AD-A222 565

(2)

OPERATIONAL SUITABILITY GUIDE

Volume II - Templates

MAY 1990

THE OFFICE OF THE DIRECTOR OF
OPERATIONAL TEST AND EVALUATION

THE OFFICE OF THE SECRETARY OF DEFENSE
WASHINGTON, DC 20301

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE
Operational Suitability Guide:
Volume II - Templates

5. FUNDING NUMBERS

6. AUTHOR(S)
Dr. Elizabeth Rodriguez

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Office of the Director of Operational Test and Evaluation
The Office of the Secretary of Defense
Washington, DC 20301-1700

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

VOL II of the Operational Suitability Guide was prepared by the Office of the Director of Operational Test and Evaluation (ODOT&E) to provide an overview of the major issues and potential trouble spots within the operational suitability arena. It addresses specific activities that are performed by the DOT&E Staff Assistant. The activities addressed are the review of TEMPs, OT&E plans, and OT&E reports. Volume II also addresses on-site observations of operational test activities.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Operational Test and Evaluation, Operational, Suitability, Test and Evaluation Master Plan, Test Report, Test Plan	100
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	PR	- Project
G	- Grant	TA	- Task
PE	- Program Element	WU	- Work Unit Accession No

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- **DOD** - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- **DOE** - See authorities.
- **NASA** - See Handbook NHB 2200.2.
- **NTIS** - Leave blank.

Block 12b. Distribution Code.

- **DOD** - Leave blank.
- **DOE** - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- **NASA** - Leave blank.
- **NTIS** - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

FOREWORD

The effective operational test and evaluation (OT&E) of defense systems is a critical part of the long-term program to provide for the proper defense for the United States. The Department of Defense has an established process for planning and conducting operational tests and for evaluating the data that result from those tests.

Volume I of this Operational Suitability Guide was prepared by the Office of the Director of Operational Test and Evaluation (DOT&E) to provide an overview of those issues that are included in the general subject of operational suitability, and to provide background information for DOT&E Staff Assistants to use when examining operational suitability subjects.

Volume II focuses on the specific activities that are performed by the DOT&E Staff Assistants. The activities that are addressed are the review of the Test and Evaluation Master Plans (TEMPs), the OT&E plans, and the OT&E reports. The Volume also addresses the on-site observation of the operational testing activities. The information in the Volume II is intended to supplement the policy and procedures contained in the DoD Directives and the guidelines that are contained in the DoD Manuals. This document does not establish new requirements for operational test and evaluation.

Volume II is organized in the following manner. Chapter 1, Introduction, discusses the arrangement and method for using the Volume. Chapter 2 covers the requirements for Test and Evaluation Master Plans (TEMPs). Chapter 3 addresses the content of OT&E Plans. Chapter 4 contains the information needed for observing the conduct of OT&E. Chapter 5 contains the information needed for reviewing OT&E Reports.

If questions or comments arise while reviewing or using this guide, they should be forwarded to the primary author:

Dr. Elizabeth Rodriguez
The Office of the Director of Operational Test
and Evaluation
Office of the Secretary of Defense
The Pentagon, Room 1C730
Washington, DC 20301
(202) 697-3895

THIS PAGE
INTENTIONALLY
LEFT BLANK

CONTENTS

LIST OF TABLES

THIS PAGE
INTENTIONALLY
LEFT BLANK

Chapter 1
INTRODUCTION

1.1 PURPOSE

The purpose of this volume of the Operational Suitability Guide is to assist Staff Assistants in the Office of the Director of Operational Test and Evaluation (DOT&E) in their review of OT&E documents and in observing OT&E events. It highlights important operational suitability factors for consideration during these activities. The guide is structured to aid in the review and evaluation of the three principal operational test and evaluation documents: Test and Evaluation Master Plans (TEMPs), OT&E plans, and OT&E reports.

1.2 DEFINITIONS

Operational Test and Evaluation is conducted to determine the operational effectiveness and suitability of weapons, equipment, or munitions for use in combat by typical military users. The Department of Defense defines operational suitability and operational effectiveness in the following manner:

Operational Suitability: "The degree to which a system can be placed satisfactorily in field use with consideration given to availability, compatibility, transportability, interoperability, reliability, wartime usage rates, maintainability, safety, human factors, manpower supportability, logistics supportability, documentation, and training requirements."

Operational Effectiveness: "The overall degree of mission accomplishment of a system when used by representative personnel in the environment planned or expected for operational employment of the system considering organization, doctrine, tactics, survivability, vulnerability, and threat (including countermeasures, nuclear, and chemical and/or biological threats)."

Operational effectiveness and operational suitability, in a strict sense, cannot be separated. There are elements within each that easily could be included in the other.

1.3 OVERVIEW

This guide is structured around the four key activities of the staff assistant: the three documents DOT&E must review (TEMPs, plans, and reports) and the observation of actual testing, with the intent of focusing on specific suitability issues that will be reported in the review of each. The guide is organized as follows:

Chapter 2 contains guidance information for the review of TEMPs.

Chapter 3 contains information for conducting the review of OT&E Plans.

Chapter 4 contains information for observing the conduct of Operational Testing.

Chapter 5 contains information for conducting the review of the OT&E Reports.

The subsections of Chapters 2 through 5 consist of templates that are subdivided into the following sections:

- The Introduction presents an overview of the subject of the template, e.g., what the particular part of the OT&E document should contain.
- The Area of Risk identifies the risk that might be encountered if proper test planning, conduct, or reporting of suitability concerns is not followed.
- The Outline for Reducing Risk directs the Staff Assistant to what might be included in a proper response to the subject of the template. Each of the items under the Outline for Reducing Risk is further emphasized by having an example (enclosed in a box) of its application to a DoD system.

1.4 HOW TO USE THIS GUIDE

The understanding of operational suitability and the meaning of the suitability elements is a critical component of suitability OT&E. The relationship of these elements to each other and to the successful introduction of the system into the operating forces are important and necessary parts of system acquisition.

To gain full value from the information within this guide, it is recommended that the reader review and evaluate the complexity of operational suitability by briefly scanning the templates in Chapters 2, 3, 4, and 5. A thorough front-to-back study of the document will then provide the reviewer with an understanding of the information contained in the guide, thus providing a mental road map for quick access to required reference material when and as it is needed.

The reader may wish to review what the thirteen elements of suitability are, what parameters are appropriate to measure those characteristics, and how thresholds for those parameters might be checked. The tutorial volume (Volume I) of the Operational Suitability Guide can be reviewed for this purpose.

If a specific document is under review, turn to the appropriate chapter: Chapter 2 for TEMPs, Chapter 3 for OT&E Plans, or Chapter 5 for Test Reports.

If information relative to observing Operational Testing is required, then review Chapter 4.

Chapter 2

TEST AND EVALUATION MASTER PLAN (TEMP)

The TEMP is an essential test and evaluation (T&E) document used by the Office of the Secretary of Defense (OSD) to support milestone decisions by the Defense Acquisition Board (DAB). The TEMP is the basic planning document for all T&E activity related to a particular system acquisition. It defines both Developmental Test and Evaluation (DT&E) and Operational Test and Evaluation (OT&E) associated with system development and acquisition decisions. The TEMP relates program structure, decision milestones, test management structure, and required resources to critical operational issues, critical technical issues, evaluation criteria and procedures.

One of the more significant functions of the TEMP is to document test and evaluation issues and criteria that will be considered in acquisition decisions. Thus, the reviewer of a TEMP must realize that the TEMP serves not only as a major control mechanism, but also to provide a clear correlation between issues and program objectives through test-verifiable criteria. In reviewing a TEMP in the area of suitability, one must ensure that it contains pertinent suitability-related information on the system's required operational and technical characteristics, test objectives, and the evaluation process.

Suitability-related system requirements, program structure, technical and operational characteristics, and associated thresholds in the TEMP must be reviewed to ensure that they are consistent with the Requirements Documents, Acquisition Decision Memorandum (ADM), and approved System Concept Paper (SCP)/Decision Coordinating Paper (DCP). T&E must be defined sufficiently to ensure that the test program will assess the effects of human performance on the weapon system's ability to meet all suitability standards, including reliability and maintainability.

The TEMP must include the system's suitability-related critical technical and operational issues and thresholds and their relationship to the system's requirements. It should clearly outline the planned T&E process.

A TEMP must describe the kind and amount of suitability test and evaluation, required resources, planned test locations and schedules. It must clearly relate T&E activity to suitability-related critical technical characteristics and operational suitability issues. It must describe the evaluation of the system relative to the suitability-related issues and the testing to be conducted to provide data to accomplish the evaluation. It must show the relationship between T&E schedules and program decision points and address the T&E to be accomplished in each program phase. It should identify the planned test articles to satisfy test objectives, as well as identify the number and rate of systems to be produced during the Low Rate Initial Production (LRIP) phase. Test resource requirements must be addressed, including known test resource shortfalls that may impede the full test and evaluation of the system suitability.

Finally, in reviewing a TEMP, one must be careful to avoid the pitfall of allowing the document to become an end unto itself. The document should define a test program that, when properly executed, will provide for accurate and efficient determination of a weapon system's operational effectiveness and suitability.

2.1 PART I, SYSTEM DETAILS

Part I of a TEMP provides details of the weapon system, its intended mission, and the required technical and operational characteristics. The mission must be adequately defined and key hardware and software features of the system must be described. Technical and operational characteristics discussions must include their relationship to the suitability aspects of system performance, as well as the effectiveness aspects.

AREA OF RISK

Heavy emphasis on effectiveness requirements may lead to oversights in suitability requirements.

