	OP-01	Reviews the draft ship manning document (SMD) and approves PSMD.
6	PM	Develops draft Navy Training Plan (NTP) and issues for review.
7	OPNAV Staff	Returns comments on the NTP to the PM with OP-03 providing direction and supervision.
8	OP-03	Convenes and chairs a NTP conference.
9	PM	Prepares a proposed NTP and submits it to OP-01 via OP-03.
10	OP-01	Reviews, approves, and promulgates the NTP.

Table I (Continued)

Step	Organization	Activity
		Demonstration and Validation Phase (Continued)
11	OP-03	Submits a proposed DCP outline to the DAE for OSD approval.
12	OSD and SECNAV Staff	Reviews and approves the DCP outline and schedules the DNSARC review and the DSARC II review.
13	CNM	Conducts review of program through the ARB and the LRG.
14	OP-03	Prepares a DCP cover sheet revision and distributes for review.
15	OPNAV Staff	Returns comments on the DCP cover sheet revision to OP-03.
16	PM and PC	Make DCP presentation to the CEB.
17	CEB	Reviews the DCP and formulates a CNO recommendation.
18	CNO	Approves the DCP and forwards recommendation to the NAE.
19	PM and PC	Make DCP presentation to the DNSARC.
20	DNSARC	Reviews the DCP and formulates a SECNAV recommendation.
21	SECNAV	Approves the DCP and forwards recommendation to the DAE.
22	DAE	Obtains comments from the OSD Staff and the OJCS and returns the DCP to OP-03 with proposed revisions (repeat steps 14 to 22 if major issues need to be resolved internally as a result of the OSD revisions).
23	OP-03	Prepares a draft "For Coordination" DCP for OSC review and approval.
24	OP-01	Reviews and verifies the "For Coordination" DCP.
25	PM and PC	Make pre-DSARC presentation to the CEB and DNSARC for final CNO/SECNAV coordination.
26	PM and PC	Make DCP presentation to the DSARC (DSARC II).
27	DSARC	Reviews the DCP and formulates recommendations for SECDEF decision.
28	SECDEF	Approves the DCP and authorizes the initiation of the full-scale engineering development phase.
		Full-scale Engineering Development Phase
1	PM	Updates PSMD and submits it to OP-01 for review and approval.
2	OP-01	Reviews, approves, promulgates the PSMD.
3	OP-01	Prepares manpower authorization change request (OPNAVFORM 1000/4A).
4	PM	Develops and issues draft updated NTP for review.
5	OPNAV Staff	Returns comments on the updated NTP to the PM with OP-03 providing direction and supervision.
6	OP-03	Convenes and chairs a NTP Conference.
7	PM	Prepares a proposed updated NTP and submits it to OP-01 via OP-03.
8	OP-01	Approves and promulgates the updated NTP.
9	OP-03	Submits a proposed DCP outline to the DAE for OSD approval.
10	OSD and SECNAV	Reviews and approves the DCP outline and schedules the DNSARC and the DSARC III review.
11	CNM	Conducts review of the program through the ARB and the LRG.
12	OP-03	Prepares a DCP cover sheet revision and distributes for review.
13	OPNAV Staff	Returns comments on the DCP cover sheet revision to OP-03.
14	PM and PC	Make DCP presentation to the CEB.
15	CEB	Reviews the DCP and formulates a CNO recommendation.
16	CNO	Approves the DCP and forwards recommendation to the NAE.
17	PM and PC	Make DCP presentation to the DNSARC.
18	DNSARC	Reviews the DCP and formulates a SECNAV recommendation.
19	SECNAV	Approves the DCP and forwards recommendation to the DAE.
20	DAE	Obtains comments from the OSD Staff and the OJCS and returns the DCP to OP-03 with proposed revision (repeat steps 12 to 20 if major issues need to be resolved internally as a result of the OSD revisions).
21	PM	Completes the remaining parts of the ILS plan.
22	SEA-04/MAT-04	Reviews the ILS plan.
23	OP-04	Reviews the ILS plan.
24	NAVMAT/PM	Prepares the logistic support plan summary.
25	OP-03	Submits logistic support plan summary to OP-04.
26	OP-04	Provides report of logistics program readiness.
27	OPTEVFOR	Submits operational test and evaluation report to OP-098 prior to CEB review.
28	NAVMAT/NAVSEA	Requests "Provisional Approval for Service Use" through OP-03.
29	CEB	Reviews the OP-04 and OPTEVFOR reports and makes recommendation to CNO.
30	CNO	Grants "Provisional Approval for Service Use" based on the CEB recommendation.
31	OP-03	Prepares a draft "For Coordination" DCP for OSD review and approval.
32	OP-01	Reviews and verifies the "For Coordination" DCP.
33	PM and PC	Make pre-DSARC presentation to the CEB and DNSARC for final CNO/SECNAV coordination.
34	PM and PC	Make DCP presentation to the DSARC (DSARC III).
35	DSARC	Reviews the DCP and formulates recommendations for SECDEF decision.
36	SECDEF	Approves the DCP and authorizes the initiation of the production phase.

Program Initiation Phase

The major activities performed in this phase include (1) reestablishment of the mission need, (2) a survey of available technology to identify areas of technological inadequacy of proposed systems, (3) the beginning of a definition of an acquisition strategy, and (4) preparation and issuing of documentation required for the milestone I decision. That documentation consists of the following:

1. Developmental proposal (DP).

2. Decision coordinating paper (DCP), which (a) provides primary documentation for use by DSARC, and (b) summarizes program, acquisition strategy, alternatives considered, issues, direction needed by decision authority, and additional requirements issued only by acquisition executive.

3. Integrated program summary (IPS), which (a) summarizes an implementation plan for life cycle of system, and (b) provides information for management overview of entire program and on detailed requirements in DoD Directive 5000.2 and additional requirements issued only by acquisition executive.

4. Milestone reference file (MRF) (established at each milestone in central location), which provides program documentation (by DoD component) referenced in DCP and IPS. This information is provided to DSARC (or equivalent) executive secretary at time "for comment DCP and IPS" submitted and for use by DoD personnel needing detailed information.

5. Test and evaluation master plan (TEMP), which describes (a) the system and intended operational mission, (b) critical T&E issues, (c) project objectives and thresholds, (d) required technical and operational characteristics, (e) environmental impact assessment of T&E, (f) integrated schedule, and (g) T&E resources required.

6. Life cycle cost estimate.

The DP, which is prepared by the Naval Material Command (NMC) (the command responsible for system acquisition), presents alternatives and tradeoffs to achieve a range of capabilities in response to the MENS/OR. It forms the basis for the DCP or NDCP, which is the most important documentation in terms of reaching a decision.

The DP is reviewed from a behavioral standpoint to ensure that manpower estimates included in the alternatives presented are accurate and that the Other Factors section of the DP includes consideration of training, support, and human resources factors that would affect the introduction of the new system. The DCP/NDCP is reviewed to ensure that (1) estimated manning levels are included, (2) manpower requirements are compared with those of a baseline (predecessor) operational systems if one exists, (3) potential tradeoffs among manpower, design, and logistical elements are analyzed, and (4) the training concepts that will be analyzed during the demonstration and validation phase (see below) are identified. The life cycle cost analysis in the various alternatives must include manpower and training components.

Demonstration and Validation Phase

At this point, alternative weapon system concepts have been determined, the subsystems targeted for advanced development have been identified, the mission need has been defined, and acquisition plans have been developed. During the demonstration and validation phase, the following steps are performed:

1. The preliminary design is initiated.
2. The management plan is developed.
3. The test and evaluation management plan is established.
4. The integrated logistics support (ILS) plan is established.
5. Requests for proposals for system/subsystem development are written.
6. Prototypes of systems under development are constructed.
7. Preparations are made for the milestone II decision.

Two important documents required during this phase are the test and evaluation master plan (TEMP) and the ILS plan. The ILS plan is particularly important from the standpoint of behavioral inputs because two of its principal elements are personnel and training. It is reviewed to ensure that manpower implications, including life cycle costs, are adequately addressed. This includes manning estimates in terms of numbers and skills, unique personnel resource constraints (e.g., introduction of new skills or critical skills not in the Navy inventory), life cycle cost estimates for personnel and training, the training concept, and the scheduling of manpower, training, and equipment so that all three coincide.

TEMP is also important for behavioral inputs because two of the major testing elements are determining whether personnel can perform required tasks adequately in the new system and testing the system to demonstrate and verify that its characteristics do not negatively impact on the ability of personnel to perform their jobs.

DoD Directive 5000.3 requires that developmental and operational testing be accomplished to ensure that engineering is reasonably complete and that all significant design problems (presumably including human factors/personnel problems) have been identified. TEMP, which is written about the time of DSARC I by the program manager, is used both as an input to the DSARC decision process and as a T&E management plan.

From a behavioral standpoint, TEMP is a critical document because one of the major subsystems to be tested during development is the personnel subsystem. With the development of TEMP, it becomes necessary to identify human factors T&E problem areas, principally in the area of design. The questions to be answered during developmental testing include the following:

1. Is there some aspect of design that could negatively and significantly affect performance of system personnel?

2. Do system personnel perform well enough to satisfy system requirements?

3. Are special skills required of system personnel that may not be within the Navy personnel inventory?

4. Are there any functions to be performed by personnel that may pose special difficulties for them?

A special section of TEMP must be reserved for personnel/human factors. The behavioral input to that section will include objectives to be satisfied by personnel/human factors testing and procedures by which personnel/human factors testing will be accomplished as part of the overall engineering test.

Another major input into DSARC II is the Navy training plan (NTP), which is developed to describe the training resources required to support the manpower requirements specified in the preliminary ship manning document (PSMD). The NTP also supports the ILS concept. It describes (1) the training concept to be pursued, (2) training device and equipment requirements, (3) required class size, (4) military construction if needed, (5) instructor requirements, (6) method of training, (7) location of training, (8) number of billets for which personnel will be trained, and (9) cost of training.

After the demonstration and validation phase is over and SECNAV is prepared to recommend the preferred system for full-scale engineering development, this recommendation is documented in an updated DCP and reviewed by DNSARC and DSARC prior to SECDEF decision. DSARC II reaffirms the mission element need and updating of the threat and asks whether:

1. The system in development meets mission element needs.

2. Systems tradeoffs have produced the optimum balance in cost, performance, and schedule.

3. Risks are acceptable.

4. Planning for selection of major subsystems is underway.

5. Testing and evaluation have been completed and the results support the recommendation.

6. TEMP identifies the T&E to be accomplished prior to DSARC II and III.

With regard to behavioral issues, DSARC II manpower documentation should (1) provide the rationale for manpower estimates, (2) identify any unique skills required, (3) estimate manpower requirements for maintenance to be performed below depot level, and (4) discuss T&E plans, etc. The specific questions that must be answered in DSARC II as they relate to behavioral inputs will be discussed later.

Full-scale Engineering Development Phase

The following activities are performed during this phase: (1) detailed ILS specifications are developed, (2) requests for proposals (RFPs) for the system are written, (3) full-scale engineering development of the system is completed and production planning/preparation begins, (4) T&E (developmental) continues, and (5) preparations are made for the milestone III decision. The various document inputs that were developed previously (e.g., ILS, TEMP, etc.) are updated in the light of the information gained during actual physical development of the system and the tests that have been performed.

Production and Deployment Phase

The milestone III decision will indicate whether the system under development will go into full-scale production and be operationally deployed.

Behavioral Inputs to Required WSAP Documentation

The behavioral inputs that can be made to the documents required at the various WSAP phases are described below.

1. MENS/OR. The MENS or OR seeks to establish a need for a new system. Behavioral inputs will usually not be relevant to the determination of such a need, unless it reflects an inability to use a present system because of personnel difficulties (which, realistically, rarely occurs). However, one section of MENS describes known constraints to solutions. One such constraint may be the unavailability of personnel with required skills, excessive difficulties in training required personnel, or a required training curriculum of prolonged duration that might increase personnel cost to unacceptable levels. Another section of MENS requires that a plan be developed for identifying and exploring alternative systems. Such a plan should consider personnel and manpower factors when alternative systems are being conceptualized.

HARDMAN 79-02 recommends that the MENS/OR promulgating letter contain the necessary instructions and reporting formats for providing manpower requirements in the DP and that the MENS contain MP&T and life cycle cost constraints. The OR should include a commitment to full consideration of manpower costs, hardware/manpower tradeoff analysis, and the feasibility of providing personnel with required skills.

2. DP/DCP. An important aspect of the OR is the establishment of the DP, which describes (a) the technical approach (or alternative approaches) to satisfy the operational requirements, (b) an economic analysis and relative benefits of the alternative technical approaches, and (c) a recommendation for the technical approach selected. The DCP's principal purpose is to support SECDEF and DSARC in determining program continuation. It contains an updated MENS and descriptions of alternative programs, an acquisition strategy, a management plan, risks, and T&E planning status.

For each technical approach and program, a verified manpower estimate based on specified data is required. The Other Factors section must include lists of training, support, and human resource factors that could impact on the effective introduction of the system. The DCP must be checked to ensure that (a) estimated manning levels will meet peacetime and wartime requirements, (b) manpower requirements of the new system are compared with those of the baseline (predecessor) system, (c) potential tradeoffs among manpower, design, and logistic elements are identified, and (d) training concepts to be analyzed during demonstration and validation are identified.

3. ILS Plan. The ILS plan should include (a) manning estimates (number and skill levels), (b) a description of personnel resource constraints (if any), (c) life cycle cost estimates for personnel and training, (d) description of training concept, and (e) a schedule for obtaining manpower, training and equipment, so that all coincide in time. As part of the ILS review, the accuracy of MP&T estimates must be verified.

4. TEMP. Behavioral inputs to TEMP are described in some detail in Appendix G.

5. NTP. Behavioral inputs to NTP are described in some detail in Appendix K.

6. Other. Other required documentation during the WSAP consists largely of updates of the previously described documents and do not require qualitatively new behavioral inputs. Hence, they will not be discussed further.

DSARC QUESTIONS AND THEIR BEHAVIORAL EQUIVALENTS

At DSARC milestones I, II, and III/IIIA, a number of questions must be addressed to satisfy the requirements of meeting these milestones. Although these questions do not have a behavioral orientation, they can be translated into personnel, training, HFE, and T&E equivalents. In Table 2, which draws considerably on previous work by Holshouser (1975), the DSARC questions are listed, together with their behavioral equivalents.[1] In addition, the analytic, developmental, or test activity required to supply the answer to the behavioral equivalent is specified and described at some length in Appendices A through N. These descriptions are by no means complete because this is not an HFE textbook; the reader should not feel that he is qualified to perform these analyses as a consequence of merely reading these appendices. Only qualified HFE specialists should be permitted to make the behavioral inputs described herein.

To make the necessary behavioral inputs, a great deal of work described in MIL-H-46855B (Department of Defense, 1979) must be performed. Some of that work, particularly in the very early predesign stages prior to letting contracts, is the responsibility of government planners; some must be performed by behavioral specialists working with development contractors.

---

[1] Questions designated by an asterisk are critical as far a behavioral inputs are concerned because they permit and, indeed, require the full spectrum of behavioral analysis and measurement.