The suitability implications of specified mission definitions and the system's technical and operational characteristics often are not well understood or clearly visible. During the early stages of a program, there is a tendency to place heavy emphasis on the definition and understanding of operational effectiveness. Later in the program, suitability deficiencies that significantly detract from system capability are identified.

OUTLINE FOR REDUCING RISK

The system's operational, support, and maintenance concepts should be examined to identify important suitability considerations, as well as the testability of key characteristics and requirements.

a. **Mission Description. (Sec. 2.1.1)**

Can key operational suitability issues be identified from the mission description in the TEMP?

The mission of the system and the planned support concept should be described in enough detail to permit the reviewers and the decisionmakers to understand the critical operational issues (COIs), including the suitability COIs. If the mission contains new or unique requirements, these should be examined to ensure that the test program has the structure to deal with these features.

> The mission of a weapon system is to penetrate enemy defenses and to conduct offensive actions behind enemy lines in an independent and self-sustaining mode for a period of 60 days. Critical operational issues must address the suitability areas of reliability, maintainability, and graceful degradation of the system as required by the self-sustaining nature of the system s mission

b. System Description. (Sec. 2.1.2)

Are the key system features that drive the suitability requirements included?

Critical operational suitability-related system characteristics may be omitted, resulting in inadequate planning for that portion of OT&E. The suitability issues may be improperly identified; as a result, a critical issue will be missed. Inadequate attention might be placed on the operational suitability issues.

> A key feature of the missile warning system (MWS) is improved operational availability. With the use of fault-tolerant computer hardware and software, the system R&M will significantly improve end-to-end availability for the MWS. The improved R&M is critical to meet the high level of operational availability.

c. Critical Technical Characteristics. (Sec. 2.1.3)

Do the technical characteristics support the operational suitability requirements?

The key technical characteristics should be described clearly. The rationale for each identified critical technical characteristic is important in understanding how the technical testing program fits into the overall acquisition program. This knowledge should then be used to understand the relationship of the technical testing to the key operational suitability requirements.

> The system has four critical areas of technical performance. The classified appendix lists these characteristics, and quantitative measures against test locations, schedule and the decisions supported, which provide the technical basis against which the system performance will be evaluated. These four areas are undetected message error rate, operational availability, processing time and system growth.

d. Required Operational Characteristics. (Sec. 2.1.4)

Are the required operational characteristics and their associated parameters listed?

The operational characteristics--with associated operational effectiveness and suitability parameters--that are critical to the mission performance and the ability to place the system into field use should be listed. Thresholds, which represent the level of system performance acceptable to the user to successfully execute the mission, also should be listed. The Service-prepared TEMP identifies the key required operational characteristics needed for operational mission accomplishment. One must ensure consideration is given to the testability of the requirements, utilizing the expertise of the development and test and evaluation communities, and check that the key required operational suitability characteristics deemed of critical importance to meeting the mission requirements are identified.

> The key required operational characteristics of the system include the following suitability requirements: system reliability, system maintainability, and logistics supportability. The values are contained in the classified appendix.

2.1.1 MISSION DESCRIPTION

The Mission Description section of the TEMP should briefly describe the mission of the deployed system, the threat it is required to be effective against, any threats that must be countered during the accomplishment of the mission, and the range of environmental conditions (weather, terrain, oceanographic, space) over which the system should be effective and suitable. As necessary, it should reference other appropriate, approved program documents. This might include the Mission Need Statement (MNS), System Concept Paper (SCP), Decision Coordinating Paper (DCP), or Service need statements, such as the Required Operational Capability Statement (ROC) or Statement of Operational Need (SON).

AREA OF RISK

The test program may not be structured to evaluate the system for its actual operational mission.

If the mission profiles are not clearly defined and understood by both the user and the developer early in the acquisition cycle, not only may the analyses and design activities be jeopardized, but also the test program may be planned in a manner that is not fully relevant to its intended use. In the suitability area, this mission definition needs to identify any specific reliability, maintainability, logistics, or other suitability area that has a critical relationship to the successful accomplishment of the system's mission.

OUTLINE FOR REDUCING RISK

Reducing the risk associated with the mission description and the associated suitability issues requires that the system's operational and support concepts be examined. Those factors that are important suitability issues should be identified and documented in the test and evaluation master plan.

a. Mission Definition.

Does the mission description include items that could be suitability issues?

The mission or missions that are planned for the system should be described in sufficient detail to ensure that critical operational suitability issues can be identified, and the context of these issues understood. The TEMP should be compared to other program documentation to ensure that it agrees with the previous mission descriptions. If the detail is inadequate, then the need for additional detail should be judged, based upon what information is available from other sources and if this information is in documentation that is controlled as to content.

> The GBU-15 is a modular, unpowered, air-to-surface guided munition designed for external carriage on F-4 and F-111 aircraft. It is designed for employment from low and high altitude and appropriate standoff ranges against high value targets. The addition of an infrared guidance section provides day/night capability as well as limited adverse weather capabilities. Suitability areas of particular importance include: maintenance concepts, environmental considerations, safety, and transportation (including handling).

b. Environmental Conditions.

Is the range of environmental conditions, over which the system will be effective, identified?

The mission description should include a discussion of the range of environmental conditions over which the system must be effective. This should include weather, oceanographic, space, terrain, obscurants, vegetation, illumination, etc., as is appropriate for the particular system in question.

> The vehicle loaded to the "full combat load" gross weight should be able to ford water to the depth of four feet and climb a 30 degree incline. The low-light-level rangefinder should be able to support the scout mission under the ambient nighttime conditions that are representative of conditions in Europe and the Middle East.

c. Definition of Logistic Support Concept.

Is the logistic support concept defined in sufficient detail to allow the planning of the operational suitability portion of the OT&E?

The planned logistic support concept for the system may require that unique aspects or approaches for the support of the system be realized. The mission description in the TEMP should highlight any of these aspects that are unusual for the type of system being discussed. The need here is to identify any unique aspects of the support scenario or concept in enough detail so that the test program can be planned to address these aspects. The detail must be at a level that shows the unique aspects and gives the planners a foundation to structure the OT&E.

> The F-XX is a light attack aircraft with leap-ahead combat effectiveness and battlefield survivability to defeat the threat of the mid-1990s. The F-XX has a worldwide operational capability and high sortie generation rates greater than 5 sorties per day in sustained operation. It will perform combat tasks in the close-in, deep, and rear battle environments. The support concept for the aircraft is a three-level concept. In contrast, major avionics items have the reliability and maintainability required to be supported using two levels of maintenance. (Avionics items to be supported by two levels of maintenance are listed in table V-1.)

2.1.2 SYSTEM DESCRIPTION

The System Description section of the TEMP briefly describes the system's design including key features and unique characteristics, interfaces with other systems, and unique support concepts. Key features and subsystems should include both hardware and software elements, as appropriate. Unique characteristics or unique support concepts should be identified if they result in the requirement for special test or analyses during test and evaluation. The relative maturity, integration, and modification requirements of any Non-Developmental Items (NDI) should be addressed. Interfaces with existing or planned systems that are required for mission accomplishment should be identified. Any interoperability with existing and/or planned systems of other DoD Components or allies should be identified.

AREA OF RISK

A poor system description may lead to improperly identified suitability issues or missed critical issues.

The significance of the System Description, relative to operational suitability, is that critical system characteristics may be omitted and the requirements for the operational suitability portion of OT&E may be improperly planned. The suitability issues may be improperly identified; as a result, a critical issue will be missed. Test planners and the decisionmakers might place inadequate attention on the operational suitability issues.

OUTLINE FOR REDUCING RISK

It is necessary to review the supporting documentation and ensure that the system description adequately describes the system's operational suitability features, and what is different about this system.

a. Key Features Description.

Does the description of key features of the system and its subsystems include those that relate to the operational suitability issues?

Key features and subsystems are those that allow the system to perform its required operational mission. The descriptions of the key features and subsystems provide a basis of information for assessing the adequacy of the OT&E program that is described in the TEMP.

> The mortar consists of a 55mm tube, a telescopic sight, and a two-piece baseplate. Eight are issued to 20-man mortar platoons. The mortar is designed to be hand-carried by two men. The ammunition weighs 25 lbs. per round The platoon will have sufficient HMMWV vehicles to transport every mortar plus its basic load of ammunition (50 rounds).

b. Relationship with Existing or Planned System(s).

Are the interfaces with existing systems identified?

The identification of the interfaces with other existing or planned systems results in a list of potential requirements for interoperability and/or compatibility issues. This list forms the basis for systems that must be acquired for use during the operational testing. If the systems are not available when required, then the potential for a serious test limitation must be examined. The electromagnetic environmental effects (E3) area can be a significant issue for some systems. Having complete knowledge of all other systems that will be operating in proximity to the system under test will help in defining the test objectives.

> Two communications systems, each with a different mission, entered development at approximately the same time The first was subjected to operational testing without the second system being present When a joint test finally was conducted, there was a serious incompatibility problem between the two systems. It was not possible to operate the first system unless the second system was turned off.

c. Unique Characteristics of the System.

Are unique characteristics of the system or unique support concepts identified?

Characteristics that are unique, different, or better in relation to operational suitability should be identified. Any unique suitability characteristics should be considered when the suitability COIs are defined. To form a basis for this consideration, these unique characteristics should be listed in the System Description. Also, these unique characteristics or support concepts should be considered when planning for the operational test events. Having an adequate description in the TEMP allows the remaining planning to proceed.

> A missile system has a unique Go/No-Go test to indicate the status of the system. It is important that the test be designed to clearly isolate a fault to the missile or the launching system Failure to isolate the fault in this manner will result in delayed or ineffective missions or the need for additional maintenance capability. The TEMP describes these features in sufficient detail to ensure such unique tests are addressed in the test plan

d. Changes over Previous System.

Are there significant changes or improvements over the predecessor system?

Major improvements in performance over the predecessor system can result in risk areas that should be the focus of operational testing. If such improvements are predicted in some of the suitability elements, the remaining elements will be adjusted to remain in balance. Thus, the major improvement must be confirmed during OT if the system is to be operationally suitable.

> An electronics system using new technology electronic components is forecasted to have a Mean Time Between Failure that is seven times that of the system it is replacing The maintenance manpower and the logistics support that is planned are significantly less than that of the predecessor system. The reliability of the system is identified as a COI.

2.1.3 CRITICAL TECHNICAL CHARACTERISTICS

The Critical Technical Characteristics section identifies the technical characteristics whose measurements are the principal indicators of the system's technical achievement. The section should identify performance thresholds and the milestones at which each of the thresholds has been, or is scheduled to be, demonstrated. Characteristics should be quantified when possible. Other program documents may be referenced for the technical characteristics, particularly if the technical requirements are classified. Technical characteristics usually are contract specifications and are derived from operational user requirements which precipitated the need for the system. Therefore, they may be traceable to the required operational effectiveness and suitability characteristics.

AREA OF RISK

Poor definition of critical technical characteristics can lead to a test program that is unable to assess program risk.

TEMPs frequently do not demonstrate the relationship between the technical characteristics and the operational characteristics, including operational suitability. The technical thresholds provided oftentimes are presented without a clear justification for their choice or an explanation of their significance or testability prior to the intended milestone. The purpose of technical testing is to reduce risk by assuring that a portion of the challenge of achieving the operational requirements has been met. To understand what degree of risk has been reduced, the relationship between the technical characteristics and the operational requirements must be discussed.

OUTLINE FOR REDUCING RISK

Having an understanding of the technical characteristics and the rationale or specific reference relating each identified critical technical characteristic is important to understanding how the technical testing program fits into the overall acquisition program. This knowledge should then be used to understand the relationship to the required operational suitability capability.

a. Critical Characteristics Listing.

Are the key technical hardware and software characteristics and thresholds listed?

Key technical characteristics are found in the system specifications and requirement documents. The TEMP should summarize them and relate them to the milestones in a matrix format. The matrix should indicate the characteristics that have been evaluated or that will be evaluated during the remaining phases of the developmental test (DT). Each technical characteristic should have a threshold value. If the technical characteristics are summarized in another acquisition document, such as a Baseline Correlation Matrix (BCM), this document should be compared to the TEMP. The technical characteristics that relate to operational suitability generally are quantifiable items, such as reliability and maintainability. One should examine these characteristics and determine how they relate to the operational suitability characteristics. What is their basis? Are they engineering estimates of expected performance required by contract specifications?

Since technical characteristics often are measured in a more pristine manner than exists in the operational units, the reliability and maintainability characteristics tend to be more optimistic than the operational characteristics. They should be examined to determine if the proper relationship exists.

CRITICAL TECHNICAL CHARACTERISTIC [Measurable Characteristic with Reference]	TEST EVENTS [Single Event or Test Phase]	TECHNICAL THRESHOLD FOR EACH TEST EVENT [Measurable Technical Achievement]	LOCATION [Test Facility]	SCHEDULE [Test Period]	DECISION SUPPORTED [Milestone, In-Process Review, or Major Event]
Reliability - Mean Time Between Operational Mission Failure (MTBOMF)	RGT I	200 Hours	Contractor's	1/92 - 7/92	
	RGT II	325 Hours	Contractor's	10/92 - 3/93	Service IPR
	FSD SYS DT	455 Hours	Test Bed	7/93 - 11/93	MS IIIA

b. Milestone Intervals.

Are milestones for demonstration of key suitability-related technical thresholds identified?

The technical thresholds should be tied to specific milestones. At those milestones, sufficient testing of the respective technical areas should have been accomplished and the results submitted to the decisionmakers. This schedule also should be compatible with the operational T&E of the companion operational suitability characteristics. The program milestone documentation should be checked to ensure that the technical characteristics that are important to critical suitability issues are addressed on a schedule consistent with the rest of the acquisition program.

> The Technical Characteristics listing indicates the decision that is supported by each of the technical thresholds. Phases of testing are indicated that will yield information to judge the system against the thresholds at each decision point.

c. Suitability Characteristics.

Do the technical suitability characteristics describe minimum acceptable system performance?

The technical characteristics should be quantifiable, where appropriate, related to system operational suitability characteristics, and provide the quantitative measures against which system performance will be evaluated during system-level technical testing. Technical suitability characteristics usually are defined differently than operational characteristics, and, as such, result in higher values.

> The Technical Characteristics listing indicates that the diagnostics false alarm rate threshold is 2 percent of all system test failures. For two BIT actuations, the false alarm rate shall be less that 0.1 percent. Diagnostics between different levels of testing shall incorporate tolerance levels that preclude retest-OK and cannot duplicate anomalies above the 2 percent rate.

2.1.4 REQUIRED OPERATIONAL CHARACTERISTICS

This section of the TEMP should list the key operational characteristics and their associated parameters as identified in the user's need statement and approved by the Service. A key operational characteristic is a principal element of the system's ability to accomplish its mission (operational effectiveness) and to be supported (operational suitability). These characteristics usually are defined by parameters that are indicators of the system's ability to accomplish its mission. If the Service has an approved matrix of operational requirements in the requirements documents, it may be used to display the required operational characteristics.

AREA OF RISK

Key operational characteristics may not be highlighted and the planned testing may not be sufficient to ensure achievement of the mission requirement.

Key required operational characteristics needed for operational mission accomplishment may address attention to the deficiencies of the current system. While characteristics often are drawn from Service-approved formal documents, the danger exists that the formal document may be incomplete or may be written in generalized, non-quantitative terms. Consequently, the operational requirements parameters selected for highlighting may not be sufficient to ensure achievement of the mission requirement. Since the operational suitability characteristics are seen by many people as secondary considerations, there needs to be assurance that adequate description and definition of the these items are included in the TEMP.

OUTLINE FOR REDUCING RISK

Key required suitability characteristics that are of critical importance to meeting the mission requirements should be identified. The description of each key suitability characteristic should include the appropriate parameters and thresholds.

a. Operational Suitability Requirements.

Are all operational suitability requirements identified?

Suitability issues are identified in the TEMP. The COIs should cover all suitability elements; the reason for eliminating any element should be discussed. Other operational requirements might be in the diagnostics area or in levels of support needs, e.g., numbers of maintenance people, test equipment, levels of maintenance, software change requirements. In reviewing the TEMP, the key is to ensure that the critical suitability areas are defined by operational requirements. Definitions should include all operationally relevant situations; e.g., all failures that can occur in operational use should be included in the definition for the operational suitability characteristics and parameters. This information should be a direct follow-on from the system's user requirements. It is very difficult for the operational test community to plan and conduct a meaningful operational test if the user community has not completely defined realistic quantitative needs.

The reliability of the communications system should be indicated by the mean time between operational mission failure (MTBOMF) values demonstrated The user equipment-requirement MTBOMF must be greater than 500 hours for both the manpack and the aviation sets The test is planned to demonstrate an 80 percent confidence in the MTBOMF values.

b. Parameters and Thresholds.

Are suitability characteristics supported by parameters and thresholds?

Each of the key operational characteristics should have an associated parameter and the parameters should be accompanied by a threshold (usually but not always quantitative). The thresholds should not be engineering estimates of expected performance or the performance that is required by contract specifications; they should be the minimum system performance acceptable to the user to successfully execute the mission. If only a component of a system is being developed, the thresholds should represent the essential operational requirements of the total weapon system. Failure definitions should be included by reference, and should reflect all failure modes or events that can be expected to occur in operational service.

REQUIRED OPERATIONAL CHARACTERISTICS
WITH THEIR ASSOCIATED PARAMETERS AND THRESHOLDS

CHARACTERISTIC	PARAMETER	THRESHOLD
Reliability	Mean time between operational mission failures (MTBOMF)	500 hrs.
	Mean time between maintenance actions (MTBMA)	32 hrs.
Maintainability	Maintenance manhours per operating hour	2.5 hrs.

c. Qualitative Suitability Requirements.

Are critical suitability areas expressed qualitatively?

With some systems, there are critical aspects of the system's suitability performance that can only be expressed qualitatively. There should be no reluctance to include such items in the list of required operational characteristics. The items that are critical to the system's ability to perform its mission within the intended field environment may require the attainment of objectives that are describable only in a qualitative manner. This could include requirements for compatibility with other systems or with certain skill level operating and maintenance personnel, or requirements that are in areas that have not progressed to the level where the requirement can be accurately quantified.

The system must be air-transportable This requires that the system must be able to be prepared for air shipment by organizational-level personnel and the system be air-transportable by C-130, C-141, and C-5 aircraft.

2.2 PART II, PROGRAM SUMMARY

Part II of a TEMP identifies responsibilities of the participating organizations, as well as the management and scheduling aspects of the weapon system program. Key factors to be discussed include T&E and acquisition strategy relationships, key decision points and associated reports, T&E requirements to support LRIP, constraints, test article and critical support resource availability, and the associated T&E responsibilities of all participating organizations.

AREA OF RISK

Suitability testing will not be performed at a time to support the program's decision milestones.