Table 2

DSARC Milestone Questions and Their Behavioral Equivalents

DSARC Question	Behavioral Equivalent
Milestone I--Program Initiation	
1. Define the expected operational environment in which the system will function.	1. Will the anticipated working environment, *including the physical* aspects (weather, illumination, temperature, humidity, ventilation, noise, vibration, ionizing radiation, etc.) and the operational aspects (high density of threat, operational communications, work loads, duty cycles, etc.) adversely effect operator performance? a. What types of effects are to be expected (reduced visual/auditory field, reduced tracking ability, or reduced joint mobility)? b. If the effects are critical to *mission performance*, how can these expected reductions in performance be minimized?
	Activity required: Environmental analysis (Appendix A)
2. Can the expected operational environment be duplicated using existing test facilities?	2. Are any unique facilities necessary for an adequate exercise of those aspects of the operational environment that may adversely impact on operator performance? Are the expected decrements in performance great enough to warrant the use of simulators or environmental chambers, test ranges, etc? a. If so, will it be necessary to exercise the entire system including the crew? b. Can part-task simulators or testing be utilized?
	Activity required: Mission/function/task analysis (Appendix B)
3. Have the operational requirements defined by the CNO been translated into test objectives?	3. Have quantitative criteria of human performance been developed?
	Activity required: Criteria development (Appendix H)
	Has a TEMP (test and evaluation master plan) been prepared and does it provide for HFE testing relevant to the operational requirement and the human performance criteria? a. Are tests identified that will determine how well the human operator has been integrated with the system elements? b. Will tests permit determination of whether operational requirements have been met?
	Activity required: Test planning (Appendix G)
4. Does the proposed program interface with any other program under development? If so, how, and what are the alternatives?	4. Have the human factor problems (if any) encountered in the deployment of similar systems currently in use been identified? Have the human interfaces between programs been determined?
	Activity required: Historical analysis (Appendix L)

Table 2 (Continued)

DSARC Question	Behavioral Equivalent
Milestone I--Program Initiation (Continued)	
*5. Have critical questions and issues to be resolved during the development phase been adequately defined? Are there alternative approaches to resolve critical questions and issues?	5. Are man-machine interfaces defined and areas critical to success of system missions pointed out?
	Activity required: Mission/function/task analysis (Appendix B)
	Are tradeoff studies (such as alternate function allocation schemes or alternate hardware/software designs) for man-machine interfaces been planned and discussed?
	Activity required: Tradeoff studies (Appendix C)
	Are the analyses and activities described in MIL-H-46855B underway?
	Activity required: Timeline and similar analyses; design analysis. (Appendices D and C)
*6. What manner of testing is to be accomplished in order to resolve these questions and issues?	6. What testing technique/procedure most efficiently answers the critical HFE questions and issues (laboratory testing, simulations, operational testing, etc.)?
	Activity required: Test planning (Appendix G)
	Have analysis and studies been accomplished on the equipment design to determine whether the equipment characteristics demand operator performance that exceeds human capabilities or approaches limitation that may significantly jeopardize mission performance, degrade system accuracy, result in excessive maintenance and down time, etc.?
	Activity required: Workload analysis (Appendix E)
7. What test environment, facilities, instrumentation, etc., will be required?	7. Have the test plans been detailed sufficiently to describe the conditions of test, control of variables, data collection techniques, and method of analysis of results? Will the conduct of tests produce results that will identify deficiencies, difficulties, limitations, and shortcomings?
	Activity required: Test planning (Appendix G)
*8. Are operational aspects of critical issues and questions addressed?	8. Have the critical issues and questions regarding the impact of the human operator on system operations been addressed in terms of manning levels, skill levels, work loads, duty cycles, stress, and the extremes of environment?
	Activity required: Manpower/personnel, workload and environmental analyses (Appendices F, E, and A)

Table 2 (Continued)

DSARC Question	Behavioral Equivalent
Milestone I--Program Initiation (Continued)	
*9. Have development milestones been clearly defined?	9. Has the acquisition project specified a schedule of events with sufficient detail to plan HFE activities? Are the milestones to be met through testing attainable within time and money constraints? If not, what alternative plans are being considered? Activity required: Scheduling (Appendix M)
*10. Are technical and operational thresholds or constraints specified where appropriate?	10. What limits have been established for the system in respect to human performance (i.e., detection ranges, lock-on ranges, response or reaction times, update times, etc.)? Are these limits within the range of capability of the human operator? What are the probabilities of system failure resulting from failure to satisfy these limits? Activity required: Criteria development; human performance reliability prediction (Appendices H and J)
11. Has liaison been established with the independent test agency, manufacturers, users, and developing agencies? Has risk technology been adequately determined?	11. Not applicable
*12. Has there been any preliminary or research testing? If so, what were the results?	12. Has preliminary or research testing considered or identified potential HFE areas where additional emphasis could result in improved system performance? If so, what are the results? Activity required: Testing and data analysis (Appendix I)
*13. Are logistic approach and personnel requirements, including training, addressed?	13. Has the human factors engineering role been considered in the ILS (integrated logistic support) plan? What human factors efforts are planned for determining personnel and training requirements with respect to operating and maintenance personnel number and skill level, training aids and devices, and operating and maintenance courses? Activity required: Training analysis (Appendix K)
14. Is the preliminary fleet introduction schedule addressed?	14. Is the deployment date consistent with the planned HFE activities scheduled for DT&E (development, test, and evaluation)? Will all the personnel critical issues and questions be resolved prior to fleet introduction? Activity required: Scheduling (Appendix M)
15. Has an assessment been made of the ways that this proposed system can be defeated or derogated by enemy countermeasures?	15. Not applicable
16. Does the program contain provisions for the incorporation of counter-countermeasures? Has it been shown that the proposed system can be integrated within itself? With other systems?	16. Not applicable

Table 2 (Continued)

DSARC Question	Behavioral Equivalent
Milestone I--Program Initiation (Continued)	
17. Does the technical and operational testing schedule contain specific provisions for the assessment of electromagnetic interference problems?	17. Not applicable
Milestone II--Demonstration and Validation II	
*1. Were the technical questions and critical issues posed at Milestone I adequately resolved?	1. Are all critical functions required of system personnel capable of being performed?
	Activity required: Criteria development and task analysis; testing and data analysis (Appendices, H, B, and I)
2. Have technical and operational thresholds and constraints been refined?	2. Have the limits of the system in respect to human performance been verified (e.g., turnaround times, reacquisition time, loading times, etc.)? Have demonstrations or test data shown that these limits are acceptable for both the human operator and the system? Have test results indicated that the task loading is acceptable for the human operator?
	Activity required: Testing and data analysis (Appendix I)
*3. Has the development testing conducted been sufficient to reasonably ensure that specified requirements can be met?	3. Have the human factors design analyses and tests required by MIL-H-46855B been performed to satisfy personnel requirements implied by system objectives? Do these analyses and tests indicate that specific requirements can be met?
	Activity required: Workload analysis; testing and data analysis (Appendices E and I)
4. Did test sample sizes during the development testing support the required confidence level?	4. Have sufficient numbers of representative subjects (pilots, ground crews, etc.), missiles, targets, environmental conditions, tactical situations, and combinations thereof been employed to provide data to ensure a valid, overall evaluation of total system performance?
	Activity required: Testing and data analysis (Appendix I)
5. Are test data available to back up contract requirements for reliability?	5. Have human error data been used in making predictive estimates of system reliability? Have failure analyses been performed or failure data collected to differentiate between failures due to equipment alone, man-equipment incompatabilities, and human error?
	Activity required: Human performance reliability prediction (Appendix J)
6. Is there any development testing or evaluation of development testing accomplished still to be completed? If so, what impact could surprises have on the next phase?	6. What new personnel problems have been identified as a result of the developmental tests? What is the significance of these to the full scale development decision? Have previous analysis and testing of man-machine interfaces provided sufficient information to select the best arrangement in keeping with program costs and direction?
	Activity required: Test planning; testing and data analysis (Appendices G and I)

Table 2 (Continued)

DSARC Question	Behavioral Equivalent
Milestone II--Demonstration and Validation (Continued)	
7. If there are untried materials employed anywhere in the system or its components, what tests have been conducted to determine the ability of these materials to adequately withstand the operating and environmental conditions expected?	7. Not applicable
8. Where critical materials are employed to maximize performance, has performance with more readily available materials been tested and assessed?	8. Not applicable
9. Are there other advances in technology that could be introduced to enhance equipment performance? Does such performance enhancement buy additional capability that is needed? Is it feasible or cost effective to attempt to introduce these advances?	9. Has a survey of the state-of-the-art been conducted to identify new and innovative man-machine interface devices and techniques? If so, have studies been performed to determine their applicability to the system under development? Activity required: Design analysis (Appendix C)
10. Have all inter- and intrasystem interfaces been sufficiently defined to ensure compatability (hardware and software)? Are other systems involved that are not service approved? If so, what effect would their nonavailability have on the effectiveness of the proposed system? What degree of risk is anticipated in their availability/compatability? Has operational testing been planned to demonstrate the integrated performance of the major system with its internal subsystems and interface?	10. Is the contractor-furnished equipment that provides the hardware interface with the operator/maintainer qualified or otherwise certified? If not qualified, what studies should be performed to either qualify it or quantify the effect on operator/maintainer performance, cost, and delivery time? Have the human factors plans been detailed for incorporation into the operational test plans for the system integration tests? Activity required: Test planning (Appendix G)
11. Has an initial integrated logistic support program been defined to support operational needs? Have fleet introduction and life cycle costs been refined?	11. Have the human factors aspects of logistics such as maintenance, training, manuals, and personnel been defined and incorporated into the ILS program plan? How do these aspects effect the deployment date and life cycle cost? Activity required: Integration analysis (Appendix N)
*12. Have the critical questions and issues to be resolved prior to Milestone III been refined and adequately addressed?	12. Have the critical human factors questions and issues to be resolved prior to Milestone III been refined and adequately addressed? What additional questions or issues need to be studied before the production decision? Activity required: Test planning (Appendix G)

Table 2 (Continued)

DSARC Question	Behavioral Equivalent
Milestone II--Demonstration and Validation (Continued)	
13. Are test programs and objectives to address these questions and issues established? Are the sample sizes specified and are they sufficient to provide the level of confidence desired? Will successful accomplishment of test programs demonstrate technical and operational suitability?	13. Have human factors engineering requirements been incorporated into the updated TEMP (test and evaluation master plan)? Does the revised TEMP cover new issues that may impact on the system operator/maintainer? Do the human factors tests specify the test conditions, parameters to be measured, and method of data analysis? Are the tests detailed sufficiently to provide data to demonstrate technical and operational suitabiity?
	Activity required: Test planning (Appendix G)
14. Have the following requirements been jointly identified by the developing agency and COMOPTVEFOR (Commander of Operational Test and Evaluation Force)? a. Test programs, participants, schedules. b. Number of test units. c. Test environment. d. Facilities, targets, countermeasures. e. Fleet personnel and training. f. Test bed. g. Installation. h. Instrumentation. i. Data acquisition, reduction, and evaluation. j. Miscellaneous contractor support. k. Funds for the above items.	14. Have the human factors test requirements in respect to items (a) through (k) been acknowledged by the program coordinator and the independent testing agency? Have these items been incorporated into the TEMP? Activity required: Test planning (Appendix G)
15. Has the developmental and operational testing schedule been outlined? Does it preclude duplication? Does it permit at the earliest possible time an initial phase of operational testing? How representative of anticipated follow-on production units are the units to be tested?	15. Does the human factors input to the TEMP outline the tests to be performed during the advanced and engineering development of the system? Are the tests outlined for the initial and follow-on test and evaluation different from those described for the developmental tests? From the operator/maintainer point of view, how representative of anticipated follow-on production units are the units to be tested?
	Activity required: Test planning (Appendix G)
Milestone III/IIIA--Full-scale Engineering Development/Production and Deployment	
*1. Have all Milestone II critical questions and issues been adequately answered?	1. Do any of the problem areas or issues on system operation that require critical human performance, identified during developmental testing, remain unresolved at this time?
	Activity required: Test planning (Appendix G)
2. Which components involve new state-of-the-art or new state-of-manufacture? What production testing remains to be accomplished on these? Is there an alternate or back-up component, material, or system?	2. Not applicable

Table 2 (Continued)

DSARC Question	Behavioral Equivalent
Milestone III/IIIA--Full-scale Engineering Development/Production and Deployment (Continued)	
3. Has a complete data package been prepared and tested to determine its adequacy to support production, acceptance (inspection), operation, and maintenance of the system?	3. Has the data package documented the type of testing, test results, and conclusions in terms of ease of unambiguous operation, high reliability, ease of checkout, removal, and replacement? Is the documentation comprehensive enough to support the decision to accept or reject the system? Activity required: Testing and data analysis (Appendix I)
4. If there are contract requirements for reliability and maintainability, have any test requirements to support these been waived? Which? Have the reliability and maintainability been adequately demonstrated in an operational rather than laboratory environment?	4. Have the behavioral tests of user acceptability or operability been integrated with the reliability and maintainability tests? Have these tests been performed in a simulated or actual operational environment utilizing service personnel? Activity required: Testing and data analysis (Appendix I)
5. Is there any developmental test and evaluaiton (or evaluation of tests accomplished) still to be completed? If so, what impact could surprises here have on the production program?	5. Have all the HFE tests provided sufficient data for resolving the critical issues? Have new problems been identified as a result of the developmental tests? If so, what is the significance of the problem to system production? Activity required: Testing and data analysis (Appendix I)
*6. Have all significant problems found in development been studied and fixes designed? To what extent have they been?	6. Have fixes to behavioral problems, identified during developmental testing, been validated by further testing in a simulated or actual operational environment? Activity required: Testing and data analysis (Appendix I)
7. To what extent has there been testing with other systems that must have a technical interface with the new systems? What were the results?	7. Not applicable
8. What operational testing has been accomplished to date on the system or its major components and what has it revealed? What type of the hardware/software was used? Was it development prototype, and how similar is it to the expected production article?	8. Has analysis of the developmental or operational test data been performed to determine the status of critical behavioral issues identified in the advanced development and engineering model? Have studies been conducted to determine the correlation of behavioral test data obtained from an engineering development model to the final production model? Activity required: Testing and data analysis (Appendix I)
9. Wherein did the physical and stress environments in which tested differ from those that could be expected when operationally deployed?	9. Were the behavioral tests designed to duplicate or simulate the anticipated operational environment and mission variables? What differences exist between the environment in which the behavioral tests were performed and the expected operational environment? What impact will the differences have on the operational use of the system? Activity required: Testing and data analysis (Appendix I)

Table 2 (Continued)

DSARC Question	Behavioral Equivalent
Milestone III/IIIA--Full-scale Engineering Development/Production and Deployment (Continued)	
*10. What user-type participation was employed in the OT&E (operational test and evaluation)? In what respect did the personnel participating in the test differ from those who could be expected to operate and maintain it? What needs were revealed for special requirements with respect to personnel qualifications, training, or conditions for maintaining, such as accessability of components, special tools, etc.?	10. Did the operational tests demonstrate that the systems can be effectively operated and maintained by the level of personnel skill, manning levels, work loads, and duty cycles anticipated to be available under service conditions? Have the demonstrations affected the training plan and maintenance concepts? In what respect did the personnel participating in the test differ from those who could be expected to operate and maintain it? What needs were revealed for special requirements with respect to personnel qualifications, training, or conditions for maintenance, such as accessibility of components, special tools, etc.? Activity required: Testing and data analysis (Appendix I)
11. Wherein did the tests depart from being two-sided or otherwise realistically representative of the operational system and enemy reaction?	11. Did the tests demonstrate that the operator/maintainer can perform the required task on his own without help from outside advisors or at least with only that support that would be available in the operational environment? Activity required: Testing and data analysis (Appendix I)
12. If the system must interface closely with other fleet systems, has the employment doctrine been developed and tested? If not, what analysis has been made to ensure no seriously degrading impact will occur? Has operational testing been conducted to demonstrate the integrated performance with other fleet systems?	12. See item 8. Activity required: Test planning (Appendix G)
13. Has dependence of the program on supporting developments been assessed and are there provisions to monitor the status of these supporting developments so as to ensure early detection of deleterious trends?	13. Not applicable
14. If the system is part of a larger major weapon system, how were its interfaces with other elements of the major weapon system tested?	14. Have modeling, simulation, and/or other tests concerned themselves with subsystem interfaces in respect to special requirements in terms of personnel skill and manning levels, function allocation, etc.? Activity required: Test planning (Appendix G)
15. What confidence has been established by the OT&E as to productibility, reliability, and maintainability? What data supports the O&M (operational and maintenance) cost estimates? What analysis has been accomplished to estimate the degradation field use will cause, and what testing is scheduled to confirm the estimate?	15. Have the planned OT&E tests been designed to produce human performance reliability and maintainability data? Are test conditions such as to reproduce the stresses of operational field activity? Activity required: Test planning (Appendix G)

Table 2 (Continued)

DSARC Question	Behavioral Equivalent
Milestone III/IIIA--Full-scale Engineering Development/Production and Deployment (Continued)	
16. Has the technical/maintenance data package been finalized and have the operational and maintenance portions been tested and evaluated by representative user-type personnel? Have provisions been made to promptly incorporate in the maintenance manuals any changes incorporated in the system?	16. Have tests been planned or performed to evaluate the adequacy of technical/maintenance data from a personnel standpoint, using representative Navy users?
	Activity required: Test planning; testing and data analysis (Appendices G and I)
17. What are the countermeasures of greatest potential concern? What tests have been conducted in the presence of these countermeasures? Check for representatives of countermeasures used in terms of technical validity and density of countermeasures deployment.	17. Not applicable
18. If a tactical missile, gun, bomb, etc., against what targets representative of what threat targets has it been tested and is it to be further tested?	18. See item 8.
	Activity required: Test planning; testing and data analysis (*Appendices G and I*)
*19. Are there critical questions concerning the effectiveness of the system as it is used under actual operational conditions?	19. Have all the critical issues identified in the TEMP concerning the operator/maintainer role in system effectiveness been resolved? Are there any new or additional issues that should be tested prior to production?
	Activity required: Testing and data analysis (Appendix I)
20. What comparative tests have been conducted with existing systems that are planned to be replaced? What were the results?	20. If such tests have been conducted, did these tests involve behavioral comparisons?
	Activity required: Test planning; testing and data analysis (Appendices G and I)
*21. Does the system have degraded modes of operation? Were the modes tested sufficiently and realistically enough to validate their adequacy? What were the results of tests conducted in these modes?	21. Did any tests with degraded modes of operation involve personnel-related factors (e.g., loads, stress, environmental conditions)? Did the results of these tests have an impact on the production decision?
	Activity required: Testing and data analysis (Appendix I)
22. What tests and assessment of storage life of the system, both shelf life and service life, have been made? Are degrading effects acceptable? Are corrective improvements planned? What are the implications of changes planned, if any?	22. Not applicable
23. Has the test item been tested in an "all-up" system? If this is not feasible at this state, have the individual increments been tested and an analysis conducted of how an "all up" system would perform?	23. Not applicable

Table 2 (Continued)