There is significant program risk if the suitability of the system is not demonstrated in time to support decision milestones. Typically, the operational suitability aspects of a weapon system program lag the effectiveness aspects of the program. Design of support equipment cannot commence until the design of the actual weapon system begins to stabilize. Maintenance manuals cannot be written until the support equipment is defined. As a result of this inherent lag, actual testing of the suitability aspects more often than not does not occur in concert with testing the effectiveness of the weapons system. If not appropriately planned for and scheduled, critical suitability elements, including diagnostics capabilities, support equipment, technical manuals, etc., may not be available. Therefore, suitability testing will not be performed in a timely manner to support the program's decision milestones.

OUTLINE FOR REDUCING RISK

Management responsibilities must be clearly delineated and include emphasis on the suitability aspects of the planned test program. The integrated schedule should indicate that the test program is not schedule driven, but accomplishment driven.

a. Integrated Schedule. (see 2.2.1)

Are all the required suitability activities clearly identified, appropriately time-phased, and adequately resourced so as to provide the required results at the decision milestones?

The integrated schedule should demonstrate that suitability testing will be consistent with the need for information at the key program decision milestones. The relationship of the test periods to the schedule should give an idea of the amount of test time that will be available to support the various milestone decisions and the associated reporting requirements. The integrated schedule should give an indication of when the various suitability elements will be available for OT&E; this includes when the support equipment and representative maintenance personnel will become available for the various levels of support, etc. The placement of these items on the integrated schedule will give the reviewer an idea of when various portions of the operational suitability evaluation will be done and when the results of this activity will become available for presentation to the decisionmakers. One must review the schedule to ensure the needed support to obtain suitability objectives has been identified; this includes testing and associated hardware and facility requirements down to the lowest practical level.

The integrated schedule for the M165 Truck program displays "on-dock" dates for the test articles to be used for maintenance tear-down and test personnel training. The sequencing of the events indicates that the required support elements will be available in a timely manner to support operational testing

b. Management. (see 2.2.2)

Are responsibilities for the operational suitability areas clearly identified and appropriately assigned?

One must ensure that the responsibilities are depicted, giving suitability objectives the needed attention, and that a sufficient number of test events are scheduled to provide a level of confidence in the resulting suitability statistics. The relationship between the key decision points and the specific T&E reports also should be examined.

Overall responsibility of the LX-21 Helicopter program is the responsibility of the Program Manager (PM). The PM establishes and chairs the TIWG and assures that adequate technical testing is accomplished. TRADOC represents the user and also is responsible for ensuring that the training of all test players meets or exceeds the minimum acceptable standards used to measure training effectiveness. Technical testing under TECOM's direction, includes...Reliability, Availability and Maintainability (RAM) measures. AVSCOM also is responsible for management oversight...to assure ..RAM maturity and growth is continuing as it pertains to flight handling qualities and airworthiness. U.S. Army OTEA will conduct Continuous Comprehensive Evaluation and the initial OT&E using LRIP aircraft USALEA is responsible for...ensuring appropriate logistics testing and evaluation are accomplished.

2.2.1 INTEGRATED SCHEDULE

The entire program schedule should be displayed on one page to include the integrated time sequencing of critical T&E phases or events, related activities, and planned cumulative funding expenditures by appropriation. Included should be event dates such as program decision milestones, test article availability, critical support resource availability; appropriate phases of DT&E, live fire T&E, and OT&E; rate production deliveries (i.e., low and full); Initial Operational Capability (IOC); Full Operational Capability (FOC); Low Rate Initial Production (LRIP); and beyond LRIP.

AREA OF RISK

Insufficient time and resources may be planned for suitability testing.

The integrated schedule may not cover all the necessary events or dates to communicate all of the aspects of system testing. The TEMP integrated schedule may be at such a summary level that it is difficult to determine the time and resources available for the suitability portion of the OT&E. Without this level of visibility, it might not be possible to assess the planned testing for the suitability objectives. In addition, critical suitability assets may not be available when required during the testing period. Such factors could result in the non-completion of testing and increase the risk of making inappropriate decisions at the decision milestones.

OUTLINE FOR REDUCING RISK

The compatibility of the schedule with the program decision milestones and the compatibility of the hardware development and testing with the software development and testing also should be evaluated. One should review the supplementary schedules and data source matrix (DSM) in the TEMP, as well as the timing sequences relating software test to hardware testing. The integrated schedule should be examined to determine if the time and resources that are available for the suitability portion of the OT&E are adequate.

a. Schedule Supports the Program Milestones.

Has sufficient time been scheduled to ensure the collection of meaningful suitability-related data? Will results be available prior to the requisite milestone?

The integrated schedule should demonstrate that the testing that will yield suitability results at a time consistent with the need for information at the key program decision milestones. The relationship of the test periods to the program schedule should give an idea of the amount of test time that will be available to support the various milestone decisions and the associated reporting requirements. (The details for suitability testing are discussed in sections 2.4 and 2.4.4.)

> The system IOT&E consists of 750 flight hours and will be completed at least 120 days prior to Milestone III; thus, there should be many opportunities for maintenance actions. The test phase will provide adequate maintenance data to present the suitability results to the Milestone III DAB.

b. **Scheduling Suitability Elements.**

Does the schedule include an on-dock date for support equipment? Will maintenance personnel training be completed in time to support OT&E maintenance-related testing?

The integrated schedule should give an indication of when the various suitability elements will be available for OT&E; this includes when the support equipment will become available for the various levels of support, when representative maintenance personnel will become available for the various levels of support, etc. The placement of these items on the integrated schedule will give the reviewer an idea of when various portions of the operational suitability evaluation will be completed and when the results of this activity will become available for presentation to the decisionmakers.

> For the IOT&E, the first- and second-level maintenance of the system will be performed by representative military personnel. Prototype test and support equipment will be available for the test phase. This phase will provide the first operational testing of the test equipment

c. **Scheduling Adequate Time.**

Is adequate time scheduled for the suitability portions of operational testing?

One should review the schedule to ensure that the support elements required to meet the suitability objectives have been identified; this includes testing and associated hardware and facility requirements down to the lowest practical level. If program schedules have slipped, then rescheduling of reviews should be examined to ensure that adequate time for OT&E conduct is planned prior to the scheduled review.

> The 750 flight hours that are planned for IOT&E will provide a sufficient number of the two designated system missions to demonstrate the system's reliability for the missions. Scheduled "on-dock" dates for support equipment, spares, T.O s, and training are supportive of the planned start of flight testing

d. **Software.**

Are adequate time and necessary resources scheduled for the planned operational testing to allow for software testing concurrent with hardware testing? Will the software be baselined and under configuration control prior to the start of OT&E?

One must ensure that the TEMP includes key software releases, subsystem/system tests, and sufficient key system events to coordinate software testing with the system schedule. The calender time available for software testing should be indicated. Resources necessary to support software testing must be planned for and scheduled.

> The integrated schedule provides visibility of the schedule for planned software block changes and clearly indicates a logical time-phased sequence for their introduction into the OT&E process in a controlled and organized fashion.

2.2.2 MANAGEMENT

The purpose of the Management section is to outline the T&E responsibilities of primary participating organizations (developers, testers, evaluators, and users). The T&E strategy should be related to the acquisition strategy of the program (any concurrency of production and testing should be discussed). The key decision points should be listed, along with the T&E reports that will support those decisions. Terms such as "Low Rate Initial Production," "Full Rate Production," and "Initial Operational Capability" should be defined quantitatively (rate and total quantity). The scheduled date (e g., fiscal year quarter) for the decision to proceed beyond LRIP should be identified. The management of schedule, resource, or budgeting constraints that may impact the adequacy of planned testing should be addressed.

AREA OF RISK

Inadequate emphasis may be placed on addressing the suitability issues of the OT&E.

The risk to operational suitability in the T&E management areas is that the management structure of the program may place inadequate emphasis on the suitability issues of the T&E. Some managers tend to focus attention on items that are important to evaluating operational effectiveness and pay little attention to the operational suitability objectives. This management emphasis is the result of the chronological sequence that places effectiveness before suitability.

OUTLINE FOR REDUCING RISK

Reducing of risk requires the manager to be held accountable for both effectiveness and suitability objectives over the life of the test. Accordingly, these responsibilities are to be clearly identified in the TEMP. Resources and schedules should provide the capability to perform the required testing.

a. **Management Responsibilities.**

Is there an outline of management responsibilities to ensure suitability objectives receive the proper attention?

The responsibilities that are spelled out should give the suitability objectives the needed attention. The proposed management team and approach will indicate the organization that will be conducting and contributing to the T&E in the suitability area. In some cases, suitability elements may be evaluated by supporting organizations. These organizations need to be identified in early versions of the TEMP to ensure that resources are available when required.

> The Test Integration Working Group (TIWG) will interface with the following functional groups: the Integrated Logistics Support Management Team (ILSMT), the Training Support Working Group, the MANPRINT Joint Working Group, the Computer Resources Working Group, and the Operational Test Readiness Review Working Group.

b. **Proposed Testing.**

Is it clear that testing is adequate in scope to provide confidence in the planned suitability results?

A sufficient number of test events should be scheduled to provide a level of confidence in the resulting suitability statistics. If significant levels of risk are the result of limited test assets, then these risks should be discussed.

> The IOT&E, as currently estimated, consists of ground tests, jettison/separation tests, captive flights and the launch/flight testing of 17 missiles using the F-18 as the carrier aircraft.

c. **Schedule Compatibility.**

Is the proposed relationship between the decision milestones and the T&E reports compatible?

The management section of the TEMP should indicate the planned sequencing of T&E reports from testing phases that support the major program milestones. The definitions of the program milestones will provide a context for assessing the adequacy of the suitability information that will be available at that milestone decision point.

> The schedule for an aircraft test and evaluation indicated that the test articles would be delivered to the Service DT&E on 1 June, where DT would be conducted until 1 August. The test articles were to be delivered to the OT&E site (400 miles away) on 4 August for 90 days of IT&E. The Milestone III was scheduled for 10 December. This schedule did not allow sufficient time for the deficiencies found in DT to be corrected before OT and did not allow sufficient time to complete the evaluation at the end of OT&E before the milestone decision.

2.3 PART III, DT&E

Part III of a TEMP is devoted to the Development Testing and Evaluation of a weapon system program. Responsibilities for its content rests with Deputy Director, Defense Research and Engineering (T&E). This section should summarize those activities planned for the development test phase(s). Knowledge of those activities can be of value to the DOT&E Staff Assistant in understanding the overall testing concepts and potential availability of data to accomplish early operational assessments.

THIS PAGE
INTENTIONALLY
LEFT BLANK

2.4 PART IV, OT&E

Part IV of the TEMP focuses on the Operational Test and Evaluation portion of the overall test program. It highlights the Critical Operational Issues (COIs), summarizes the OT&E that has been performed to date, and describes the OT&E that is planned.

AREA OF RISK

An ineffective test program may be structured.

The OT&E description may lack detailed information or communicate information that is unclear and ambiguous. Critical Operational Issues (COIs) may not be identified. Problems or limitations with past OT&E may not be included in the OT&E summary. Limited definition of the planned OT&E may result in the T&E "contract" being assumed to be one thing, while the plan is for something else.

OUTLINE FOR REDUCING RISK

a. OT&E Overview. (see 2.4.1)

Is the OT&E Overview complete?

This section should provide a summary of how the OT&E is structured. It should show how the program structure, test management structure, and the required resources are related to the system requirements, COIs, test objectives and decision milestones. It also should show how the completed OT&E has evaluated the system and how the future OT&E will evaluate the system. This overview should give adequate attention to the suitability COIs and discuss how the suitability evaluations/assessments will be provided for each of the decision milestones.

> A review of the TEMP indicated that the OT&E overview was an attempt by the system developer to justify performing a scaled-down OT&E.

b. Critical Operational Issues. (see 2.4.2)

Are all COIs identified?

Factors which could preclude suitability performance, as required by the users, must be identified. Identification and description of the related COIs are required to focus test resources and the attention of the decisionmakers on these important issues. Program aspects that may result in suitability-related COIs include a statement of higher levels of suitability performance than that required in previous similar systems, as well as the introduction of new or unproven technology into a system.

> Can the units that are equipped with the system achieve their peacetime and wartime system readiness objectives (SROs)? The system shall demonstrate a peacetime operational availability of 0 86 and a wartime operational availability of 0.78. The probability of successfully completing a three-hour mission shall be 0.70, with a mean time between operational mission failure of 8 5 hours

c. OT&E to Date. (see 2.4.3)

Is the summary of OT&E to date complete and accurate?

The description of prior OT&E should not be a duplication of the detailed OT&E reports. The discussion should summarize the prior operational tests, including what portions of the support system were tested and what the results were; what suitability COIs have been fully or partially addressed?; what were the results?; what planned activities were not performed, and why?

> The IOT&E was initiated in January of 1985 and was conducted for a two-month period. Insufficient data were accumulated to provide quantitative measures for reliability or maintainability. Five qualitative maintainability deficiencies were documented in the report for this short test phase.

d. Future OT&E. (see 2.4.4)

Does the discussion of future OT&E include a complete description of each OT&E test phase?

All future OT&E should be described and its specific purpose identified. Any major deficiencies should be addressed, as well as when their correction will be verified. Each phase of OT&E should be discussed separately. The configuration to be tested in each phase of OT&E should be identified. The portion of the support system that is present in each of these test phases should be discussed. The suitability objectives for each phase of OT&E should be listed. A brief description should portray how each phase of testing will be conducted (events to be performed, types of representative support personnel to be used, how the system maintenance and logistics support will be evaluated in this phase, and the role of suitability modeling and simulation in this phase). Those factors that limit the full and completely realistic operational test of the suitability aspects of the system should be identified.

> The IOT&E supports Milestone III and will consist of a 60-day, 300-flying-hour effort using three prototype aircraft. The test will be conducted between July and October 1990. Logistics and maintainability demonstrations will be conducted to provide an early assessment of the user's ability to maintain the system under a concept of two levels of maintenance.

2.4.1 OT&E OVERVIEW

The OT&E Overview should provide a summary of how the OT&E is structured. It should show how the program structure, test management structure, and the required resources are related to the system requirements, suitability COIs, test objectives and decision milestones. It also should show how the completed OT&E has evaluated the system suitability and how the future OT&E will evaluate the system. This overview should give adequate attention to the suitability COIs and discuss how the suitability evaluations/assessments will be provided for use in each of the decision milestones.

AREA OF RISK

An inadequate description leads to the approval of an unacceptable suitability test and evaluation program.

If the discussion in the OT&E outline is inadequate, or if the coverage of operational suitability is not adequate, then the information that is used to judge the acceptability of the TEMP and the test program also will be deficient. As a result, the test program that is approved may be unable to meet the needs of the decisionmakers.

OUTLINE FOR REDUCING RISK

To reduce the risk associated with inadequate testing, it is important to have an understanding of the testing already conducted and that which is planned. The OT&E Overview should include the contractor testing and early technical testing as well as early operational test and evaluation. The operational evaluation must take advantage of testing data from all appropriate sources up to and including the independent operational test and evaluation. Follow-on testing should be used as soon as it is performed to assist in the evaluation of the production articles. The Overview indicates the level of development of test articles being used in the operational test; for example, LRIP or prototypes. If the test articles will not be full-rate production articles, there should be a discussion of the differences between the test and full-rate production articles and the effects of their use during OT&E.

a. OT&E Overview.

Is the OT&E Overview in the TEMP a complete summary of the suitability T&E program?

The OT&E Overview should provide a clear understanding of the testing already completed, the testing yet to be completed, and who will process and evaluate the data. The Overview should be organized by acquisition phase and specific information requirements; for example, information to the system's developer should be annotated. Special suitability testing requirements should be discussed; for example, the impact on the test of an inadequate test support package, an independent contractor to perform IV&V on the system's software, or the use of the developing contractor to operate or maintain the tested system during the IOT&E.

The pre-Milestone II TEMP for an aircraft indicated in the OT&E Overview that an effort would be made to conduct early operational capability tests to provide the user with a perspective of the potential effectiveness and suitability during system's development It was stated that OT evaluators will participate in contractor and technical testing exercises, to include demos, surveys, and mock-ups. The Overview stated that an independent contractor would be used to conduct IV&V, and this contractor would be available to the OTA to assess systems' software suitability Because operating the system in a chemical environment is a major concern, the Overview provided a discussion of the operational testing of system's maintenance while in this environment The Overview also stated that the IOT&E would be conducted using LRIP aircraft (with a brief description of the difference between LRIP and Full-Rate production systems), and that resulting data would support the full-rate production decision at Milestone III

b. OT&E Summary.

Does the DOT&E Overview summarize the OT&E that has already been conducted, to include test articles descriptions and the future OT&E?

The Overview should discuss the OT&E conducted to date, by phase, and outline the successes and failures in achieving the operational suitability characteristics. If reviews and decisions were made that altered future system development or affected testing, there should be a discussion of how these issues will be addressed. The Overview should provide a test article description that states the level of development of the tested system, and a synopsis of the OT&E events with a discussion of the results of each event. Finally, the Overview should provide a summation of program management decisions that will impact on the acquisition schedule or operational testing and thereby require adjustments in test resources or test design and execution.

The future OT&E planned for the system must be applicable for the phases of the system's acquisition and in the time frames required. This section should provide a test article description of the system that will be tested in each of the future phases and a discussion of the OT&E objectives for each test. There also should be a detailed discussion of the OT&E events and scope of the testing, as well as a discussion of the basic scenarios that will be followed for each of the tests.

In evaluating the TEMP for a system, it was discovered that the future OT&E section stressed the completeness of past testing and attempted to justify a lack of planned future suitability OT&E for the system. The TEMP stated that because Service personnel would be participating in the DEM/VAL phase, have access to contractor IV&V data, and participate in demos, IOT&E should be reduced in scope to a 30-day field exercise. The TEMP proposed concurrent DT/OT prior to Milestone III, with each test organization having equal access to results Finally, it proposed that an Early Operational Capability Unit participate and maintain the system in IOT&E because this would provide a head-start on the training of the Initial Operational Capability Unit. Many of these statements tend to dilute the operational test by reducing the test time and resources, and by providing non-typical user troops.

2.4.2 CRITICAL OPERATIONAL ISSUES (COIs)

The Critical Operational Issues (COIs) are key operational effectiveness and operational suitability issues that must be examined in OT&E to determine the system's capability to perform its mission. The COIs are not characteristics, parameters, or thresholds, but they may have associated characteristics, parameters, or thresholds. COIs should cover all areas that critically affect the system's ability to accomplish its mission in the intended environment. The TEMP should identify which phase of OT&E will address each COI.

AREA OF RISK

Adequate attention may not be focused on some specific point that is important to the successful fielding of the system.