DSARC Question	Behavioral Equivalent
Milestone III/IIIA--Full-scale Engineering Development/Production and Deployment (Continued)	
24. Has the system been tested with the ancillary equipment proposed for service use, such as special test sets, automatic test, dummy loads, data links, special assembly equipment, etc.?	24. See item 8. Activity required: Testing and data analysis (Appendix I)
25. Is the test item in any way terrain, visibility, or weather sensitive? If so, were the marginal areas of performance thoroughly tested?	25. Have the effects of terrain, visibility, or weather, upon the operator's performance been adequately determined? Activity required: Testing and data analysis (Appendix I)
26. What tests were conducted for environmental effects (temperature, humidity, water, snow, ice, vibration, shock, dust, mud, EMI (electromagnetic interference), etc.)? What deviations were necessary from conditions representative of the conditions the item will encounter in actual practice? What are the effects on system operation and maintenance?	26. What tests were conducted to establish the ability of the system and checkout/support equipment to be operated/maintained reliably by system personnel under various ground, sea, and air environments representative of the operational environment? Activity required: Testing and data analysis (Appendix I)
27. What tests of system safety were performed? If safety certification is required, have necessary tests been performed and adequate procedures prepared for loading, assembly, storage or stowage, handling, operational use, testing for contaminants, leakage, emergency disposal, etc.? Does the system present hazardous mechanical, electrical, chemical, explosive, acoustic, radiation, or other effects to maintenance or operational personnel, and has adequate provision been made therefore?	27. Were the appropriate criteria specified in relation to the operator/maintainer environment in a system safety program? Activity required: Criteria development (Appendix H)
*28. Are there any logistic or manpower problems that would result from planned fleet introduction and deployment of these systems?	28. Are there any logistic problems involving manpower, availability, skilled personnel, training, or trainers? Activity required: Manpower/personnel analysis; training analysis (Appendices F and K)
*29. Has adequate operational testing and evaluation been accomplished to date on the system or its major components to reach the production decision?	29. Has this testing included behavioral elements? Activity required: Testing and data analysis (Appendix I)
30. What additional OT&E is programmed and when? Does it include provisions of updated and projected operational/threat environment to ensure that the production hardware will be operationally effective and suitable?	30. Has the TEMP been updated to include HFE operational tests in actual and simulated combat environments utilizing the limited production of full production model? Does the current version of the TEMP provide for follow-on HFE testing? Activity required: Test planning (Appendix G)
31. Has ILS again been addressed to determine adequate life-time support and cost for this system?	31. Have sufficient data been obtained from developmental tests to provide valid estimates of logistic costs including personnel factors? Activity required: Testing and data analysis (Appendix I)

DISCUSSION

The picture presented in this report of the spectrum of behavioral inputs to WSAP is somewhat abstract and idealistic. That is, these are inputs that could and would be made, provided a sufficiently high level of WSAP management in DoD and the Navy really wished to receive them and act upon them. Active implementation of these inputs will be hampered if the following conditions continue to exist:

1. WSAP management remains indifferent to the necessity of incorporating behavioral inputs into the process. A firm commitment on the part of that management is required if these inputs are to be made. This may require adding behavioral specialists to some of the Navy staffs responsible for generating and reviewing WSAP documents. In the past, DoD management has given lip service to the incorporation of behavioral inputs in the WSAP but has failed to follow through with substantive support. Project managers often are not inclined to allocate funds for the performance of human factors analyses and evaluations. There have been instances where minimal funding was provided for these activities and later cancelled at the first sign of budget duress. In general, WSAP management has been highly resistant to behavioral work, despite the prodding of such agencies as the General Accounting Office (1981). However, recent interest expressed by the Chief of Naval Material in incorporating behavioral inputs into the WSAP is indeed encouraging.

2. Given that managers receive behavioral inputs, it is necessary for them to implement (e.g., base their decisions on) at least some of the recommendations made in these inputs. If such inputs are to be ignored, there seems to be little point in providing them in the first place. If, however, managers are provided these inputs in a timely fashion, there will be an impulse to act upon at least some of them.

3. Behavioral inputs must be technically sound and must be presented to WSAP managers and decision makers in sufficient time for the latter to consider them. Despite some inadequacies in behavioral technology, the analyses and outputs described can be performed reasonably effectively and within required time constraints, provided that the personnel given responsibility for them are technically qualified. In the past, technically unqualified personnel have been assigned the task of providing behavioral inputs with very depressing results. However, problems concerning technical competence occur frequently.

RECOMMENDATIONS

Personnel involved in the DSARC decision-making process should help ensure that the behavioral inputs described in this report are provided at the specified DSARC milestones.

REFERENCES

Baker, C. C., Johnson, J. H., Malone, M. T., & Malone, T. B. Human factor engineering for Navy weapons system acquisition. Alexandria, VA: Essex Corporation, July 1979.

Bureau of Naval Personnel. The Bureau of Naval Personnel new development human factors program (Report ND 64-51). Washington, DC: Author, New Developments Research Branch, Personnel Research Division, February 1964.

*Bureau of Naval Personnel. Manual of Navy officer manpower and personnel classifications. Volume I: Major code structures (NAVPERS 15839C). Washington, DC: Author, 7 July 1975.

*Bureau of Naval Personnel. Manual of manpower and personnel classifications and occupational standards. Section I: Navy enlisted occupational standards (NAVPERS 18068D). Washington, DC: Author, September 1975.

*Bureau of Naval Personnel. Manual of manpower and personnel classification and occupational standards. Section II: Navy enlisted classifications (NAVPERS 18068D). Washington, DC: Author, January 1981.

Condon, C. F. M., Hayes, J. F., Turner, W. S., & Walder, G. L. Guide for conducting personnel and training research in support of advanced ship programs (Rep. PTB-70-7). Alexandria, VA: Matrix Corporation, May 1970.

Department of Defense. Human engineering design criteria for military systems, equipment and facilities (MIL-STD-1472B). Washington, DC: Author, 31 December 1974.

Department of Defense. Major system acquisitions (DoD Directive 5000.1). Washington, DC: Author, January 1977. (a)

Department of Defense. Major system acquisition process (DoD Directive 5000.2). Washington, DC: Author, January 1977. (b)

Department of Defense. Test and evaluation (DoD Directive 5000.3). Washington, DC: Author, January 1977. (c)

Department of Defense. Human engineering requirements for military systems (MIL-H-46855B). Washington, DC: Author, 31 January 1979.

Department of Defense. Human engineering design criteria for military systems equipment and facilities (MIL-STD 1472C). Washington, DC: Author, 2 May 1981.

General Accounting Office. Effectiveness of U.S. forces can be increased through improved system design (Rep. PSAD-81-17). Washington, DC: Author, 29 January 1981.

Greer, C. W. Analyst's guide for the analysis sections of MIL-H-46855. Seattle: Boeing Aerospace Company, 1976.

Note. References marked with an asterisk are cited in appendices only.

Greer, C. W. Navy manager's guide for the test and evaluation sections of MIL-H-46855. Seattle: Boeing Aerospace Company, 1977.

HARDMAN Project Office, Chief of Naval Operations (OP-112). Documented description of the weapon system acquisition process (Rep. HR 79-01). Washington, DC, 25 July 1979. (a)

HARDMAN Project Office, Chief of Naval Operations (OP-112). OP-01 HARDMAN actions and tasks for surface ship acquisitions (Report HR 79-02). Washington, DC: Author, 28 September 1979. (b)

*HARDMAN Project Office, Chief of Naval Operations (OP-112). Training requirements determination handbook. Volume I: Overview. Volume II: Training concept. Volume III: Training resource requirements. Volume IV: Total ship system (Rep. PAR-059-80). Washington, DC: Author, 17 November 1980.

Holshouser, E. L. Translation of DSARC milestones into human factors engineering requirements (Tech. Publication TP-75-58). Point Mugu, CA: Pacific Missile Test Center, 30 September 1975.

Holshouser, E. L. Guide to human factors engineering general purpose test planning (GPTP) (Tech. Publication TP-77-14). Point Mugu, CA: Pacific Missile Test Center, 30 June 1977.

Malone, R. B., Gloss, D. S., & Eberhard, J. W. Human factors techniques employed in deriving personnel requirements in weapon system development (Rep. PRR 68-3). Alexandria, VA: Matrix Corporation, October 1967.

Meister, D. Human factors evaluation in system development. New York: Wiley, 1965.

Meister, D. Human factors: Theory and practice. New York: Wiley, 1971.

Price, H. E., Fiorello, M., Lowry, J. C., Smith, M. G., & Kidd, J. S. Department of defense and service requirements for human factors R&D in the military system acquisition process (Research Note 80-23). Alexandria, VA: U.S. Army Research Institute, DoD Human Factors Engineering (HFE) Technical Advisory Group (TAG), July 1980. (a)

Price, H. E., Fiorello, M., Lowry, J. C., Smith, M. G., & Kidd, J. S. The contribution of human factors in military system development: Methodological considerations (Tech. Rep. 476). Alexandria, VA: U.S. Army Research Institute, July 1980. (b)

*Swain, A. D. A method of performing a human factors reliability analysis (Rep. SCR-685). Albuquerque, NM: Scandia Corporation, 1963.

*Swain, A. D., & Guttman, H. E. Handbook of human reliability analysis with emphasis on nuclear power plant applications (NUREG/CR 1278). Washington, DC: U.S. Nuclear Regulatory Commission, 1980.

*Van Cott, H. P., & Kinkade, R. G. (Eds.). Human engineering guide to equipment design (Rev. ed.). Washington, DC: U.S. Government Printing Office, 1972.

*Woodson, W. E. Human factors design handbook. New York: McGraw-Hill, 1981.

Note. References marked with an asterisk are cited in appendices only.

APPENDIX A

ENVIRONMENTAL ANALYSIS

ENVIRONMENTAL ANALYSIS

Environmental analysis is the determination of the effect of an environmental condition on personnel and system performance. The major conditions examined may be any of the following: noise, temperature, vibration, acceleration, lighting, weather and atmospheric conditions, and sea state. The acceptable range of conditions (i.e., the range of values that poses no hazard to personnel or their performance) can be determined from handbooks on human factors (c.f., Van Cott & Kinkade, 1972; Woodson, 1981). For example, extended exposure to a noise level greater than 90 db will require ear protection; levels greater than 135 db will be hazardous to hearing of personnel without ear protection even for brief exposures.

In performing this analysis, the environment in which the system will be operated is examined to determine whether unacceptable environmental conditions exist because of the way in which the equipment is designed or must be operated (mission requirements). Either the unacceptable environment can be changed by changing equipment design or location (also, but less likely, by changing the system mission) or by developing or using some means of protecting against that environment. In cases where the environment is not potentially lethal but could result in degradation of personnel performance, the analyst should, if possible, predict the amount of degradation and its probable effect.

In preparing an input to the DSARC committee, the following items should be covered:

1. Description of the environment in which the new system will be used and those conditions that will impact significantly on system personnel.

2. Quantitative values of the environmental conditions.

3. Quantitative acceptable and nonacceptable values for impact of specified environmental conditions on personnel.

4. Anticipated effect of environmental conditions on personnel and system performance.

5. In cases where new technology creates potential new risks for personnel, the types of data/research required to evaluate hazards and protection requirements should be indicated.

6. Recommendations for action to be taken.

APPENDIX B

MISSION/FUNCTION/TASK ANALYSIS

MISSION/FUNCTION/TASK ANALYSIS

In developing any man-machine system (MMS), questions dealing with the allocation of functions between personnel and equipment, analysis of tasks, and identification of man-machine interfaces must be answered. The answers are obtained by analyzing mission requirements and specifying the functions to be performed by system personnel. Mission and function analyses are performed early in the feasibility/analysis stage of system development and serve primarily as prerequisites to task analysis. Function flow diagrams should be available by DSARC I and the results of the task analysis by DSARC II at the latest.

The process of performing a mission/function/task (MFT) analysis can be summarized as a series of steps, although each step involves a number of subordinate activities.

1. Analyze system requirements. To implement system requirements, certain functions are necessary. For example, if a ship is to perform an ASW mission, it must detect and locate submarines, steer a particular track, attack with available weapons, etc. Each such activity or function also implies more molecular functions that must be performed if the more molar (superordinate) function is to be accomplished.

2. Determine system functions. For each system mission (e.g., drop bombs, perform ASW patrol), the individual major operations that must be performed to implement the mission are listed sequentially. The resultant functions are described in the form of a function flow block diagram (FFD) (see Figure B-1). The effect on system functions of any environmental factors, performance requirements, and constraints are then determined. For example, if the system must perform in extremely cold weather (e.g., 10° below zero), what effect will the cold environment have on how the system actually performs? In particular, what effect will the requirement have on personnel performance? When additional functions are required by this analysis, they are inserted into the FFD. New functions are determined by specifying the inputs to and outputs from each already available function; the input actions are those required to initiate the function and the output actions result from performance of the function. For example, to perform the function of submarine detection, the operator must first scan the sonar scope. Therefore, scanning becomes a separate but subordinate (more molecular) function to the detection function. One output of the detection process is a verbal report that the submarine has been detected; the verbal report describes a communication function. Each function is examined in terms of alternative ways in which inputs and outputs can be supplied. These may eventually become design alternatives. All feedback loops are specified.

3. Allocate functions between men and machines. The allocation process is actually one step in establishing design alternatives. Although the original system concept usually has already implied certain function allocations, others are still undetermined. The function allocation process is designed for these as yet unspecified alternatives.

For those system functions that have not yet been allocated, operator and equipment functions are differentiated by describing all the possible ways in which mission objectives can be implemented (within the general categories of automatic, semiautomatic, and manual). Each alternative is then examined to ensure that it can satisfy systm requirements. Then, the msot cost-effective alternative is selected by a complex process that has been described in detail by Meister (1971). In brief, the process requires the selection of evaluative criteria (e.g., cost, reliabiltiy, producibility, etc.) including those relating to behavior (e.g., operability, maintainability, and training). Each of these criteria is given a weight corresponding to its perceived importance. Each design alternative is then judged against every other in terms of these critera and an appropriate weight for that alternative is assigned. The weights, when added together, provide a numerical value for each alternative that can be compared. For example, if alternatives I, II, and III (corresponding to, let us say, automatic, semiautomatic, and manual modes) receive values of 6.8, 7.3, and 4.4 respectively, alternative II should be implemented.

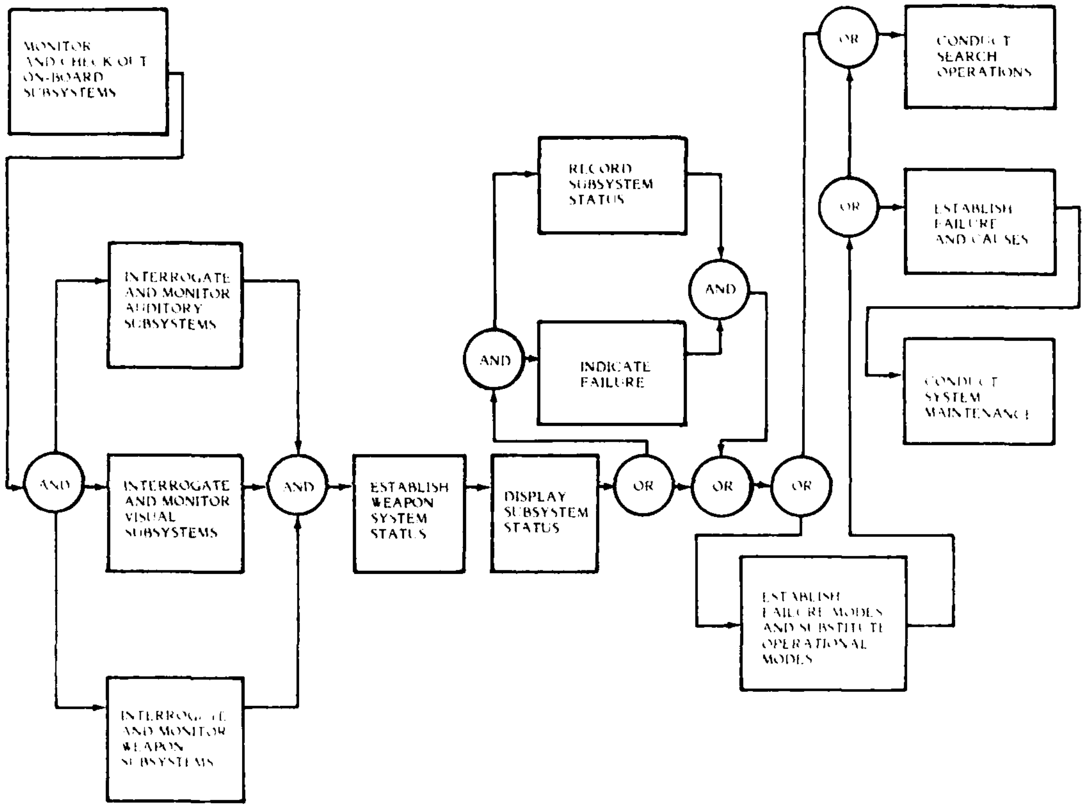


Figure B-1. Sample function flow diagram for weapon system (Modified from Greer, 1976).

4. Determine and describe the task. Because of the complexity of task analysis (TA), no abbreviated description of the process can be entirely satisfactory. Hence, the reader is urged to consult Baker, Johnson, Malone, and Malone (1979) (pp. 3-14 to 3-26).

A task is an action taken to implement a system function. It is defined by the immediate purpose for an action, a machine output or consequence of that action, and the human inputs, decisions, and outputs necessary to initiate the action and accomplish its purpose.

The starting point for TA is the list of functions developed during MFT. For each function, it is necessary to list in sequence all the actions that must be performed by system personnel to implement that function. The identification of the task poses few problems. What may be difficult is deciding on the level of detail to be described; that is, the analyst must decide whether he wishes to describe the action (task) at the level of the individual control or display (e.g., read meter) or at the somewhat more molar level (e.g., determine that fuel level is adequate). The particular level used is arbitrary, but it is best to be as detailed as possible.

5. Analyze the task. After the task has been described, it becomes necessary to analyze it by drawing certain inferences from the task description. To assist in this analysis, which is performed in terms of the demands imposed by the task on the operator/maintainer, analytic tools such as the operation sequence diagram (OSD), may be developed. The OSD provides a graphic display of task component interrelationships, thus making such interrelationships more visible. Ultimately, however, it is necessary to ask the questions listed below (taken from Meister, 1971) about each task. The answers to these questions indicate either that the system poses no behavioral problems or that a potential problem exists that must be investigated further. It should be noted that the design alternative(s) selected are evaluated for operability and maintainability effectiveness by answering relevant questions.

a. Functions/Tasks.

(1) Are functions/tasks to be performed within operator capability? Consider requirements in the following functions:

(a) Sensory/perceptual.
(b) Motor.
(c) Decision-making.
(d) Communication.

(2) Do task characteristics impose excessive demands on the operator?

(a) Task duration (possible fatigue effects?).
(b) Frequency of task performance (possible fatigue effects?).
(c) Information feedback (insufficient operator guidance?).
(d) Accuracy (too demanding?).
(e) Error probability (supportable?).
(f) Error criticality (effect on task performance?).
(g) Concurrent multitask requirements (effect of one task on another?).

b. Environment.

(1) Events requiring operator response:

(a) Speed of occurrence (too fast?).
(b) Number (too many?).
(c) Persistence (too short-lived?).
(d) Movement (excessive?).
(e) Intensity (too weak to perceive?).
(f) Patterning (unpredictable?).

(2) Physical effects:

(a) Temperature, humidity, noise, vibration (excessive?).
(b) Lighting (substandard or special effects?).
(c) Safety (problems?).

(3) Mission conditions:

(a) Potential emergencies (can operator recognize and overcome rapidly?).
(b) Mission response characteristics:
  1. Accuracy requirements (excessive?).
  2. Speed requirements (excessive?).

(4) Event criticality (effect on error probability?).

c. Equipment.

(1) Display information requirements:

(a) Too much to assimilate?
(b) Difficult to perceive/discriminate/track?
(c) Require excessively fast operator response?
(d) Too much memory required?

(2) Control requirements:

(a) Excessively fine manipulations required?
(b) Too much force required?
(c) Must be responded to too rapidly?
(d) Too many to perform in sequence?

MFT analysis is an integral part of the development process. In providing a DSARC input, the following--derived from the MFT analysis--should be emphasized:

1. Statement that analyses required by MIL-H-46855B have/have not been performed or are/are not underway.

2. Specification of which analyses (e.g., workload, information, task) have been or will be performed; if not to be performed, indicate why.

3. Adequacy of the system design in general to satisfy system/mission requirements. Exceptions to the previous statement.

4. Man-machine interface problem areas (critical questions and issues to be resolved) discovered by performing the MFT analysis.

5. Recommendations for action, where required; risk and costs involved.

APPENDIX C

DESIGN ANALYSIS

DESIGN ANALYSIS

Introduction

Design analysis is defined here as those analyses relating to the design, selection, and evaluation of specific hardware and software mechanisms for exercising the system, including procedures and methods for using the hardware and software. Specifically included are:

1. The development and evaluation of operating and maintenance procedures.

2. The design and evaluation of hardware and software man-machine interfaces (e.g., control panel, methods of accessing software files).

3. Human factors-related tradeoff studies between alternative design configurations or procedures.

Timeline and workload analyses (described in Appendices D and E) are related to design analysis because they seek to determine how personnel will be influenced by design characteristics.

Procedures

MIL-H-46855B (Department of Defense, 1979) requires the developer to apply human engineering principles and criteria to the development of operating and maintenance procedures based on human performance functions and tasks. In actual practice, almost all procedures are developed by the engineer responsible for designing the hardware/software with which the procedure is to be used. Behavioral design analysis enters the picture when the behavioral specialist reviews the procedures in draft form to determine and eliminate:

1. Discrepancies between the procedure and the requirements of the equipment design that could lead to error or inoperability.

2. Failure to include all equipment-required operations in the procedure; frequently the written procedure overlooks certain operating requirements that could again lead to error or operational failure.

3. Informational ambiguities that can result in misinterpretation of procedural requirements; when procedures are unclear, the probability of error is increased.

4. Excessive demands on the operator, where procedures require (a) excessive strength, speed, and/or frequency of response, (b) excessive perceptual capability, or (c) accuracy/precision of responses.

5. Lack of feedback information needed if the operator is to know that he is performing correctly.

The analysis is performed by examining the procedure to determine its behavioral elements and comparing those elements with the behavior that can reasonably be expected of the operator. Although this comparison involves considerable intuitive judgment, it is possible to make use of lists of relative capabilities (see Table C-1).

Table C-1

Relative Capabilities List

Man	Machine
Can monitor low probability events not feasible for automatic systems because of number of events possible	Limited program complexity and alternatives; unexpected events cannot be handled adquately
Absolute thresholds of sensitivity are very low under favorable conditions	Generally not as low as human thresholds
Can detect masked signals effectively in overlapping noise spectra	Poor signal detection when noise spectra overlap
Able to acquire and report information incidental to primary activity	Discovery and selection of incidental intelligence not feasible in present designs
Not subject to jamming by ordinary methods	Subject to disruption by interference and noise
Able to recognize and use information, redundancy (pattern) of real world to simplify complex situations	Little or no perceptual constancy or ability to recognize similarity of pattern in spatial or temporal domain
Reasonable reliability in which the same purpose can be accomplished by different approach (corollary of reprogramming ability)	High reliability may increase cost and complexity; particularly reliable for routine repetitive functioning
Can make inductive decisions in new situations; can generalize from few data	Virtually no capacity for creative or inductive functions
Computation weak and relatively inaccurate; optimal game theory strategy cannot be routinely expected	Can be programmed to use optimum strategy for high-probability situations
Channel capacity limited to relatively small information throughput rates	Channel capacity can be enlarged as necessary for task
Can handle variety of transient and some permanent overloads without disruption	Transient and permanent overloads may lead to disruption of system
Short-term memory relatively poor	Short-term memory and access time excellent
Can tolerate only relatively low imposed forces and generate relatively low forces for short periods	Can withstand very large forces and generate them for prolonged periods
Generally poor at tracking though satisfactory where frequent reprogramming required; can change to meet situation.	Good tracking characteristics over limited requirements
Performance may deteriorate with time because of boredom, fatigue, or distraction; usually recovers with rest	Behavior decrement relatively small with time; wear maintenance and product quality control necessary
Relatively high response latency	Arbitrarily low response latencies possible
Relatively inexperienced for available complexity and in good supply; must be trained	Complexity and supply limited by cost and time; performance built in
Light in weight; small in size for function achieved; power requirement less than 100 watts	Equivalent complexity and function would require radically heavier elements, enormous power, and cooling resources
Maintenance may require life support system	Maintenance problem increases disproportionately with complexity

Note: Taken from Baker C. C., Johnson J. H., Malone M. T., & Malone, T. B. Human factor engineering for Navy weapons system acquisition. Alexandria, VA: Essex Corporation, July 1979, p. 3-11.

Human Engineering of Design

MIL-H-46855B also requires that human engineering principles and criteria be applied during detail design to equipment drawings. (See Meister, 1965, for a detailed description of human engineering activities during system development.)

The tool most often used in evaluating design is the checklist, which is a series of written statements that describe the individual characteristics that an equipment ought to have to be properly human engineered. The checklist is a highly condensed form of MIL-STD-1472C (Department of Defense, 1981). The items found in the checklist represent a selection from the total population of equipment characteristics that might affect performance on the equipment under consideration. They are selected on the basis of their presumed effect on operator performance.

Checklist use is quite simple. The drawing (or the mockup or the prototype equipment) is examined to see if it has the set of characteristics included in the checklist. If any drawing or equipment characteristic is not in accordance with the standard specified in the checklist, a potential human engineering problem exists and must be resolved.

Human Factors Tradeoff Studies

All design is one tradeoff after another. Opposing considerations are balanced against each other, and a decision is reached in favor of one or the other. However, at certain points in the development of a system, a decision as to which of several alternate feasible approaches must be followed is necessary for further system definition.

A tradeoff, of course, is almost never between just two factors. Every factor has implications for other factors so that, if the tradeoff is to be properly analyzed, the latter must be brought into the analysis; for example, in choosing between two ways of performing an equipment function, the engineer may select one method because it has higher reliability, but he will have to balance the higher reliability against the additional cost of that reliability. Again, an engineer may prefer one method over another because he will have to balance this consideration against higher cost and the fact that more automatic equipment generally requires more maintenance.

There are, of course, no absolute tradeoff priorities. Every tradeoff is performed in the context of a specific equipment design problem. Although, for experimental purposes, one can develop abstract tradeoff problems, it has been found that engineers need the equipment development context to assign an appropriate weighting to the factors. Priority number one in one design situation may be third or fourth in another.

There is, of course, a kind of tradeoff in which the weight of the evidence so clearly favors one alternative over another that the decision in favor of that alternative is irresistible. An example of this is the decision between having a man lift a 500-pound weight unaided and having that weight lifted by automatic lift. This decision is obvious. Tradeoff studies must be performed only when the conclusions to be drawn from the available data are obscure or when they do not point logically to a solution, or when data are insufficient. Of course, the problem at issue must be important enough to warrant a study; the decision between two types of controls (e.g., a toggle switch or rotary) or their location on a panel would not ordinarily warrant a study.

Although relatively few design tradeoffs are centered around human factors variables, many of them have implications for personnel functioning and therefore require behavioral inputs. The following criteria should be applied to determine when a problem requiring a tradeoff study requires such inputs:

1. Will the decision reached or the solution selected have significant effects on:

   a. The number of personnel required to operate and/or maintain the equipment?

   b. The skill level of personnel required to operate and/or maintain the equipment?

   c. The amount of training these personnel would need?

   d. The manner in which they would be required to perform (i.e., the nature and difficulty of their tasks; this in turn would affect the efficiency with which personnel can perform their tasks)?

   e. Their safety?

2. Will a major design requirement impose a constraint on the number and type of personnel (e.g., is the developer required to design so that two men can operate four pieces of equipment?)? The question that must be answered is what effect this constraint will have on design.

3. Will the problem involve the capability of personnel to perform a job? The question to be answered is whether personnel will be overloaded by a particular design solution.

In most cases, the tradeoff study will not be experimental in nature (i.e., require the gathering of laboratory-controlled data). Rather, it will be a largely logical, systematic examination of alternatives using available data, an attempt to anticipate the potential consequences of design factors. Therefore, the need for formal tradeoff studies (i.e., those consciously and deliberately performed) will not be frequent. Such studies are performed relatively early in system development (e.g., in predesign) because major problems requiring tradeoff studies occur primarily in early design. At later stages, the problems have been solved, or design has proceeded so far that only a major perturbation would require such studies.

There are no special techniques for the tradeoff study. The general steps for performing any tradeoff study (regardless of its degree of human factors involvement) are listed below:

1. Determine the goals of equipment design or the functions that equipment must perform.

2. Determine the problems involved in meeting these goals or the problems that prevent an immediate decision.

3. Determine the tradeoff factors to be considered and the relative weight that should be assigned to each.

4. Determine whether data are available to make a decision, as well as the adequacy and implications of these data.

5. Determine the alternative design solutions that permit achievement of system goals.

6. For each alternative, anticipate the functional performance consequences (in terms of equipment and human factors) that follow.

7. Determine the advantages/disadvantages of each design alternative relative to the weighted criteria.

8. Select the design alternative that most nearly achieves system goals.

Figure C-1 illustrates the tradeoff comparison process. This method lists the functional and technical design requirements and compares the alternative approaches with respect to the degree to which they satisfy individual requirements.

Information relating to design analysis communicated to DSARC should include the following:

1. Indicate the general types of human factors design analyses performed with regard to which subsystems and equipments.

2. Indicate all critical behavioral issues for which tradeoff studies were performed and decisions made.

3. Indicate any unresolved critical issues requiring further human factors design analysis.

4. Assess the state of the art in man-machine interface devices and techniques as these relate to the sytem being designed.

NOMENCLATURE Trade-off Release Energy to Unfasten Stage 1		Comparison Matrix of Design Approaches		Selection
	*Functional & Technical Design Requirements	1	2	
Operability		Solenoid Valve	Explosive Valve	
Insure response of actuator within 200 milliseconds of valve release.	Stored pneumatic energy shall be released by a valve to activate the Stage 1 unfastening mechanism. Valve flow parameters must insure positive and rapid energy release to the unfastening mechanism. All connectors shall comply with military standards for threads and fittings (Reference function---etc.) 28 VDC ± 1.5 will be required to actuate the release valve. (Reference function---etc.)	A solenoid valve operated by a 28 VDC signal from guidance and control will release energy to the actuating mechanism.	An explosive valve operated by a 28 VDC signal from guidance and control will release energy to the actuating mechanism.	Performance 2, 1 Reliability 2, 1
		*Discussion	*Discussion	
		Pro 1. Solenoid valve is an off the shelf component with standard line connectors. Con 1. Solenoids are heavy. 2. Voltage spikes would be caused in the DC line. 3. A heavy load would be imposed on diodes in the guidance and control autopilot programmer switch.	Pro 1. Explosive valve is is an off the shelf component with standard line connectors. 2. Low power is required to fire the explosive squibb. 3. No valve leakage prior to actuation. 4. Instant respose. 5. Light weight. 6. No voltage spikes would be caused in the DC line. Con 1. The valve cannot be functionally checked out, but has a good reliability record.	Procurability 2, 1 SELECTION Solution 2 NOTE: See section 4 of trade study report for reasons of selection.

*For example purposes only, representative requirments and partial discussion are listed.

Figure C-1. Comparison matrix of design approaches (taken from AFSCM 375-5).

APPENDIX D

TIMELINE ANALYSIS

TIMELINE ANALYSIS

Timelines are one of the basic analytic techniques used by human factors specialists to predict the incidence of time and errors. Timelines serve two purposes. First, they permit an appraisal of time-critical activities to verify that all necessary events can be performed. Second, they provide an integrated task-time chart to assess the occurrence of incompatible tasks and to serve as a baseline for workload evaluation.

The most common source of material for a timeline analysis is a detailed-level functional flow diagram, in sufficient detail that tasks are allocated to operators. Timelines are most effectively used during the concept formulation phase of system development, after DSARC I but before DSARC II. They require comparatively little time to develop, are of only moderate complexity, are equally useful for analysis of gross or detailed operator procedures, and can be used either for individual operator tasks or team tasks, as long as all the tasks are placed on a single timebase.

A typical timeline chart is shown in Figure D-1. Each timeline must be related to a higher level functional requirement. The functional flow title and number should be indicated on the timeline sheet for reference. Other information, such as location of the function and the type of function, are desirable. Each of the subfunctions or tasks are numbered and listed along the left side of the sheet. The time units of interest--hours, minutes, or seconds--are indicated, and, at the same time, a scale of suitable length is selected such that the total time period of interest fits onto the worksheet. Once the scale for a sheet is chosen, it should be adhered to for all portions of that timeline sheet. Timebases can be either clock- or scenario-referenced, but should not use units so small as to imply a degree of precision that does not exist in the data.

Since the timeline is used in conjunction with and to develop a workload evaluation, no report specific to this analysis need be made to DSARC. Items of information that it produces will be summarized with the workload report.

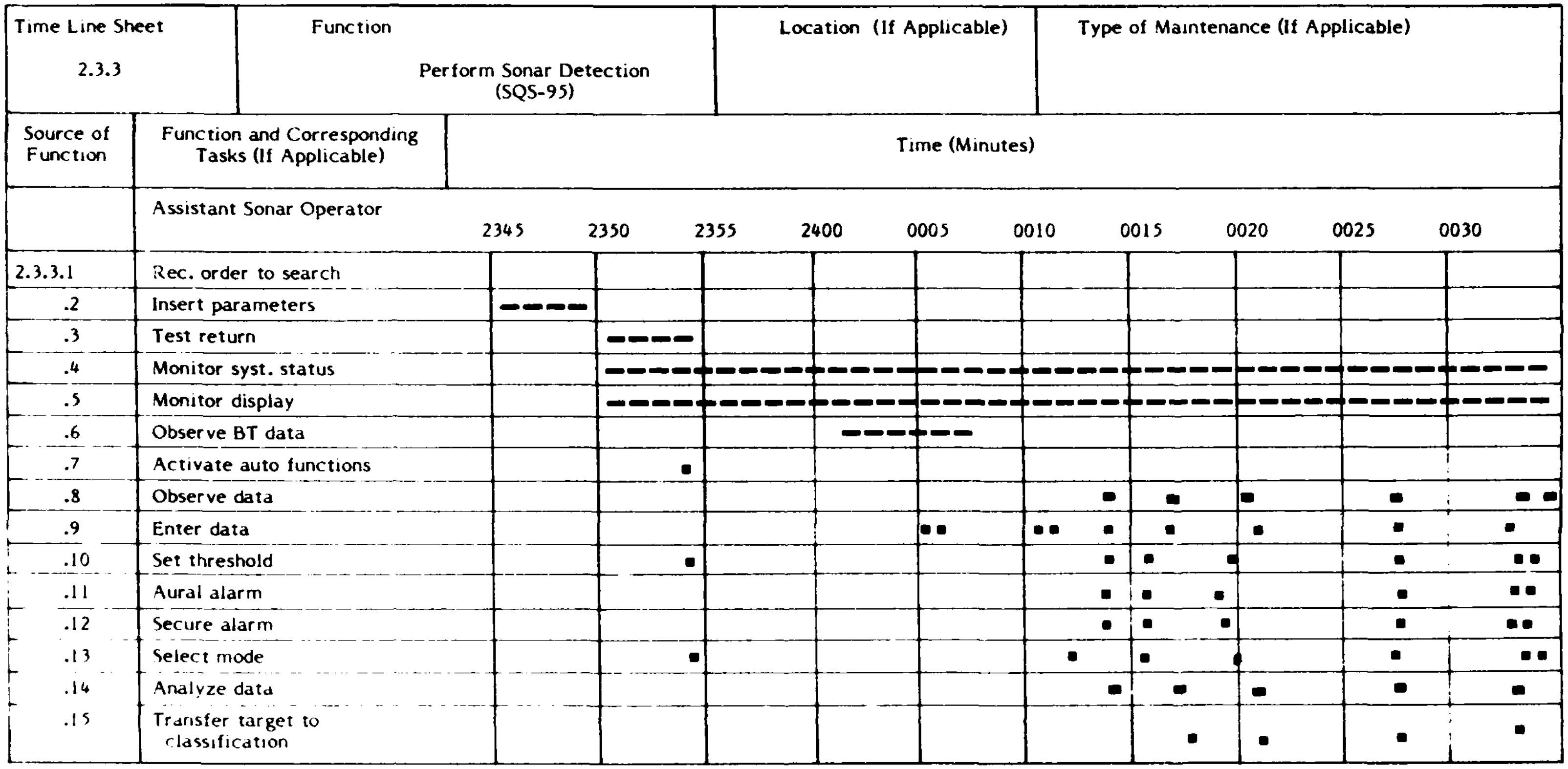


Figure D-1. Sample timeline sheet (Modified from Greer, 1976).