The significance of the COIs is that they help focus resources and management attention on items that are important in evaluating the system's progress toward attaining its operational objectives. The COIs allow the decisionmakers to have a smaller set of issues to address when making decisions on the acceptability of system development to date. Within this context, the risks are that the COIs may be improperly identified, or that a critical issue will be missed and the decisionmakers will not focus adequate attention on some point that is important to the successful fielding of the system.

OUTLINE FOR REDUCING RISK

The Critical Operational Issues (COIs) are the critical aspects of the system's operational effectiveness and operational suitability that are intended for examination and resolution during OT&E. Critical operational issues are developed by the tester and may be represented as questions that must be answered at the next acquisition decision milestone. The COIs are not characteristics, parameters. or thresholds, but they may have associated characteristics, parameters, or thresholds. The issues should cover all areas that critically affect the system's ability to accomplish its mission in the intended environment. The TEMP should identify which phase of OT&E will address each COI.

The emphasis of COIs is on the determination of the attainment of certain key performance levels and on surfacing potential problems that could interfere with successful mission accomplishment. Critical operational issues may change from one milestone to the next as some are resolved and new ones emerge in keeping with the systems development status. COIs must be structured to ensure that the information needs of the acquisition review body can be addressed for the milestone at hand.

Reducing the risk associated with OT&E COIs requires that the system's requirements, mission, and operating and support concepts be understood. Factors that are important critical issues related to operational suitability should be identified and documented in the TEMP. The full range of the intended operational environment must be considered. The list of COIs should be thorough enough to ensure that, if every COI is resolved favorably, the system should be operationally suitable when employed in its intended environment by typical users.

a. Completeness of List of COIs.

Is the list of suitability COIs complete?

The major risk in this area is the situation where some important suitability area has been overlooked and thus not identified in the list of COIs. As a result, attention is not given to this area when the Operational test plan is prepared, and the operational testing provides inadequate data to evaluate the condition at the time of an important milestone decision. The review of the TEMP should focus on identifying any critical suitability area that is not included in the list of COIs. Suitability risk areas may be identified from the areas of highest risk, or from the areas that are of the highest criticality within the support plans for the system.

> A remotely piloted vehicle (RPV) test program was planned to address three primary objectives, which were labeled as "critical" issues. Other issues including survivability, RAM, training, and human factors were to be addressed only to the extent that they affected the RPV's ability to meet the criteria of the three "critical" issues. Program milestone documentation did not contain explicit suitability criteria. The approach to suitability testing was to observe the ability of the RPV to support the mission under sustained combat operations, noting any shortfalls that could be attributed to suitability problems. This approach was justified by arguing that only "critical" issues required explicit criteria, and therefore suitability did not require explicit threshold values.

b. Suitability Requirements.

Are the suitability requirements at high levels compared to previous systems? Are they identified as COIs?

The planned levels of suitability performance (particularly reliability and maintainability) may be COIs if the system's successful operation is dependent on achieving a markedly higher level of reliability, maintainability, etc.

> The reliability requirement for the new system is significantly higher than the existing system (failure rate is one-third of the previous rate). A COI has been identified as "has the system achieved the planned level of reliability?"

c. System Technology Risk Areas.

Is new or unproven technology required for this system, or planned for use? Are there COIs on these technologies?

The use of advanced technologies in a system may introduce risk both in achieving system performance levels and in achieving the capability to support such new technology. Are there relatively unproven technologies in the system? Are these technologies understood from a reliability standpoint? Is the process to support these technologies (maintenance procedures, test methods, maintenance training, human factors, etc.) understood or demonstrated?

> A new, unproven cooling method is proposed for the IR seeker of a guided missile. The reliability of the cooling and the seeker is critical to the missile's operation. Therefore, the reliability and maintainability of the seeker and its cooling is specified as a COI.

2.4.3 OT&E TO DATE

Each completed phase of OT&E should be summarized. Descriptions of the hardware and software actually tested should be provided. Differences between the system used in testing and the configuration expected to be fielded should be highlighted, and potential impacts to suitability resulting from the differences should be discussed. The actual suitability testing that occurred should be summarized, including events, scenarios, resources used, test limitations, results achieved, and the evaluations conducted. Planned suitability objectives that were not met should be highlighted and explained. Status on the resolution of all suitability-related Critical Operational Issues should be discussed.

AREA OF RISK

The suitability-related requirements that remain for the future OT&E phases may be misjudged.

Discussions may concentrate more on the operational effectiveness requirements and ignore the suitability requirements. In some cases, the information may be condensed to such a degree that there is very little information provided to form a basis for judging the suitability-related requirements that remain for the future OT&E phases. If numerous phases of OT&E have been completed, the appropriate OT&E reports may be referenced and only the most recent or pertinent OT&E results included in the TEMP discussion.

OUTLINE FOR REDUCING RISK

Reducing the risk associated with suitability requires that the system's suitability-related requirements be tested. The test results should be examined to ensure the objectives of availability, compatibility, transportability, safety, human factors, interoperability, reliability, wartime usage rates, maintainability, manpower, training, supportability, logistics supportability, software supportability, and documentation have been met.

a. Summary of Actual Testing.

Is the previous testing for suitability summarized?

The previous testing should be summarized to include events, scenarios, resources used, test limitations, evaluations conducted, and results achieved. Specific OT&E reports that contain the detailed results should be identified. This summary should address the suitability areas adequately and should not be a summary of the effectiveness testing to date. Suitability elements that are included in each of the test phases should be identified.

> During phase IA of the OT, the suitability evaluation did not address the direct level of maintenance support because the test equipment that was to be delivered to the test site was not available in time to meet the dates that were mandated by the scheduled range times.

b. **System Configuration.**

Is the configuration of the test systems used in previous phases identified? Were similarities and differences with production or later test articles summarized?

The early phases of the testing are necessarily performed on systems that are not of the production configuration. Nonetheless, the data from these test phases are valuable. The key is to place the test results into context with the configuration of the test articles and the realism of the test environment. The summary of the prior OT&E phases should document these factors.

> The systems used in the Phase I IOT&E did not have the diagnostics software included in operational software for the mission computer. Therefore, the diagnostics objectives were not evaluated.

c. **Suitability Objectives.**

Are the suitability objectives of the prior OT&E activity described?

The discussion of prior activity can be relatively brief or can include references to previous reports. The critical questions are--what suitability objectives were planned to be addressed by the prior test phases, and what were the results?, Is there a discussion of significant events, test conditions, scenarios, resources used, limitations, and results as related to the system's suitability requirements?

> Phase I IOT&E was conducted on two of the initial prototype systems. Reliability data were collected during the test phase. Since the DT&E test team provided the system maintenance, no maintainability data were collected. A qualitative assessment was made by observing the maintenance as it was performed.

d. **Suitability Results.**

What were the results of the prior OT&E phases in the area of operational suitability?

The summary of the previous operational testing phases should be summarized and provide insight into the suitability issues that have been addressed and the results that were achieved. A listing of the COIs (including the suitability COIs) should be included. This listing should indicate which of the COIs were resolved (satisfactory, unsatisfactory, yes, no, etc.), partially resolved, or unresolved at the completion of that phase of testing.

> The threshold for this time frame was established as a system reliability of greater than 0.80, so the issue was satisfactory for the Phase I OT&E. (Since the thresholds will change for the various phases of operational testing, this success does not mean that the suitability will automatically remain satisfactory. The thresholds will become more demanding in the later phases of OT.) In the case of the maintainability COI, the qualitative maintainability assessment concluded that the access provisions for the required preventative maintenance tasks were unacceptable. Since the design of the access provisions were not projected to be revised in later development configurations, this deficiency was highlighted as a potential limitation on the system's operational suitability once it was fielded. The program manger was directed to resolve the deficiency.

2.4.4 FUTURE OT&E

All remaining OT&E required to resolve suitability-related Critical Operational Issues should be discussed. Operational testing should be described to verify the correction of major suitability deficiencies that were previously identified. A major deficiency is one that precludes the system from being designated as "operationally suitable."

AREA OF RISK

Testing may be structured so that a critical suitability issue will be missed.

Future testing may not be adequately described or the OT&E may not be directed at some key issue that must be examined before the appropriate decision milestone. The results of earlier testing may identify the need for testing of deficient areas that have since been corrected.

OUTLINE FOR REDUCING RISK

Suitability objectives for each critical issue must be reviewed. For the objectives not met to date, one must ensure that the future test program will adequately address the issues prior to the appropriate milestone. Those deficiency corrections dealing with suitability issues should be identified and testing planned to verify the correction. Each operational test should be described by the hypothesis being tested, the system configuration that will be tested, the scenario(s), and the sample size of the test (e.g., repetitions, hours, etc.).

a. OT&E Objectives.

Are the OT&E objectives in the area of suitability listed in sufficient detail to be addressed during the next phase of testing?

This section should summarize the objectives that will be the focus of the future phases of OT&E. The discussion should be a summary of the detail that will be found in the OT&E test plan. However, it must provide enough detail to ensure that the planned test program will meet all of the management information objectives. Objectives in the suitability area should be summarized to ensure that all critical areas will be addressed.

> Objectives included evaluating the R&M of the RPV system in its operational environment and the adequacy of the planned logistics support for the system

b. System Configuration.

Is the configuration of the test systems to be used identified? Are similarities to and differences between the production or later test articles summarized?

The key to placing the test results into context is the configuration of the test articles and the realism of the test environment. The differences between the planned test systems and the production systems should be clearly identified.

The OT Phase B will be conducted with 25 developmental sets. These sets will approximate the production configuration, except for the computer operating system. The software that will be used will have the Block II operating system instead of the Block III planned for the production units.

c. OT&E Events, Scope of Testing, and Scenarios.

Is the test scenario representative of the actual support environment and are suitability issues addressed for planned test events?

Evaluating suitability requires that the system be performing at a tempo that is representative of actual operation, and that the support needs posed by the test systems be comparable to that projected for the actual operation. The summary should indicate the type of personnel who will maintain the system, the status of the logistics support (e.g., unit level maintenance, unit and direct, etc.), the maintenance documentation that will be used, and the environment under which the system is to be employed and supported during the testing. If information from outside of OT will be used by the OTA to supplement the data from this OT phase, e.g., DT data, modeling and simulations (M&S), etc., these sources also should be identified. Any planned use of M&S should be identified, along with a reference to the M&S verification plan or, in the case of existing M&S, an explanation of when and by whom they were accredited.

The five test aircraft will be flown on representative missions of the planned duration. The systems will be exposed to the predicted combat operational stress. The sortie rate will be less than predicted, (only 1 to 1.5 sorties flown each day for each aircraft). The test stations that will be available at the test site are the projected support level for a 24 UE squadron; therefore, there will be second level maintenance capability that is in excess of that planned for the operational squadron. The evaluation will be adjusted for this difference.

d. Test Limitations.

Are factors identified that may preclude full and complete operational testing involving/concentrating on suitability issues?

Limitations listed in this section should indicate which of the OT&E objectives and which COIs will not be addressed during this test phase. The limitations could include threat realism, resource availability, limited operational environments (military, climatic, etc.), limited support environment, maturity of the test system, safety, etc., that will preclude a full and completely realistic operational test. The discussion of the limitations should address the impact of the limitations on the resolution of the affected COIs and the conclusions regarding the operational effectiveness and operational suitability of the system. The COI(s) that are affected should be identified after each limitation.

The number of available prototypes will limit the test to an examination of the performance of a single combat team. The interoperability of multiple teams will not be evaluated during this test period (COI S-6). The maintenance test equipment for the direct level of maintenance support will not be available in this test phase and the mean time to repair at the direct level will not be evaluated (COI S-2). The test site for this phase of OT does not include the variation in topology to evaluate the transmission capability under all of the terrain conditions specified in the users' need statement (COI S-8).

2.5 PART V, T&E RESOURCE SUMMARY

A summary of the suitability-related resources to be used during the test program should be provided. This should include major range and unique instrumentation requirements necessary to accomplish the OT&E suitability-related objectives. As system development progresses, test resource requirements must be reassessed and subsequent TEMP updates must reflect any changes.

AREA OF RISK

Timing and quantities of resources may not be adequate to ensure a realistic test.

The principal resource areas of operational suitability risk lie in the lack of resources to provide adequate test time or an environment that has insufficient realism. With the wide and complex array of responsibilities, test planning may fail to focus upon some of the operational suitability objectives that can be of critical importance in the operation of the system.

OUTLINE FOR REDUCING RISK

Sufficient time, test articles, and other resources must be scheduled to ensure adequate sample size and testing of the suitability characteristics of the system. Planned test article quantities, test phase duration, and other critical suitability-related test resource parameters should be identified.

a. Resources.

Are the number, timing, and configuration of hardware and software test articles specified? Are unique and/or modified hardware and software test support equipments identified? Are requirements for critical operating force support and special requirements identified? Are test support spares and repair parts provided for? Have system simulation requirements been identified?

The number and timing of hardware- and software-required test articles must be identified, and be adequate to provide the required suitability test data. Differences between test and production articles should be defined, with probable impact on suitability test results identified. Key support equipment and technical information, required for testing in each phase and for each major type of OT&E, should be identified. Descriptions should include any appropriate measure of test duration, e.g., hours, sorties, etc.

Unique or modified support equipment should be identified by test phase. Equipment requiring special calibrations should be identified by source, and calibration requirements specified.

The operational force support requirements should be identified for each phase for testing, e.g., aircraft flight hours, ship steaming days, T&E units, etc. The support to each test element or unit and test phase should be adequate for a credible operational suitability test in light of the operation and maintenance concepts planned for the system.

Any planned operational suitability system simulations, including computer-driven simulation models and hardware-in-loop test beds, should be defined. The system simulation requirements should be compared with existing and programmed capabilities. The process used to establish the credibility of the tool should be identified.

Special data processing equipment, special databases, and restricted/special use air/sea/land spaces should be identified and specified. The overall test data gathering, processing, and quality control procedures should be explained and should be adequate.

> One prototype system (which is functionally identical to the production configuration) with Block I software will be dedicated to maintainability testing during Phase I. A production model with Block III software (which incorporates automatic equipment reconfiguration) will be dedicated for continued maintainability testing during Phase II. The PROSTAR model will be used in conjunction with the ALTSTAR simulation for a broader assessment of logistics supportability issues. This combination of model and simulation has been verified and validated on the Air Force SEEK ELF program using actual field data. Required test support spares and repair parts have been identified. No special equipment, databases, or test support is required for OT&E suitability testing.

b. Budgeting and Scheduling.

Are appropriation line numbers provided for suitability-related resources and identified by fiscal year and program element numbers? Are need dates scheduled for key suitability-related test resources?

All costs for testing should be accurately identified by program element. All items, services and/or commodities should be included. Need dates for key suitability-related test resources should be documented, including such things as unique instrumentation, support equipment, simulators, models, and test beds.

> Funding for the IOT&E is identified as follows.
>
> | | FY89 (M$) | FY90 (M$) | FY91 (M$) |
> |---|---|---|---|
> | RCLR-1 (PE 64321) | 350.0 | 610.5 | 240.3 |
> | RCDR-1 (PE 64789) | 132.24 | 436.2 | 670.0 |
>
> Note: Funding for specialized maintainability testing instrumentation is included in PE 64321. Associated simulation, modeling, and test bed support is included in PE 64789. Reference the program schedule for key need dates.

THIS PAGE
INTENTIONALLY
LEFT BLANK

Chapter 3
OT&E PLAN

Test and evaluation plans are formal planning documents that provide a description of the test to be conducted and the evaluation methods. The test plan provides sufficient details about the planned test to assure the approval authority that test objectives will be addressed satisfactorily. It provides guidance to the Test Director with regards to test execution, test approach, sample sizes, operational environment, how the threat will be portrayed, instrumentation requirements, data collection, data handling, and test data presentation during testing. It also provides the measures of effectiveness and measures of performance, as well as the comparisons to be made. The evaluation plan, which may be a separate document, describes data handling and processing, and evaluation methods.

The operational test and evaluation, and the OT&E plan, support both development and production decisions made by decisionmaking authorities during the system acquisition process. The plan details the extent to which the system's issues and criteria will be addressed during individual phases of operational testing. Operational testing should address each critical operational issue and thereby support the evaluation of system operational effectiveness and suitability and the decisionmaking process.

There is no standard Department of Defense format for the OT&E plan. Each Service has developed its own approach to organizing the information that must be contained in the plan. Table 3-1 was constructed from the Services' policies on operational test and evaluation; it shows how the structure of this operational suitability guidance document relates to the major sections of the respective Services' operational test and evaluation plans.

This guidance document provides a process for reviewing an OT&E plan to ensure that the test program will provide for efficient and effective suitability test and evaluation. Critical Operational Issues are reviewed to identify potential implications in the operational suitability area. Once the focus of the suitability effort is determined, then the test objectives and the supporting measures of suitability are determined.

As much of the test and evaluation planning activity is directed at identifying and organizing the test assets (e.g., test systems, test ranges, supporting personnel, etc.) into planned test events and sequences, this area must be thoroughly reviewed. The structuring of the test activity should ensure that what is performed during testing is adequate; this requires realistic testing with respect to the planned operational environments and sufficient test data to give confidence in the results. An accurate and credible evaluation of the system's suitability for operational use then can be conducted.

Table 3-1 OT&E Plan Formats

OPERATIONAL SUITABILITY GUIDE	ARMY TEP (DAP 71-3)	AIR FORCE (AFOTECR 55-1)	NAVY (OTD GUIDE)	MARINE CORPS (MCOTEA)
3.1 Description of Test Articles	1.0 Introduction	I Introduction	1.0 Introduction 2.0 Administrative Information	1.0 Introduction 2.0 System Description
3.2 Scope of Test			3.0 Scope	3.0 Operational Test - Objectives and Issues
3.3 Operational Issues		II OT&E Concept		
3.4 Test Limitations				
3.5 Test Conduct	2.0 T&E Strategy 3.0 Test Design	III Methodology IV Administration V Reporting	4.0 Operational Effectiveness 5.0 Operational Suitability 6.0 Reports 7.0 Security	4.0 Evaluation Procedure 5.0 Operational Test Conduct
3.6 Data Management		Supplement E Data Management Plan		
	Appendices	Supplements	Annexes	

The templates that follow address suitability considerations for each of the major content areas included in OT&E plans. These templates are organized into the following sections:

3.1	Description of Test Articles
3.2	Scope of Test
3.3	Operational Issues
3.3.1	Test Objectives
3.3.2	Suitability Parameters
3.4	Test Limitations
3.5	Test Conduct
3.5.1	Test Scenario
3.5.2	Test Hours
3.6	Data Management

In reviewing the OT&E plans from the various Services, the overriding objective is to ensure that clear linkage is provided throughout the total test program. Simply stated:

- the operational issues must be clearly defined and testing must focus on each issue;
- test limitations must be considered and appropriate adjustments made to offset the effect of the limitations;
- testing must be planned for conduct in a controlled fashion to ensure valid and accurate data are generated; and
- data must be collected in sufficient quantity and in a controlled fashion to provide substance for meaningful evaluation.