APPENDIX E

WORKLOAD ANALYSIS

WORKLOAD ANALYSIS

One of the critical questions to be answered during development is whether the crew of the new system will have sufficient time to perform its tasks adequately. More specifically, development management would like to ensure that personnel will not be overloaded in performing their work because an overload condition leads to less effective performance.

Workload analysis is a technique for answering the above question. It provides an appraisal of the extent of crew task loading, based on the sequential accumulation of task times, and permits the analyst to determine the crew's capability to perform all assigned tasks in the time allotted by mission constraints. Once this capability is confirmed, hardware design requirements can be more precisely designated. On the other hand, as limitations are discovered, alternate function allocations or crew task assignments must be considered and implemented. Workload appraisals are needed very early in development to assure that task loads are within the scope of the crew size and capability. Workload analysis verifies that no combination of tasks requires more task capacity, or time to perform, than is available. Although the technique has now also become part of sophisticated computer-aided techniques, such as the computer-aided function-allocation evaluation system (CAFES), this appendix describes a completely adequate manual method.

Workload analysis is a graphic presentation of an operator's workload constructed by plotting task involvement against a timebase (duration of operator activity). The analysis begins by dividing the operator's task into categories corresponding to his perceptual-motor channels. In the case of a team, the natural basis for categorization is the individual operator position (as in Figure E-1). Although workload analysis ordinarily describes individual task performance, its greatest effectiveness is realized when several crew positions are plotted together on the same graph. By doing so, any unbalanced workload distributions among personnel become readily apparent.

In some situations, operators can effectively perform more than one task at a time. However, an operator cannot perform two tasks simultaneously if both require the use of an overlapping perceptual-motor activity nearly 100 percent of the time. If the workload analysis chart is properly developed, it exposes such conditions. When such conditions are noticed, it is apparent that either a task must be given to another operator or the operator must be provided with some type of equipment assistance.

Task loading estimates may come from several sources. For example, the task may be the same as, or similar to, another task in another system that is in actual operation. Task time data from previous systems are generally the most reliable since, presumably, they have been verified in practice. When such information is not available, the next best data source is several operators who have performed similar tasks.

When experienced operators or other data sources are not available, the behavioral analyst, together with knowledgeable project designers, must make an "educated guess" about the task workload implications. The analyst will have to do what he does with all problems of this sort; he will have to break the task down into its simplest elements and extrapolate from what he knows about other subtask elements.

Workload analysis is most generally performed after DSARC I, when sufficient other work has been performed to develop the necessary input data. It may continue past DSARCs II and III. It may be used to perform a gross or top level (several minutes at a time) analysis of operator workload or a very detailed one (a few seconds at a time). If

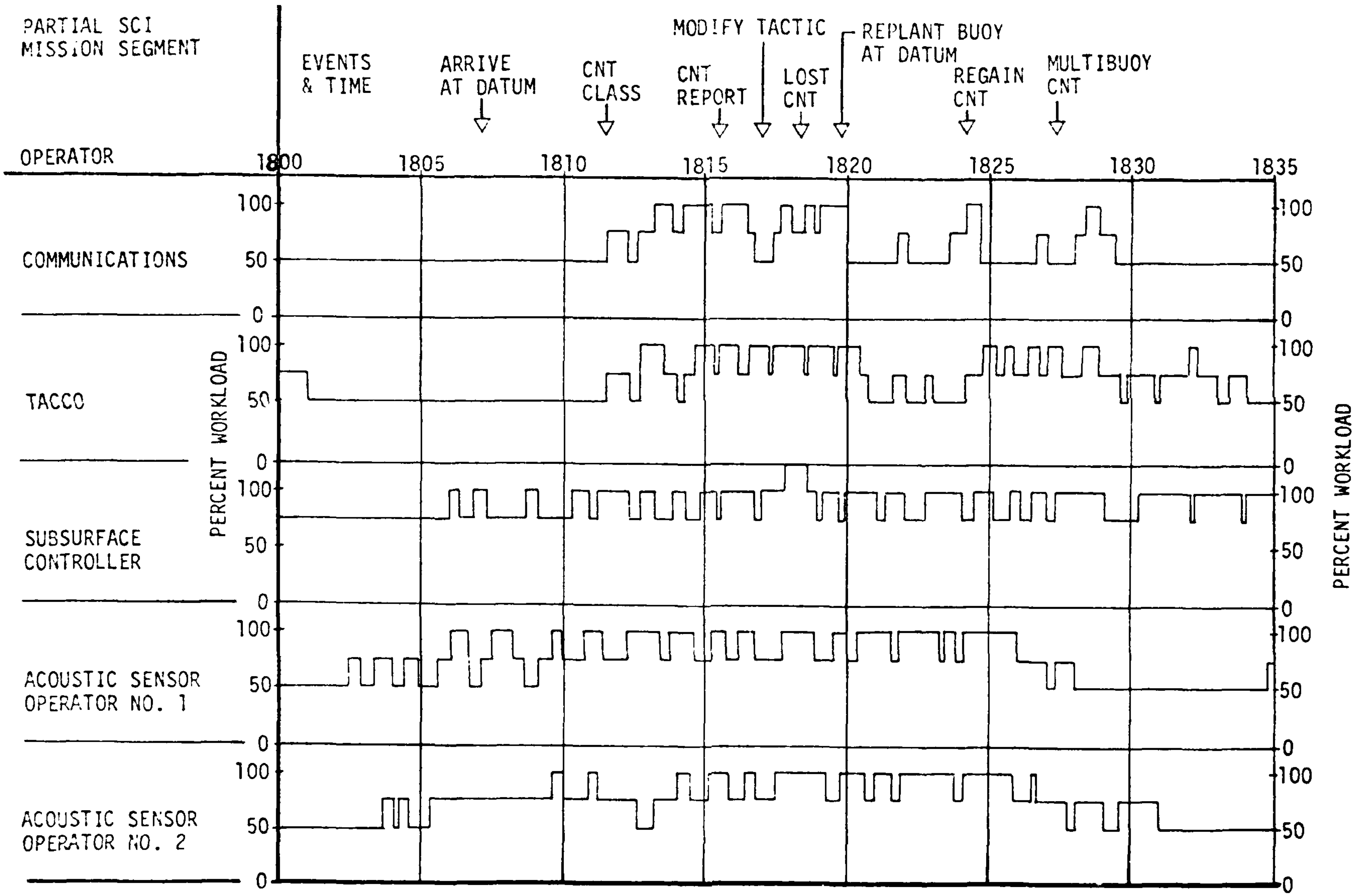


Figure E-1. Sample Workload Analysis for Sonobuoy Detection (Taken from Greer, 1976).

several workload profiles are combined in a single graph, it is possible to compare tasks being performed simultaneously. Because of the definition of work overload and the notion of the use of separate perceptual-motor channels, this technique is best used by analysts alone.

Since the process of estimating workload is based on the estimate of time required to do the task, it is only as accurate as that data. It is also limited by the knowledge of the time available to do the task and by unknown discrete channel summation effects. Depending on these variables alone, the accuracy of most workload assessments is probably in the ±20 percent range.

The workload analysis may be made up of a simple continuous chart from the beginning to end of a mission, or there may be several charts, each of which describes a particularly critical segment of the mission. The time scale in the analysis should be compatible with task complexity; for example, 15-minute intervals may be all that is necessary for simple workload analysis evaluations, and 5-second intervals, for more complex tasks. Whatever intervals are used should be common for the total group of tasks and operators when they interact.

There are also computer-controlled ways of performing a workload analysis. The most well known is the workload assessment model (WAM), which is one submodel of the more comprehensive CAFES system developed by NADC. WAM considers the human performance aspect of man-machine function allocation schemes on a time and cumulative task basis to determine whether man can perform all of the tasks derived from the allocated functions. It uses a timeline of mission tasks and determines those periods when the operator is overloaded in terms of time available versus time required to do all tasks. Workload can be analyzed for each operator in a crew to determine how changes in task allocations will alleviate overloading conditions.

Similar to the manual workload technique discussed previously, WAM is based on workload variations in each performance channel (e.g., eyes, hands, feet). It generates bar graph and histogram plots of workload data, which can be visually scanned to find heavy workload situations. WAM also provides an option for automatically shifting tasks to equalize workload.

Workload analysis results should be communicated to DSARC decision makers in relation to questions dealing with personnel limitations and critical issues to be addressed. Information should contain the following:

1. Determination that personnel are/are not overloaded and hence capable/not capable of performing their tasks.

2. Points in the scenario of system operations at which unacceptable operator workloads have been found and the reasons for such overloads.

3. Effects of the overload on system effectiveness (related to human performance reliability (HPR) prediction).

4. Recommendations for system redesign and/or further analysis/testing.

APPENDIX F

MANPOWER ANALYSIS

MANPOWER ANALYSIS

Manpower analysis is conducted to determine (1) the number of personnel (officers and enlisted men) needed to operate and maintain the system (e.g., a ship) and (2) the individual ratings and skill levels each of these personnel should have. This information is fundamental to just about every other behavioral analysis performed during system acquisition.

There are two ways in which the desired information can be gathered. First, since almost every system developed in the military has a predecessor to which it is intimately related, the manpower allocation of the predecessor can be used as the starting point for the new manpower analysis. Given that one knows (1) what the manpower allocation is for the predecessor, (2) the differences between the predecessor and the new system, and (3) the implications in terms of manpower of these differences, required manower can be derived by making manning changes appropriate to these differences. Determination of the differences between the systems and their behavioral implications is aided by the task analysis described previously. The differences between the two systems are expressed in terms of changes in hardware reflected in task changes that imply certain manpower needs. For example, if the new system is to be more highly automated than its predecessor, the automation changes will be reflected in a reduction of required manning. If equipment is added to the new system, then the tasks to be performed may demand additional personnel.

Second, the predecessor system can be ignored completely and a new manpower determination made based on a previously developed task analysis. However, this procedure ignores whatever information can be gained from the predecessor system.

The process of estimating the required manning should be accomplished within the following guidelines:

1. Manning must provide for the performance of all day-to-day activities required of the personnel in the system.

2. Manning must provide for performance of all emergency action that can feasibly be anticipated.

3. Manning must provide for scheduling of normal work periods and off-time periods, with sufficient personnel to keep the system in operation for long periods of time.

4. Manning must incorporate as few different jobs as possible.

5. Manning must require as small a number of personnel as possible.

6. Manning must require a minimum of training time to fulfill and maintain the fully manned strength of the unit.

The listing of operations and maintenance task areas and their associated equipment components should first be examined to establish the logical groupings of (1) related operational tasks to be accomplished on similar equipment at the same location and (2) maintenance tasks that involve similar equipments. Also, it is necessary to identify which of the operator groupings or positions at one location are similar to those at other locations as a first step in determining if the same set of personnel qualifications may be used to fill both positions. The following guidelines should be useful for grouping task areas into positions:

1. Activities assigned to a position should keep the man busy for a typical work period of unit operations.

2. Tasks should be grouped so they contain similar knowledges and skills, thus minimizing training requirements. Groupings of similar knowledges and skills might be based on: same equipment used, same general procedures followed, same type and level of knowledge or principles of operation required, same nomenclature and location of parts, and same man carrying through a related series of actions.

3. If otherwise appropriate, the tasks grouped in a position should resemble the groupings of comparable positions of an existing system.

4. Technical supervisory positions can be established by determining who (a) decides what men will do during a mission, (b) decides the distribution of material, equipment, etc., to maintenance personnel, (c) allocates maintenance activities in an emergency, and (d) monitors the quantity and quality of maintenance and operator work.

5. In assessing the relative supervisory level of a position, consideration should be given to the (a) relative amount of routine versus nonroutine activities, (b) relative range and complexity of situations requiring nonroutine activities, (c) numbers of factors (or sources of input information) entering into problems and decisions, (d) need for judgment and proper action based on incomplete data, and (e) amount of irreversible commitment depending on decisions.

The development of skill-level estimates for each position identified is based on the questions outlined in step 5e. above, and on the skill levels assigned to similar positions in previous systems. This latter information is further supplemented by relating task area requirements to the statements of knowledges and skills required by the various naval enlisted and officer specialities and levels described in the Manual of Navy Enlisted Qualifications (18068D) (BUPERS, 1981) and the Manual of Navy Officers Classifications (15839C) (BUPERS, 1975). An additional source, the Manual of Navy Enlisted Occupational Standards (18068D) (BUPERS, 1975) is used in the further process of determining whether (1) the new skills and knowledges required to operate, maintain, repair, and overhaul the system are within the scope of existing ratings or require revision of qualifications for existing ratings, and (2) they may be identified through the current Navy enlisted classification (NEC) and naval officer billet classification (NOBC) codes.

With the identification of the various operator and maintenance positions and skill-level requirements, the number of operators per position and the total number of each category of maintenance personnel must be determined. In connection with this activity, the analyst should obtain the system installation schedule identifying the type of Navy unit, the number of units affected, the number of installations per unit, the planned installation date, and the system being replaced, if any.

During the program initiation phase, the manpower requirement is determined by aggregation; that is, by adding the manpower for individual systems and modules together with the level of administration and support required. During later phases, equipment-related data (e.g., mean time between failures (MTBF) or mean time to repair (MTTR)) are used.

The information supplied in inputs to the various DSARC reviews include:

1. The number of personnel the new system requires.

2. Their job rating, pay grades, and corresponding skill levels.

3. The rational for deriving these manpower estimates.

4. Differences in manning between the new system and its predecessor and reasons for these differences.

5. Implications in terms of equipment, training, etc. of the manpower estimates.

APPENDIX G

TEST PLANNING

TEST PLANNING

Introduction

The two types of tests performed during system development can be categorized as developmental and operational. Developmental tests, which seek to answer engineering questions arising during design, are performed between DSARC milestones I and III. Operational tests, which seek to verify that the fully developed system satisfies system requirements, are performed between milestones II and III, although some extended operational tests may be performed ever milestone III. Both types of tests are ordinarily initiated and controlled by engineering personnel. Behavioral specialists may make some contributions to their planning but, except in the operational test, which is fairly systematic, their contributions are limited. Other tests, which one may call human factors tests, are initiated by the behavioral specialist and can be more systematically controlled by him. These tests make extensive use of full-scale wooden or plastic mockups.

Since DSARC decisions are made wherever possible on the basis of empirical data, considerable emphasis in DSARC checklist questions is placed on testing. The achievement of effective testing requires appropriate planning.

Developmental Tests

Types of Development Tests

Because developmental tests are quite limited in terms of the information they can provide the behavioral specialist, the amount of planning the human factors engineer can do for these tests is necessarily limited. The four major developmental tests or test situations that are of interest to the human factors engineer are described below:

1. Prototype tests. Tests of prototypes or breadboards are conducted in factory and special test facilities as well as in the laboratory. The human factors engineer may participate in these tests if one of the significant test parameters involves a measurement of human responses. For example, a test of a prototype vidual display serving as the basis of a new navigation system might well involve human factors evaluation because of the need to measure the visual response. However, many such tests (e.g., system computer processing speed) describe purely engineering parameters and will not require behavioral participation.

2. Design engineering inspection (DEI). A design engineering inspection is often held during the development of new systems. At a relatively early stage in system design, the customer will inspect the system's projected design configuration, as this has been extrapolated from initial functional analysis. In fact, there may be more than one such inspection, if a series of changes in system configuration (e.g., different models) is planned. Since, in most instances, hardware has not yet been built, the inspection is carried on with mockups. These mockups are demonstrated by company personnel and examined by customer representatives who typically note the revisions they would like to see incorporated in the design.

3. Engineering mockup inspection. The company may also conduct an engineering mockup inspection as a last check prior to placing design into production. The engineering mockup, which is built from production drawings, is, in essence, the first prototype of the production system. The engineering mockup contains considerable operational equipment, although the equipment at this time may not be operating. It is therefore much more

comprehensive than the usual human factors mockup would be but is used by designers in much the same way that a static human factors mockup would be employed by human factors engineers--to check the adequacy of design by visual inspection. Teams of designers review the mockup visually to verify that parts and components meet physical requirements and are in accord with engineering blueprints, that they are accessible for maintenance, etc. Human factors engineers may be invited along with other designers to inspect the mockup and make recommendations for modification.

4. Qualification tests. Qualification tests may employ the engineering mockup production prototype or first production article to qualify system design. After the first prototype or production hardware is fabricated and before it is either sent for further testing or sold, at least the major items of equipment (not piece parts or individual components) will be qualified, acceptance tested, or checked out in the factory; this is the established final stage in the production of the first article. Qualification or acceptance tests are largely functional checks; that is, tests to determine that the equipment will perform to physical criteria and without any effort to involve human factors. The tests consist of individually activating each equipment function and then comparing the recorded output with that required.

The methods involved in securing human factors data from development tests do not differ substantially from those used ordinarily to measure human performance. The tools employed include instrumentation, checklists, interviews, and recording of time and error data. The following information can be secured from qualification checkouts:

1. Performance times. These data may be used to help develop a distribution of performance times and thus to set performance standards for operational performance.

2. Errors. Number and type of errors committed by test personnel are of special importance in evaluating equipment operability and maintainability and in anticipating problems that may occur in field testing.

3. Malfunctions. Types of malfunctions encountered and resultant troubleshooting behavior may be observed.

4. Interviews. Interview data regarding reactions of personnel to equipment operation and maintenance features of the equipment are also of great importance in discovering operability and maintainability problems.

5. Technical data. The adequacy of technical data documents used by test personnel will be of interest.

Advantages and Disadvantages

The engineering developmental test has two major advantages for the human factors engineer that are, unfortunately, mitigated by corresponding disadvantages. The first advantage is that it costs nothing in development time and relatively little money to the human factors engineer; the test being performed is an integral part of planned system development. The second advantage is that the tests apply directly to system development. If problems arise in the course of their performance that the human factors engineer can help solve, he has made a direct contribution to the development of the system.

On the other hand, there are at least three disadvantages to the developmental test. First, it is not always possible for the human factors engineer to get permission to participate in or even to observe the developmental test. In some cases, even the number of observers is restricted to those with a primary engineering interest. Second, the human factors engineer has little or no opportunity to manipulate variables, since the developmental test is not under his control. Also, developmental tests performed in initial design usually possess comparatively few operational characteristics and make no effort to simulate the full range of operational conditions so important to the human factors engineer.

The skilled human factors specialist may make some plans to participate in developmental tests. However, since he has no control over the way in which these tests are conducted, it is difficult for him to do any more than rudimentary planning with regard to, say, the measures that he can take.

Planning is much more possible with those tests specifically initiated by the specialist to answer behavioral questions. These tests make use of static and functional mockups. The static mockup is a three-dimensional, full-scale model of an equipment or assembly that is to be designed or has been designed in prototype form. Because it is static, it cannot be programmed to demonstrate the functions of operating equipment nor can it be operated by personnel to perform operational routines except on a simulated ("walk through") basis. However, it can be used to (1) evaluate and decide among alternative equipment configuraitons, (2) determine workspace difficulties by simulating operating tasks, (3) discover accessibility problems by maintenance operations, and (4) plan the optimal location and routing of cabling, etc. Much more behavioral information can be gained from a functional mockup, which, in its most elaborate form, is much like a simulator. However, the behavioral specialist cannot count on the availability of such a mockup.

Operational Tests

The final stage in planning the human factors test program is the development and writing of the detailed test plan for the operational system test (OST). No systematic performance evaluation, either of hardware or human factors, can be performed without a written test plan that integrates both of these elements. Every large-scale developmental project, if conducted correctly, will include the development of an evaluation plan. This should be ready, in preliminary form at least, by the start of prototype fabrication and should be continuously updated after that.

The OST, which is performed after the initial operational system is produced (between DSARC II and before DSARC III), is a test that simulates, in as much detail as possible, the actual operation/maintenance of the system as it would occur in the operational environment. The purposes of this test (and its subvarieties, OT-I, OT-II, etc.) are to verify system adequacy for operational use and to discover those minimal modifications needed to bring the system to the point of satisfying mission requirements. The OST is usually performed by the Navy using Navy personnel but with the aid of contractor personnel. The major sections of the test plan written for the OST are listed below:

1. Purpose of test operations. This section describes both general and specific test objectives.

2. Administrative ground rules. This section describes the responsibilities of contractor customer groups and/or individuals in planning and conducting the tests, and

the way these groups and individuals interrelate. This information is required primarily for large-scale, field-evaluation tests, where masses of personnel and evaluators from different customer and contractor groups may be interacting.

3. The system being evaluated. This section includes a list of (a) the equipment subsystems to be tested, with a brief description of each, (b) equipment that will not be tested (if any), (c) special tools, test equipments, or checkout equipment required for system operation and maintenance, and (d) special test requirements for particular items of equipment. Also, it describes the criteria of successful system and equipment performance (i.e., test scenarios, criteria describing accomplishment of mission goal, and individual output readings, such as meter values that indicate that major equipment functions have been correctly performed). Finally, it describes the facilities where the test program will take place. Where test conditions impose potential hazards to personnel, safety requirements should be identified in terms of specific facility and equipment items (e.g., rescue air locks).

4. Subjects. The nature of the personnel performing in tests must be described; in particular, whether they are contractor and/or customer (military or civilian), and their relevant training and experience. If user personnel are to be used, this section should indicate at what stage in the test program they will be introduced and what test responsibilities they will have. Also the interaction between contractor and customer test personnel, any on-the-job training necessary for user personnel, and the way such training will be conducted should be described.

5. Data collection personnel and support. The number of observers and any special qualifications they must have, such as required previous training or experience, must be indicated. The activities to be performed by observers should be described in detail. Data recording and/or processing requirements for personnel and equipment should be included.

6. Program schedule and sequence. This section will list tests in which behavioral data are to be collected in relation to a calendar and milestone schedule.

7. Operational procedures. For each test operation, the procedures to be used in the test should be listed.

8. Evaluation design. In this section, the evaluation plan is described in detail with special emphasis on the following:

a. Variables. The independent variables to be tested (i.e., parameters to be experimentally manipulated) should be called out in terms of contrasts or comparisons (e.g., pressure suit vs. shirtsleeve condition). Also, the dependent variables and measures to be recorded should be listed.

b. Conditions. Conditions to be controlled and the manner in which the control is to be exercised (randomization, counterbalancing, etc.).

c. Data analysis. This will include a statement of the methodology to be used (e.g., t-tests, analysis of variance, etc.).

d. Instrumentation requirements. Lists all instrumentation, special test facilities (e.g., centrifuge), recording equipment (timing devices, magnetic tape recorders, etc.), any computers to be used and for what purpose (providing stimulus inputs, feedback, etc., or data analysis). For manual recording of data, indicate the forms to be used, who fills them out, how they are to be processed, etc.

e. Test configuration. Indicates under what mission-determined conditions (e.g., normal or emergency mode) the data will be gathered.

9. Test integration requirements. Describes whether a particular evaluation depends on the occurrence of any other test, how it interrelates, and if this will create any problems.

10. Criteria of evaluation completion. Indicates when the evaluation will be considered successfully completed (e.g., after a certain number of test replications, when certain objectives have been achieved).

DSARC Inputs

In response to DSARC questions involving test planning parameters, the following information should be supplied:

1. Indicate and describe what behavioral tests are to be or have been performed, what their specific purposes are or will be, and what parameters and variables are or have been tested, projected or actual data analysis procedures, etc.

2. Describe problems that have been identified as a result of the tests performed to date, problems that remain unresolved, and new issues that will be studied in planned testing.

3. Assess the adequacy of the behavioral test planning and actual tests conducted to date and the problems encountered.

APPENDIX H

CRITERIA DEVELOPMENT

CRITERIA DEVELOPMENT

Definition

Human performance criteria are requirements in performance terms that must be levied on system personnel if the system is to perform correctly. In other words, in system X, personnel must perform task K in 52 minutes; otherwise, the mission will fail.

Human performance criteria are derived from overall system objectives and mission requirements just as functions and tasks are derived (see MFT analysis) and in much the same way. Having determined that a function is required for system performance, one asks how that function is to be performed. If the function must be performed manually, the next question is whether there is any requirement on the personnel performing that function to do so in a specified time, a specified number of times, or to a certain degree of accuracy. Speed, frequency, and accuracy are the basic human performance criteria for most system tasks.

The reason for examining human performance criteria is to ascertain whether personnel will be able to perform the function in the specified time, number of times, or with the desired accuracy. The physical and mental limits of human capability are fairly well known. Thus, for example, if a manual response is required to a signal within 100 msec., this immediately suggests that the function has been improperly allocated, since about 200 msec. is the minimal response time allowable. If a human performance criterion is found to be impossible to achieve, the function being supported must be simplified, modified, or automated. If the criterion can be achieved manually but only with great difficulty, the design configuration requiring that criterion must be examined or the configuration must be changed somewhat to reduce the difficulties. Another reason for determining human performance criteria is that these criteria become measures of human performance when developing test plans involving measurement of system personnel.

DSARC Inputs

In reporting to the DSARC, the following information relative to criteria development should be provided:

1. List all critical human performance criteria found and indicate what design features make these critical.

2. Indicate what the effect of failing to accomplish these critical criteria will be and make recommendations for minimizing any negative effects.

3. Indicate what testing--if any--is required to verify that human performance is actually critical at certain stages of system functioning.

APPENDIX I

TESTING AND DATA ANALYSIS

TESTING AND DATA ANALYSIS

It has been pointed out that there are two types of tests that may provide behavioral performance data: (1) regularly scheduled engineering (system developmental) tests, and (2) those created specifically for human factors purposes.

System Developmental Tests

Types of Tests

The types of system developmental tests described in Appendix G can be categorized as exploratory, resolution, and verification tests. These tests are compared in Table I-1 and discussed below.

Table I-1

Major Differences Among Test Types

	Type of Tests		
Characteristic	Exploratory	Resolution	Verification
Typically performed in	Predesign and early development	Early development	Throughout development
Control of independent variables	High	Moderate	Low
Number of measures recorded	Few	Few to many	Many
Repeatability of test conditions	Any reasonable number (e.g., factorial design)	Few (gross configuration differences)	One (comparison with performance standard)
Control over test environment	High	Moderate	Low
Number of dependent variables	Few	Few to many	Many
Factors initiating test	Ambiguity of system inputs/outputs	Need for design decision	Need to verify system adequacy
Part/system testing	Part	Part to subsystem	Subsystem to system
Resemblance to operational conditions	Low	Moderate to high	High

Note. Taken from Meister, D. Human factors evaluation in system development. New York: Wiley, 1965.

1. Exploratory tests. Exploratory human factors testing seeks to determine basic operator performance requirements and capabilities, particularly as these will be used in

the system under development. It is not concerned with comparing configurations or system performance with a standard but, rather, with determining the range of operator, equipment, or system responses and how they occur in various operational situations. This type of testing is initiated by a lack of required knowledge about operator or equipment responses; for example, the need to determine to what degree of accuracy an operator will be able to track a new display. Although the answer to this type of question does not specifically evaluate a particular system design, exploratory testing is initiated by a definite problem arising during system development, the answers to which have not been provided by previous research. Exploratory testing most resembles traditional laboratory experimentation because the tester can control and manipulate his variables and assign subjects to test conditions.

2. Resolution testing. The purpose of resolution testing is to determine which one of a number of system task configurations will be the most adequate in satisfying performance requirements. Resolution testing, in contrast to exploratory testing, is tied specifically to the system under development. The resolution test situation is not set up to determine the relationship among variables but specifically to compare two or more configurations. Consequently, test control is exercised primarily to ensure that all configurations are treated identically and are compared on relevant variables. Resolution testing displays some characteristics of system testing, insofar as the configurations being compared may be those of an operational configuration but need not involve all system elements.

3. Verification testing. Verification testing determines whether or not a particular design configuration meets specified performance requirements. In contrast to resolution testing, verification testing takes the system configuration selected (prototype or production), exposes it to anticipated operational usage conditions, and determines whether it meets previously setup performance standards. Verification testing, if performed correctly, is true system testing, involving all system elements.

The term "developmental tests" is used primarily to refer to engineering equipment tests designed to study equipment design characteristics. Developmental tests include (1) bench-type tests of "breadboards" and prototypes conducted in laboratory settings (such as dynamics or wind tunnel tests), (2) inspections of the engineering mockup and/or first production article, and (3) factory qualification or acceptance tests of the first production items. Since developmental tests have a methodology and rationale particularly oriented to equipment engineering purposes, their objectives do not often deliberately involve human factors unless specifically required by the statement of work. Nevertheless, these equipment objectives may not prevent the gathering of "fall-out" data that are also relevant to human factors evaluation; in many cases, it is possble for human factors engineers to participate in or observe such tests profitably.

Evaluation Methods

The methods involved in securing human factors data from development tests which include instrumented measurements, checklists, interviews, and manual recording of time and error data, do not differ substantially from those used in other contexts. The following information can be secured from these methods:

1. Performance times. Because of the nonoperational character of these tests, these data are at best only suggestive, but they may be used to help develop a distribution of performance standards for field test operations and operational performance.

2. Errors. Number and type of errors committed by test personnel are of special importance in evaluating equipment operability and maintainability and anticipating

problems that may occur in field testing. Since personnel may not have had a great deal of experience in using equipment being qualified for the first time, however, it is likely that the number of errors made will be unrealistically high.

3. Malfunctions. Types of malfunctions encountered and resultant troubleshooting behavior may be observed.

4. Interviews. Interview data regarding reactions of personnel to equipment operation and maintenance features of the equipment are also of great importance in discovering operability and maintainability problems.

5. Technical data. The adequacy of technical data documents used by test personnel will be of interest.

Human Factors Tests

The second type of test available to the human engineer is the human factors test, which is developed specifically to answer human factors questions and is not ordinarily part of scheduled "engineering" tests (although there is no reason why human factors tests should not and every reason why they should be so scheduled). Because the human factors test is performed in response to design problems, its use is not restricted to a particular phase of development. However, if it is to be maximally effective, it must contribute to initial design.

Since the human factors test, which can be performed in mockups or simulators, is planned by and under the control at all times of the human factors engineer, he conducts it according to procedures that are largely derived from experimental methodology (i.e., methods involving the control and manipulation of variables).

The human factors test also has advantages and disadvantages. The opportunity to retain control over the test situation and thus to manipulate variables means that data obtained may be more operationally meaningful to the human factors engineer than those obtainable in developmental tests.

Categories of System Test Data

The data to be analyzed in any full-scale system evaluation are of five general types:

1. Mission accomplishment data. Data describing or related to the accomplishment of the overall mission (e.g., was the mission accomplished satisfactorily?).

2. Equipment operation data. Data describing equipment functioning (e.g., did the guidance subsystem acquire its target with minimum resolution error?).

3. Personnel performance data. Data describing the performance of system personnel (e.g., errors and response times).

4. Supporting system data. Data describing the adequacy of supporting system elements such as technical data, communications, and logistics (e.g., were all spares available when needed?).

5. System characteristics data. Data describing the characteristics of the system as a total entity (e.g., operability, reliability, and maintainability measures).

Each of the five types of data can be analyzed in terms of several measures.

Total Peformance and Diagnostic Measures

Some of the test performance and diagnostic measures (see Table I-2), particularly those dealing with the amount achieved, accomplished, or consumed, are measures that are characteristically used to evaluate total system performance. Others, such as those dealing with errors and various behavior categories, are used to describe the performance of individual system elements. The difference between the two kinds is the difference between measures of terminal system output as against intermediate criterion measures. The latter are mainly of diagnostic interest and are useful in explaining what inputs influenced the terminal outputs.

The particular measures selected will depend on the nature of the system, its mission functions, and the hypothesized importance of individual system elements and behaviors to overall system performance. For example, if a moving vehicle is the primary element of the system, fuel consumption measures will be highly appropriate; if the system involves tracking as a major parameter, time and percent on target are appropriate. In addition, measures highly correlated with terminal system outputs and describing the greatest amount of performance variance should be selected. The number of diagnostic measures selected depends on how detailed an analysis the human engineer wishes to conduct. The range of measures available as a function of various tasks is very broad. Although the data analysis should theoretically utilize all relevant measures, there is a practical limit, particularly in diagnostic measures, because of the analysis burden that they impose.

Analysis of System Test Data

Since the system is composed of interacting primary and supporting elements (equipment, personnel, procedures, technical data, logistics, and communications) that are organized at various levels of operation (mission segments and phases) and under different conditions, utilizing equipment grouped in units of increasing complexity (component, assembly, subsystem, system), and operated in terms of tasks of varying complexity (simple and complex) by various operators and crews, the data collected from the system evaluation must be analyzed at each of these interactive levels. Data from each level must interface properly to produce a meaningful system description.

In addition to providing data at various system levels, the system, considered as a superordinate entity, has attributes or characteristics, such as operability, maintainability, and reliability, which are distinct but not independent of each of the foregoing elements. These attributes must also be analyzed.

The only way the system can be fully described in such comprehensive terms is through a mathematical model. This does not mean that individual analyses of subsystem functions or conditions cannot or should not be performed through the use of the more common statistics of comparison and correlation, but only the system model is broad enough to encompass all these system elements in interaction. Whatever the means employed, the purpose of the analysis is to determine whether system performance requirements have been achieved (i.e., whether the mission has been accomplished successfully) and, even more importantly from the human factors engineer's standpoint, whether each of the system elements involving significant man-machine interaction has performed acceptably during that mission.

Table I-2

Classification of Generic Performance Measures

Time

1. Reaction time, i.e., time to:
   a. Perceive event.
   b. Initiate movement.
   c. Initiate correction.
   d. Initiate activity following completion of prior activity.
   e. Detect trend of multiple related events.
2. Time to complete an activity already in process, i.e., time to:
   a. Identify stimulus (discrimination time).
   b. Complete message, decision, control of adjustment.
   c. Reach criterion value.
3. Overall (duration) time:
   a. Time spent in activity.
   b. Percent time on target.
4. Time sharing among events.

Accuracy

1. Correctness of observation, i.e., accuracy in:
   a. Identifying stimuli internal to system.
   b. Identifying stimuli external to system.
   c. Estimating distance, direction, speed, time.
   d. Detection of stimulus change over time.
   e. Detection of trend based on multiple related events.
   f. Recognition: signal in noise.
   g. Recognition: out-of-tolerance condition.
2. Response-output correctness, i.e., accuracy in:
   a. Control positioning or tool usage.
   b. Reading displays.
   c. Symbol usage, desision-making and computing.
   d. Response selection among alternatives.
   e. Serial response.
   f. Tracking.
   g. Communicating.
3. Error characteristics:
   a. Amplitude measures.
   b. Frequency measures.
   c. Content analysis.
   d. Change over time.

Frequency of Occurrence

1. Number of responses per unit, activity, or interval:
   a. Control and manipulation responses.
   b. Communications.
   c. Personnel interactions.
   d. Diagnostic checks.
2. Number of performance consequences per activity, unit, or interval:
   a. Number of errors.
   b. Number of out-of-tolerance conditions.
3. Number of observing or data-gathering responses:
   a. Observations.
   b. Verbal or written reports.
   c. Requests for information.

Amount Achieved or Accomplished

1. Response magnitude or quantity achieved:
   a. Degree of success.
   b. Percentage of activities accomplished.
   c. Measures of achieved reliability (numerical reliability estimates).
   d. Measures of achieved maintainability.
   e. Equipment failure rate (mean time between failure).
   f. Cumulative response output.
   g. Proficiency test scores (written).
2. Magnitude achieved:
   a. Terminal or steady-state value (e.g., temperature high point).
   b. Changing value or rate (e.g., degrees change per hour).

Consumption or Quantity Used

1. Resources consumed per activity:
   a. Fuel/energy conservation.
   b. Units consumed in activity accomplishment.
2. Resources consumed by time:
   a. Rate of consumption.

Physiological and Behavioral State

1. Operator crew/condition:
   a. Physiological.
   b. Behavioral.

Behavioral Categorization by Observers

1. Judgment of performance:
   a. Rating of operator/crew/maintainer performance adequacy.
   b. Rating of task or mission segment performance adequacy.
   c. Estimation of amount (degree) of behavior displayed.
   d. Analysis of operator/crew behavior characteristics.
   e. Determination of behavior relevancy:
      (1) Omission or relevant behavior.
      (2) Occurrence of nonrelevant behavior.
   f. Casual description of out-of-tolerance condition.
2. Subjective reports:
   a. Interview content analysis.
   b. Self-report of experiences ("debriefing").
   c. Peer, self, or supervisor ratings.

Note. Taken from Smode, Gruber, & Ely, 1962 as presented in Meister, D. Human factors evaluation in system development. New York: Wiley, 1965.

To determine this, a quantitative minimum standard must have been established for each subsystem function and relationship in the model. This can be expressed in terms such as "minimum time to launch." The specification of such quantitative standards is a major function of the reason for the development and progressive testing of the function analysis and the mathematical model, first, to generate expected and required function values, and, second, to test these against empirical data.

The essence of the analysis should therefore be a comparison in terms of each system relationship between explicit criterion values (such as amount of allowable fuel consumption for a mission or minimum acceptable tracking error) and actual achieved values for these parameters. Where criterion values (e.g., required time minima for ground checkout) are not available, the analyst is often forced back to correlational analyses, particularly in relating intermediate variable to terminal outputs.

Table I-3 summarizes the major categories of analysis that can be performed using the measures described in Table I-2. The reader will note that the analyses are categorized in terms of the following primary system elements and characteristics: (1) mission, (2) equipment, (3) behavior elements, and (4) system characteristics. The remainder of this appendix discusses various kinds of evaluational analysis.

Evaluation in Terms of Mission Performance

The initial data analysis is performed in terms of the overall mission, because everything else in the system is subordinate to mission accomplishment. Mission analyses describe the (1) relationship between the number of mission attempts and the number of mission successes (i.e., achieved performance *reliability*), (2) amount of system resources expended against the maximum permitted, (3) terminal output accuracy, and (4) overall mission duration and reaction time. Obviously, in such a global analysis the impact of the individual system elements, among them human factors, is largely obscured. Nevertheless, because the ultimate criterion of system performance is whether or not the mission is accomplished, the initial analysis of human factors data must be placed within this framework. To be comprehensible, personnel errors and difficulties must be analyzed within the mission context.

Since the mission is influenced by all system elements, its success or failure cannot immediately be interpreted in terms of the effectiveness of personnel behavior. The mission may succeed even when personnel make errors, or it may fail even when they do not. The fact that there is no direct relationship between mission accomplishment or nonaccomplishment and various aspects of personnel behavior involved in implementing the mission forces the human factors engineer to examine personnel behavior in relation to every mission.

Mission success or failure reflects on the significance of operator/crew performance during the mission. Significance here refers to the impact of that performance (especially time and errors) on mission success or failure. For example, errors that contribute directly to mission failure assume a different meaning than do errors correlated with mission success. Errors therefore must be conceptualized in pluralistic and contingent terms.

There is, of course, a direct correlation between ultimate mission success and the success or failure of each mission segment, phase, and task. If each is actually necessary, the mission as a whole cannot be successful if any of its constituent functions or tasks fails completely. Since the result of individual task performance may modify or be modified by following tasks and functions, the degree of direct task relationship to system

Table I-3

Generic Categories of System Test Data Analysis

Analyses of Mission

1. Frequency and percentage of mission accomplishment in terms of:
   a. Overall mission goals.
   b. Comparison of segments and phases.
   c. Goal-relevant criteria (e.g., miss distance).
   d. Time to accomplish the mission.
   e. Expenditure of system resources (fuel, etc.).
   f. Reaction time (where applicable).

2. Measures of system reliability in terms of ratio of mission successes to failures:
   a. Achieved reliability measurement.
   b. Comparison of achieved with predicted or required reliability.
   c. Percent system availability.

3. Analysis of frequency, type, and severity of discrepancies occurring during the mission:
   a. Equipment malfunctions.
   b. Personnel error and/or difficulties.
   c. Time discrepancies.

Analyses by Equipment

1. Determination of equipment subsystem reliability:
   a. Mean time between failure for major equipment components.
   b. Comparison of minimum acceptable reliability for each component with its achieved reliability.

2. Determination of the nature, frequency, and impact of equipment failures.

Analyses by Behavioral Elements

1. Frequency and percentage of tasks accomplished by personnel:
   a. Percentage of time operator/crew track target correctly.
   b. Percentage of time operator/crew detects signals, identifies stimulus correctly.
   c. Ratio of tasks accomplished correctly to tasks attempted (achieved reliability).
   d. Probability of task accomplishment (human reliability index):
      (1) In terms of time to accomplish task.
      (2) In terms of time to react to initiating stimulus.

2. Frequency, percentage, magnitude, and classification of human errors analyzed:
   a. By task and mission phase.
   b. By operator position.
   c. By impact of error on mission or task.
   d. In emergency conditions.
   e. As human-initiated failures.
   f. By probability or error occurrence.
   g. By equipment operated.

3. Comparison of mission phases in terms of:
   a. Frequency and type of human error.
   b. Error effect.

4. Comparison of tasks:
   a. Required vs. actual duration.
   b. Required vs. actual reaction time.
   c. Other criteria requirements vs. actual accomplishments.

Analysis by System Characteristics

1. Frequency and classification of system discrepancies:
   a. Communications errors.
   b. Nonavailability of required system elements.
      (1) Prime equipment, test equipment, tools, spares.
      (2) Personnel.
      (3) Technical data.
   c. Logistics inadequacies.
   d. Technical data inadequacies.

2. Measures of operability (see behavioral analyses).

3. Measures of system maintainability (as a whole and by mission segments) in terms of:
   a. Equipment downtime.
   b. Number and duration of holds and delays in mission performance for other than maintenance reasons.
   c. Amount of preventative maintenance.

Note. Taken from Smode, Gruber, & Ely, 1962 as presented in Meister, D. *Human factors evaluation in system development*. New York: Wiley, 1965.

output may be slight. Tasks performed closer in time to the terminal output may be more closely related to that output. The same is true of higher level (e.g., subsystem) mission phases and functions. Therefore, the effectiveness of individual tasks cannot be evaluated solely on the basis of terminal mission success or failure. It is essential, however, to determine the relationship between individual tasks and system/subsystem outputs, for, unless this relationship is known, task performance cannot be interpreted. Where it is necessary to know whether particular task errors significantly influence mission success, this may be determined by correlational analysis and/or by tracing the effect of the error through successive tasks and functions. Statistical analysis of the relationship may suggest that the relationship is very strong, but only empirical examination of error effects can provide certainty.

Comparison of Mission Segments and Phases

Detailed mission analyses may involve comparisons among mission segments and phases in terms of relevant human performance criteria, for example, a comparison of manual tracking error among various phases. The human factors engineer might seek to determine whether one or more of these phases manifested a significantly greater tracking error. To compare performance among segments and phases, however, it is necessary to select criteria that take into account the effects of differences in number and compositon of the various segment and phase tasks.

Decision-making Analysis

Another form of mission analysis is in terms of critical decisions that must be made by the system. If the system is required to choose among alternative methods of operation, it is possible to determine how well the system has performed by analyzing whether, in each case, the most effective alternative was chosen. The relationship of any incorrect decision to system performance must also be determined.

A special form of critical decision is the emergency situation, such as malfunctions of a critical life support system. Decisions in response to emergencies represent the system's ability to cope with "high stress" conditions. If system deficiencies exist, they are most likely to be revealed under emergency conditions. Analysis of crew responses to emergencies must also consider the character of the mission phase in which the emergencies occurred.

Mission Reliability

Where a mission or mission segment is performed repetitively, the human factors engineer will be interested in the consistency of personnel performance over the successive trials.

In summing up, it is apparent that, although the human factors engineer may not have primary responsibility for mission analysis, he must examine his data in terms of their interrelationships with mission performance and seek to determine the impact of behavioral responses on the degree to which the mission has succeeded. Because of the lack of clearcut comparison conditions in many system tests, the analysis of relationships among intermediate and terminal inputs and outputs may have to be performed by multivariate correlation techniques and what is essentially logical analysis of the data.

Evaluation in Terms of Equipment

The adequacy of equipment functioning is obviously a major consideration to the system evaluator, but is properly the concern of the test and reliability engineer rather than of the human factors engineer. It is possible, however, for system personnel to experience special difficulties with particular equipment and for these difficulties to reflect on overall system performance. The human factors engineer will, therefore, be concerned whether more errors (or more or less of any other behavioral response) are made in relation to one particular class of equipment or individual equipment than another.

The simple determination of amount of error (or any other criterion measure) is not, however, sufficient to draw any significant conclusions concerning equipment design or operation. Determining the significance of the error for equipment performance and analyzing error relationships with other system elements and functions are critical requirements.

If the human factors engineer finds, for example, that an excessive number of errors were made in operating a particular equipment, he will have to determine whether the operation of that equipment significantly influenced mission accomplishment. This will involve determining the relationship (if any) between the amount of personnel error on the particular equipment and some system output measure related to the operation of that equipment.

Evaluation in Terms of Behavioral Data

A general framework for system evaluation in terms of behavioral responses is presented in Table I-4. The basic unit of behavioral data is the task. Behavioral data are analyzed in terms of task completion, reaction time, and duration, with errors playing a secondary role. Errors are unimportant except in terms of their effect on task performance and in relation to other system and mission elements. If behavioral variables play a secondary, dependent role, it is because system performance is organized on the basis of mission parameters, rather than on the basis of the parameters of any individual system element.

The error responses described in Table I-4 can also be categorized in terms of the crew member or the crew that made these errors. To the extent that different crews are used in performing actual or simulated missions, it is possible to determine whether there are any statistically significant differences among them in terms of the behavioral variable (e.g., errors) being measured. Such a test can easily be performed using analysis of variance, t-tests, or nonparametric tests of the significance of differences between means. A significant variation among crews might suggest that at least some of the casual factors for errors in system performance were idiosyncratic to particular personnel and not attributable to system variables. The reverse might suggest that some system characteristic was responsible. The interaction of crew variables with other system variables should be determined also. However, the meaning to the system of any statistical findings must be determined separately from the statistical test.

It is much easier, of course, to determine the frequency and type of errors made by personnel than to determine whether these errors indicate personnel incapacity. The impact of an error on system performance is much more important than its frequency of occurrence, and it is precisely in the determination of casual significance that the human factors engineer runs into difficulties. Moreover, error "standards" (in terms of the number of errors "allowable" to an individual operator or crew) are not readily available,

Table I-4

Steps in the Evaluation of Behavioral Data

Analysis	Criteria Methods
1. a. Determine frequency and percentage of tasks successfully completed. b. Establish probability of successful completion in future.	1. a. Examine tasks in terms of their terminal outputs. b. Establish failure causes; eliminate tasks failing for equipment reasons.
2. Determine effects of task noncompletion on performance of: a. Subsequent tasks. b. Other system elements. c. Overall mission, segments, and phases.	2. a. Examine related task pairs; establish relevant dependent relationships. b. Determine which tasks present serious problems.
3. a. Determine task duration. b. Determine which tasks were significantly delayed.	3. a. Determine minimum performance time. b. Compare with actual performance time.
4. Determine whether task reaction time requirements were met.	4. a. Determine task reaction time requirements. b. Compare with actual reaction times. c. Assess human factors causes of reaction time failures.
5. Determine effects of task duration and reaction time delay on: a. Subsequent tasks. b. Other system elements. c. Overall mission, segments, and phases.	5. Identify tasks causing major portion of system delays.
6. Determine frequency and types of errors: a. In types of tasks and functions. b. On successive mission trials.	6. a. Establish human factors causes of errors. b. Estimate probability of error recurrence.
7. Determine impact of errors on: a. Task in which error occurred. b. On subsequent tasks. c. On other system elements. d. On overall mission, segments, and phases.	7. Establish significance of specific errors to task completion.

especially for newly developed systems, and may only be valid in a specific mission/scenario context. To compare the performance of the different members of the same crew may not be very meaningful, when the tasks they perform differ in terms of the number of steps and their difficulty. The latter in particular poses problems because, although procedural steps can be counted and compared, the differential difficulty level of each step is sometimes difficult to assess. The situation is simpler for determining the significance of performance time deviations. The minimum required and maximum permitted times for tasks and mission segments are often available because many missions are time-dependent (e.g., "windows" in launching space vehicles). Use of timeline analysis methods permits specification in detail of time requirements. Deviations from such time standards can then be easily determined, but the casual significance is more difficult to ascertain.

DSARC Inputs

The information supplied to DSARC reviewers should include:

1. Listing and brief description of developmental/operational tests performed; representativeness of test situation, including subjects; and questions for which tests were supposed to provide information.

2. Summary description of information gained from tests, with particular attention to personnel performance determined from the tests; for example, whether personnel can or cannot perform required which test tasks adequately and, if they cannot perform, reasons for this failure.

3. Major recommended changes to hardware/software, design, procedures, task allocations, etc.; recommendations, accepted, rejected, implemented, in progress.

4. Unanswered questions requiring further testing; implications if no further testing.

APPENDIX J

HUMAN PERFORMANCE RELIABILITY PREDICTION

HUMAN PERFORMANCE RELIABILITY PREDICTION

At both DSARC I and II, it would be highly desirable to be able to predict the adequacy of system personnel performance. This would permit system development managers to (1) determine that system personnel will (or will not) be able to perform their jobs adequately, (2) compare two or more design configurations in terms of which permits more effective crew performance, and (3) indicate where design changes are desirable to reduce error potential.

Human performance reliability (HPR), used as a number, is the probability that an individual operator or the crew as a whole will perform his/its tasks correctly. There are a number of HPR predictive techniques (Meister, 1971). However, the one most in use is called technique for human error rate prediction (THERP) (Swain, 1963; Swain & Guttman, 1980), which is used for estimating human error rates and predicting the man-machine system decrement that will result from human errors. The technique employs an iterative process composed of five steps:

1. Defining the operation to be evaluated.
2. Listing all operator tasks.
3. Estimating error rates for those tasks.
4. Predicting the effect of the errors on the system operation.
5. Recommending subsequent changes to reduce the system failure rate.

The THERP model uses two measures: P_i, the probability that an activity will produce an error of a class (e.g., Class i) and F_i, the probability that an error will lead to partial or total system failure (depending upon which is being considered). The P_i statistic is derived from error rate data over a unit of time. The determination of F_i is based on analyst judgment, the estimation being made with regard to the unique characteristics of the particular system being evaluated. In applying the model to a system, an initial task analysis of actions, having some associated error potential, is begun. All operations to be performed by the human operator are enumerated, along with contingency modes. As altenative actions are described, a probability branching tree is developed, describing the relations between the various contingency events. Once this tree has been generated, application of the model follows conventional reliability prediction techniques. A computer program can be written to assist the behavioral analyst in making tradeoffs to find the optimum balance between predicted system reliability and various cost factors.

Probability data are combined by a multiplicative method when tasks are assumed to be independent, or by the solution of functional equations (e.g., Task C is a function of the combined errors of Tasks A and B) when the operations are considered to be interdependent.

As a result of using THERP, the analyst should expect to obtain failure rates associated with the particular aspect of the system under consideration. Stated more precisely, the output represents the joint probabilities of P_i and F_i, or the error rate probability estimate is derived from the proper combination of individual F_iP_i products over all individual task behaviors. The model may be used in determining if a system will perform within designated reliability limits or in estimating the absolute level of possible operator efficiency. As a design analysis tool, alternative configurations may be compared on the basis of operator error probabilities. System redesign may be justified in some cases by presenting system failure probabilities as evidence of potential design deficiencies.

The technique has been applied to systems development problems with some measure of success. The model is comprehensive in scope and can be applied to all forms of equipment, tasks, and operator behaviors. The output offers a prediction of system effectiveness. As is the case with other HPR techniques, subjective judgments are required, based on such factors as analyst experience and familiarity with similar systems. The major limitation to this technique (and all human performance reliability techniques) is in the source of accurate reliability data. Further limitations to the general use of the model stem from the difficulty in collecting experimental data during system development to validate model predictions. To date, only a few validation studies of THERP have been performed, with, however, reasonably good agreement between the model and data. Since it requires a fairly detailed task analysis, HPR prediction becomes more effective as development time proceeds. However, crude predictions can be made by DSARC I and certainly more sophisticated ones by DSARC II.

THERP is a complex technique to use, in part because it involves the manipulation of several concurrent tasks, rather than a single task. It requires more time to perform than do other predictive techniques but has the great advantage of providing quantitative predictions. Because of its complexity, the techniques should be used only by qualified personnel.

DSARC Inputs

In presenting the results of the HPR predictions to DSARC I and II, the following should be included:

1. An indication of the probability that system personnel using the new system will (or will not) be able to perform their tasks to quantitative system requirements. (Use of the technique for this purpose presupposes that quantitative criteria for personnel performance exist.)

2. Description of the effect of the HPR probability on overall system reliability; that is, the effect of crew reliability on overall system reliability.

3. Comparison of alternative proposed design configurations in terms of their individual HPR scores.

4. Specification of critical design features that require redesign because they are associated with unacceptably low HPR scores.

5. Description of data constraints that qualify the conclusions drawn.

APPENDIX K

TRAINING ANALYSIS

TRAINING ANALYSIS

The training analysis that eventually winds up as the Navy training plan (NTP) is an exhaustive, extremely detailed, and comprehensive analysis that, in the case of a major weapon system, may require as much as 6 years to complete. The NTP, in addition to establishing a training program for new acquisitions, identifies manpower needs and defines the resources necessary to satisfy training requirements. Like the other analyses described, the development of the NTP is an iterative process that becomes progressively more detailed. The following description cannot deal with the methodology involved in performing this analysis because that would require several books. The interested reader can refer to the Training Requirements Handbook, Vols. I-IV (HARDMAN, 1980), from which much of the following material has been taken.

The training analysis has two subsections. The first is the development of the training concept, which has no specified starting point in the weapon system acquisition process (WSAP) but must be completed not later than DSARC milestone I and then updated throughout system development. The second is the training resource requirements necessary to support the training concept.

Training Concept Components

A training concept is the manner in which required training is to be accomplished in terms of the following components:

1. Type of training. States requirements for operator, various levels of maintenance, team, and/or proficiency training for all categories of personnel. Categories include military (officer and enlisted), civil service, and contractor personnel.

2. Presentation environment. Establishes the environment in which each type of training will be presented. Environments include formal school training, other formal training (i.e., structured on-board training), contractor-provided training, and informal training.

3. Presentation technique. Defines the specific technique to be used in presenting material for each type of training. Techniques include group instruction and various types of individualized instruction.

4. Presentation media. Describes the media or means to convey or communicate information required for each type of training. Examples of media include printed material, training equipment, training devices, training aids, and audio/visual aids.

5. Pipeline. Establishes the sequency of courses required for initial skill and skill progression training for each type of training. Courses include factory, prerequisite, replacement, conversion, and/or combinations thereof.

6. Location. Establishes the number of and eventually the specific locations of the training facilities for each type of training. Number of locations are expressed in terms of their actual number or minimally as single-, dual-, or multisited. Specific locations are expressed by the physical location (e.g., Norfolk, Orlando, or ASW School, San Diego).

Training Resource Requirement (TRR) Components

Training resource requirements (TRRs) are defined as the materials and personnel and cost thereof with which the training identified in the training concept will be accomplished. They are expressed in terms of the following components:

1. Training device. Consists of hardware and software (including simulators) that have been designed (or modified) for training purposes, involving, to some degree, stimulation of some type in its construction or operation, and having the methodological and evaluation techniques to train (refresh or expose) personnel to some level of performance proficiency.

2. Training equipment. Consists of equipment that is designed for use in the subject of instruction used by the instructor or student in teaching.

3. Other training material/other instructional material. All items of material prepared, procured, and used in a course or program as part of the teaching or learning process. This includes the general categories of training aids (instructional aids), training aid equipment (instructional aid equipment), and instructional literature.

   a. Training aids/instructional aids. Examples of various training aids are: Audio cassette, audiovisual aid, demonstration aid, graphic aid, mock-up operable transparency, and prefaulted modules.

   b. Training aid equipment. Equipment used to display training aids but that is not itself the subject of instruction (e.g., motion picture projectors, slide projectors, opaque projectors, etc.).

   c. Instructional literature. Printed matter used in the learning process, including that developed for a specific purpose and other printed matter procured (e.g., texts, manuals, etc.).

4. Billets. A billet is the basic personnel unit of a naval organization. Training billets include instructors, training support, and students (chargeable).

5. Military construction (MILCON)/site preparation. MILCON encompasses new, expanded, or extensive modification of training facilities (buildings/structures). Site preparation is loosely defined as minor modification of existing training facilities.

Algorithms

The methodology for deriving a training concept and the training resource requirements involves a series of questions that have been combined into logic sequence (e.g., Is individual operation required? Will military personnel perform the operations?). A logic sequence (algorithm) is illustrated in Figure K-1.

The training concept algorithms are described in Table K-1; and the TRR algorithms, in Table K-2. Both types of algorithms are generic and designed to be used at any point in the WSAP. However, TRR algorithms differ significantly from training concept algorithms because they depend on input data developed in the training concept. Also,

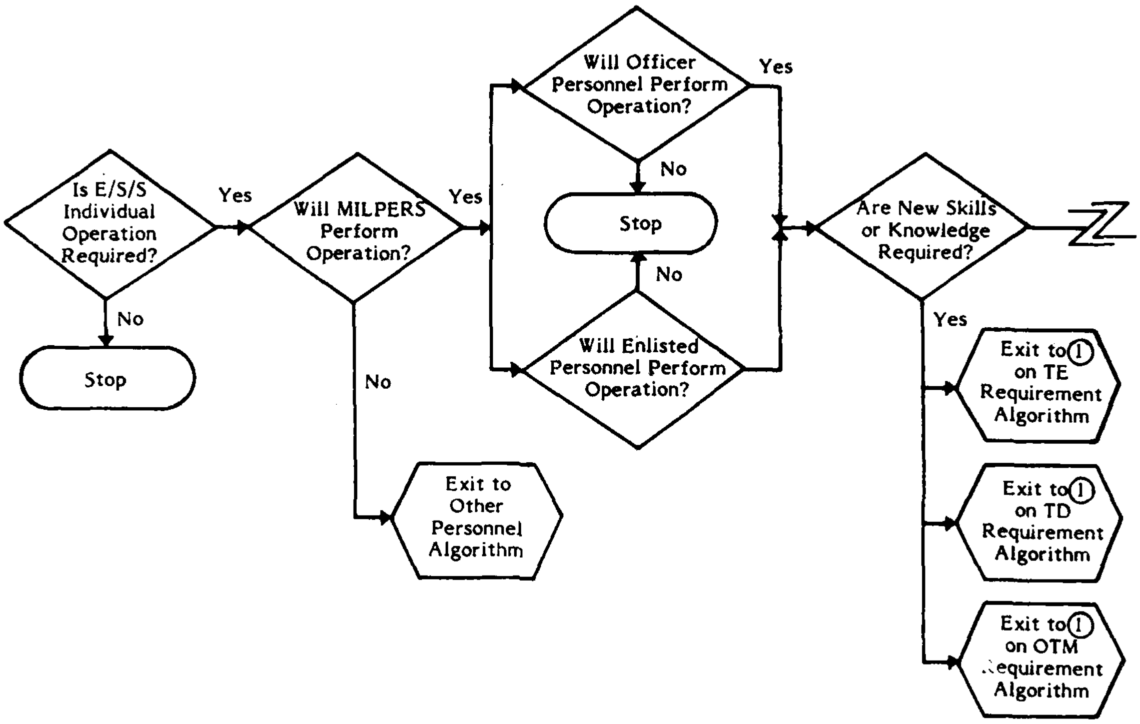


Figure K-1. Operator training requirement algorithm (partial).

they differ in structure since a combination of logic and computational sequences are used. Thus, the training concept algorithms are used to develop a narrative of the training concept; and the TRR algorithms, to derive specific quantitative values in support of the training concept.

DSARC Inputs

The following training information should be supplied to the DSARC:

1. Specification that a training analysis effort is underway, what it consists of, and the stage which it has reached.

2. Any training problems or issues that have been unearthed as a result of the training analysis, further steps being taken to solve these problems, and the milestone schedule for these steps.

Table K-1

Training Concept Algorithms

Title	Purpose
Operator Training Requirement	To develop the operator training requirement by determining the: 1. Need for individual operation of the equipment/subsystem/system (E/S/S). 2. Type of personnel required to perform the individual operation. 3. Requirement for new skills or knowledge necessary to perfom the individual operation. 4. Requirement for technical formal school training to impart new skills or knowledge. 5. Requirement for informal or other formal training to supplement formal school training.
Organizational Maintenance Training Equipment	To develop the organizational maintenance training requirement by determining the: 1. Need for corrective maintenance (CM) or preventive maintenance (PM) at the organizational level. 2. Type of personnel required to perform organizational CM and/or PM. 3. Requirement for new skills or knowledge necessary to perform organizational CM and/or PM. 4. Necessity of technical formal school training to impart new skills or knowledge for organizational CM and/or PM. 5. Requirement for informal or other formal training to supplement formal school training for organizational CM and/or PM. 6. Feasibility of incorporating organizational PM training for <u>new</u> skills or knowledge within planned/existing individual operator training. 7. Feasibility of incorporating organizational PM training for <u>old</u> skills or knowledge within planned/existing individual operator training.
Intermediate Maintenance Activity Training Requirement	The logic steps for determining the IMA training requirement are identical to those used for determining organizational maintenance.
Depot Maintenance Training Requirement	The logic steps for determining the depot-level training requirement are similar to both those of the organizational and IMA levels. However, there is no PM requirement at the depot level.
Team Training Requirements	To develop the team training requirement by determining the: 1. Need for coordinated group operation of the E/S/S. 2. Type of personnel required to perform the group operation. 3. Requirement for intraoperation between (a) equipment within subsystem/system and (b) E/S/S and other E/S/S. 4. Requirement for new skills and knowledge necessary to perform the group operation. 5. Necessity of technical formal school training to impart new skills or knowledge for group operation.

Table K-1 (Continued)

Title	Purpose
Training Device (TD) Requirement	To develop the operator, maintenance, and team training device requirement by determining the: 1. Need for formal school training. 2. Need for TD to conduct technical formal school training. 3. Existence of a similar type E/S/S. 4. Existence of technical formal school training on the similar type E/S/S. 5. Use of TD in formal school training. 6. Capability of existing TD to satisfy training. 7. Need requirement for Class "A" school or other prerequisite training.
Training Equipment (TE) Requirement	The logic steps for determining the TE requirement are identical to those for TD.
Other Training Material (OTM) Requirement	The logic steps for determining the OTM are similar to those for TD. OTM involves determining, for formal and other formal/informal training, the requirement for (1) training (instructional) aids (e.g., audio cassette, mock-up, prefaulted module), (2) training aid equipment (e.g., motion picture, overhead, and sound/slide projectors), and (3) instructional literature (e.g., texts, curricula, training manual, and lesson outlines).
Pipeline Requirement	This algorithm is entered independently after determining the requirement for operator, maintenance, and team training. It is used to determine: 1. Operator and maintenance rating. 2. Operator/maintenance course integration requirements.
Location Requirement	To develop the requirement for the following siting options for operator, maintainer, operator/maintainer, and team training: single, dual, multiple, and actual.
Prerequisite Training Requirement (TD)	To determine the training device requirements for prerequisite training by ascertaining the: 1. Need for Class "A" or other prerequisite training. 2. Need for TD on E/S/S to conduct class "A" or other prerequisite training. 3. Existence of Class "A" or other prerequisite training. 4. Use of TD in existing Class "A" or other prerequisite training. 5. Capability of existing TD to satisfy training need.
Prerequisite Training Requirement (TE)	The logic for identifying prerequisite training TE is identical to that for identifying prerequisite TD.
Prerequisite Training Requirement (OTM)	The logic for identifying prerequisite training (OTM) is similar to that for identifying prerequisite TD and TE.

Table K-2

Training Resource Requirement (TRR) Algorithms

Title	Purpose
Training Device (TD) TRR	To translate the requirement for a training device, as established in the training concept, into the specific number of training devices necessary to support the steady-state training requirements of the E/S/S or total ship system and the estimated cost (i.e., funding profile) necessary to properly phase in the training device(s).
Training Equipment (TE) TRR	Same as TD TRR. Contains additional logic step to determine if TE installed as part of the TDs can be used in a standalone mode without impacting on the programmed use of the TD.
Other Training Material (OTM) TRR	To determine type of instructional literature, training aids, and training aid equipment necessary to support the E/S/S or total ship system and the estimated cost of each.
MILCON TRR	To determine the extent of new, expanded, or modified training facilities required to support identified training for the new E/S/S and the related cost.
Instructor TRR	To determine the number of instructors required to support the annual training load imposed by the new E/S/S or total ship system.
Training Support TRR	To calculate supplemental billets required for training support (e.g., training administration, course management, etc.).
Military Billet and Civilian Personnel Requirement	To determine the annual aggregate number of military officer and enlisted billet and civilian personnel requirements needed to support fleet and shore activities. This data serves as a basis for the annual training input requirement (ATIR) algorithm.
E/S/S Existing Course	To determine the additional TRR, if any, needed by existing courses of instruction having TD, TE, or OTM evaluated as adequate to impart skill or knowledge of the new E/S/S.
Annual Training Input Requirement (ATIR)	An independently entered algorithm to calculate for each course at each location the annual training input requirement.
Class Size	An independent algorithm used to calculate the single shift class size for each location per year.

APPENDIX L

HISTORICAL ANALYSIS

HISTORICAL ANALYSIS

Historical analysis is performed in relation to predecessor systems. The aim is to determine which characteristics of the predecessor system have been retained in the new system and to note where changes between the two systems arise. Basically, we are talking about a comparison of the two systems in terms of certain dimensions (e.g., number and skill level of personnel, equipment characteristics, etc.). Where changes have been made, it is necessary to determine their planned impact on behavioral aspects (e.g., on the ability of personnel to do their jobs within time constraints, the training they require, life cycle costs, etc.). The individual analyses required to determine this impact are described in other appendices.

DSARC Inputs

The information provided to DSARC describes:

1. Changes in the values of behavioral parameters between the predecessor system and the one under design.

2. Behavioral implications of these changes.

APPENDIX M

SCHEDULING

SCHEDULING

Scheduling involves correlating (mentally, not statistically) two or more inputs and/or events with each other and/or with some predetermined milestone chart to determine:

1. Whether inputs or events that are supposed to be coincident are or will actually be coincident.

2. Whether a required schedule will in fact be met.

The analytic process involved is simple comparison of calendars.

DSARC Inputs

The information provided to DSARC is that:

1. Specified milestone schedules will be met.
2. If not, the reasons for the failure to meet schedules.

APPENDIX N

INTEGRATION ANALYSIS

INTEGRATION ANALYSIS

Integration analysis is simply the determination of whether one or more required inputs have in fact been incorporated in a specified document. It is necessary for the analyst first to be able to recognize the nature of the required input based on its characteristics and then to know that the input must be joined with another input or inserted into a particular document. The effect of integrating inputs must also be indicated.

DSARC Inputs

The information provided to DSARC will indicate that:

1. Specified human factors information had been incorporated into required documentation.

2. The behavioral information included in the document has certain implications that are described.

DISTRIBUTION LIST

Deputy Under Secretary of Defense Research and Engineering (Research and Advanced Technology), Military Assistant for Training and Personnel Technology
Assistant Secretary of the Navy (Manpower, Reserve Affairs, and Logistics)
Principal Assistant for Manpower and Reserve Affairs
Deputy Assistant Secretary (Manpower)
Chief of Naval Operations (OP-01), (OP-11), (OP-12), (OP-110), (OP-115) (2), (OP-964D), (OP-987H)
Chief of Naval Material (NMAT 04), (NMAT 0722), (Director, Strategic System Projects (SP-15))
Chief of Naval Research (Code 200), (Code 440), (Code 442), (Code 442PT)
Chief of Information (OI-213)
Director of Navy Laboratories
Commander, Naval Military Personnel Command (NMPC-013C)
Commanding Officer, Naval Training Equipment Center (Technical Library)
Director, Training Analysis and Evaluation Group (TAEG)
Superintendent, Naval Postgraduate School
Commander, Army Research Institute for the Behavioral and Social Sciences, Alexandria (Reference Service)
Chief, Army Research Institute Field Unit, Fort Harrison
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base (Manpower and Personnel Division)
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base (Scientific and Technical Information Office)
Commander, Air Force Human Resources Laboratory, Lowry Air Force Base (Technical Training Branch)
Commander, Air Force Human Resources Laboratory, Williams Air Force Base (Operations Training Division)
Commander, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base (Logistics and Technical Training Division)
Director, Plans and Programs, Air Force Logistic Management Center, Gunter Air Force Station
Director, Science and Technology, Library of Congress
Defense Technical Information Center (DDA) (12)

FIL