In the final analysis, all aspects of the test program must link together to provide for a credible and defendable evaluation of the weapon system's suitability and effectiveness.

3-1. DESCRIPTION OF TEST ARTICLES

The OT&E test plan coverage of the test articles should include discussion of the number of test articles and any significant differences between the test articles and the production system to be fielded. It should discuss the planned configuration and integration of the test units. Test articles will vary in their integration and maturity level during the acquisition process; therefore, testing the suitability of these systems must take these differences into consideration.

AREA OF RISK

Inadequacies in test articles can limit the ability to test and influence test results.

The planned test articles may not be completely representative of the planned production article. The number of test articles planned may not be sufficient to adequately test the system under the intended operating environments. These limitations must be identified and considered in the test planning.

OUTLINE FOR REDUCING RISK

Reducing the risk associated with OT&E testing requires that the plan clearly communicate the test assets and identified limitations and risk areas associated with these test articles.

a. Number of Test Articles.

Is the number of test articles discussed and justified?

The number of test articles is most often a compromise between need and cost. The operational suitability implications of the test articles is primarily in the area of the statistical measures that require a number of trials or number of test hours to reach a level of confidence in the test results. The number of test articles that are proposed should be compatible with the requirements for confidence in the measures. The level of reliability to be demonstrated can be a major factor in determining the number of test articles required.

> During the system IOT&E, only five receivers will be available. Three will be deployed in the manpack configuration and two will be used in the vehicular installation. Operational suitability data will be collected to indicate any possible difference in reliability, maintainability or logistics supportability of the two different installations.

b. **Configuration.**

Is the configuration of the test articles comparable to the planned production configuration?

The configurations of the test articles and the production systems may differ. Differences should be reviewed and implications in the operational suitability area identified. The implications might mean that the systems will not be able to demonstrate compatibility with other items, such as test equipment, or the software will not have all of the features or diagnostics capability that is planned for the production systems.

> The test plan identifies the shortage of production isolator modules. The units to be used for the Phase IB testing will not have isolator modules of the latest configuration. Reliability data for the system will be adjusted to reflect the use of these older version modules.

c. **Alternative System Configuration.**

Does the test address all planned configurations of the system?

Some systems are planned for operational use in a number of different configurations or applications. The test scenario must discuss how each configuration will be tested.

> The system configurations consist of 1-channel, 2-channel, and 5-channel navigation signal radio receiver sets. The Army is the primary tester for 1- and 2-channel sets, the Air Force is the primary tester for the 5-channel airborne set, and the Navy is primary tester for the 5-channel shipboard set. Each Service's scenarios will include varying mission roles and mission types that could impact the operational suitability of the various configurations.

d. **Suitability Assets.**

Have adequate resources been provided for testing of the suitability elements?

The evaluation of operational suitability requires that the test be conducted with adequate operational suitability assets at the testing site. This could mean test equipment, maintenance personnel and facilities, or other items such as documentation. A limited number of assets could result in misleading data and inappropriate results.

> The second level of maintenance will consist of test equipment with the capability to verify failures and to isolate the failures to one Shop Replaceable Unit (SRU) 95 percent of the time. (It is projected that 5 percent of the Line Replaceable Units (LRUs) will have to be sent to the depot level for fault isolation and repair.) During the IOT&E Phase I, only one intermediate level test set will be available at the test site. The OT&E Test Director (TD) will determine which failed test units will be troubleshot and repaired at the test site and which are to be returned to the contractor for factory-level repair. The limitation on the number of test sets will be offset by the TD's allocation of repairable units. He will assure that a cross section of the failed units are tested and adequate suitability data are gathered at the test site.

3.2 SCOPE OF TEST

The test scope will provide a summary of the relevant information regarding the number of systems involved in the test, test location, and duration. The scope also will discuss the sequence and priority of the test phases and subtests.

AREA OF RISK

Limits and risks inherent in the planned test may be overlooked.

The scope of the test may be so limited that the system will not be exercised sufficiently to demonstrate its capability to meet the operational requirements. The description may be so abbreviated that the reader is unable to judge the limits and risks inherent in the planned test.

OUTLINE FOR REDUCING RISK

The plan should contain sufficient descriptive information about the scope of the test to indicate that the test scenario that is planned is adequate, that the test environment is representative of the intended operating environment, and that the test duration will allow operational suitability elements to be taken into consideration. Are the right factors and conditions included to ensure the system will be exercised sufficiently to capture the data needed to answer the suitability issues? The tactical context for the test should be discussed in terms of type, size of the military organizations to be represented or simulated, the operations to be conducted relative to the threat, interoperability element to be represented or simulated, and their composite relationship to the system under test.

a. Test Concept and Scenarios.

Is there sufficient information regarding the system's different configurations and concept of employment during the testing?

The test plan should include sufficient information regarding the employment methods planned for the test and the use of any different configurations of the system. The planned scenarios should be compared to the doctrine in sufficient detail to permit the examination of the test realism. The scope of the planned test should be described, including the number of test articles, the arrangement of the test assets, the number and type of operational scenarios that will be employed, the manner of supporting the system during the testing phase, and the range and variations of test environments that will be used. Also see "Test Scenario" (see section 3.5.1.a, p. 3-18).

> The system requires the employment of three major elements: a master station, user units, and the direct support team vehicle. Four configurations of the user units will be employed: manpack, surface vehicle, auxiliary ground unit, and airborne (fixed and rotary wing) units. The test site configuration employed will consist of up to 370 user units controlled by a master station and its alternate.

b. Support System Concept.

Will the scope of testing exercise the support system in sufficient detail to allow evaluation of suitability issues?

The scenarios and test events should include events that will trigger the use of support resources. The scope of testing should include the use of the support structure intended to support the system once it is fielded. Any limitations within this area should be highlighted.

> Five levels of maintenance are included in the planned support concept. The test plan describes specific maintenance actions to be performed and evaluated through the third echelon. Maintenance manuals will be evaluated for completeness and consistency with the planned maintenance skills and equipment.

c. Test Environment.

Is the planned test environment representative of the environment in which the system will be operated when fielded?

The plan for employment during the test must be compared to the intended operational environment including the doctrine, tactics, and threat. The scope of testing should be analyzed to ensure operational suitability issues can be assessed.

> The test will involve a Marine Amphibious Brigade-size unit operating in rocky and sandy terrain during a scheduled field exercise. Weather is characterized by warm days and cold nights, with possible precipitation. Special events will include the requirement for operating in an NBC environment, wearing gas mask and protective gloves Another event requires cold weather testing operating the system while wearing cold weather gloves, mittens, and inserts.

d. Interoperability Issues.

Are the interactions between the system under test and other systems within the operating environment consistent with assigned missions?

The interface requirements and operating scenario of the system should be examined to ensure that required interfaces are included in the test. If the entire range of interface requirements will not be tested, are there provisions to simulate the interfaces? One should review the data requirements to ensure that information concerning interactions among systems is captured.

> The test will examine the ability of the system to effectively transfer, receive, and/or process information within the system and with external systems. The test plan includes a special interoperability test using a specified test unit, the PLRS platoon, and other personnel equipped with the EPLRS.

3.3 OPERATIONAL ISSUES

Operational issues for the system are identified in the TEMP. The list of the Critical Operational Issues (COIs) should identify operational suitability features that are critical to mission performance and the ability to place the system into field use. The issues should consider the total system, including critical subsystems and the support items, the system's wartime mission requirements, and interfaces with other systems in the operating environment. Operational suitability issues are used in developing the test objectives or test issues.

AREA OF RISK

Major issues may be overlooked.

Risks associated with the suitability issues include clarity, coverage of all missions and scenarios, and coverage of the planned operational environment. The suitability COIs may be unclear or ambiguous if they are not described completely. There is risk that the suitability testing will be focused improperly. Major issues may be overlooked. Major suitability issues related to the total operating environment may not be included.

OUTLINE FOR REDUCING RISK

In order to reduce risks associated with operational issues, the test plan must have a thorough coverage of the COIs; this includes what is operationally critical, in terms of the system, its mission, requirements, the operating environment, and the supporting organizations and structure.

a. Critical Operational Issues.

Does the plan address the COIs that were identified in the TEMP?

The OT&E plan for the system should address the COIs identified and discussed in the TEMP. The COIs should be related to a particular phase or phases of testing for resolution. Does the OT&E plan contain test objectives to address these COIs?

> The system concept calls for two levels of maintenance for the elements of the radar. This is identified as a COI in the TEMP. The test plan has a test objective to evaluate the systems reliability and availability as it relates to the feasibility of the two-level maintenance concept.

b. Operational Suitability Issues.

How will the suitability issues facilitate answering one or more system critical issues? Are the suitability issues necessary?

There should be consistency between the critical operational issues in the TEMP and the suitability issues and objectives addressed in the OT&E plan. The issues should contribute added focus on approaches to assessing the operational suitability. Specific test events should address each issue and provide data to determine whether the system has satisfied each issue.

The system OT-III has the RAM operational issue, "What is the reliability, availability, and maintainability of the user equipment? The system must be sufficiently available to support the basic mission." To support this suitability issue, eleven OT&E test objectives were developed for assessment of RAM

c. Operational Issue Development.

Are the operational issues developed sufficiently to identify the operational suitability areas that should be addressed by operational testing?

The requirements documentation and employment doctrine should identify the elements of system performance that could impact the suitability of the system. From these descriptions, the suitability COIs should be identified. The test planning should relate to each COI and the rationale for selecting that particular issue. Each issue should be a focus for the test planning.

The test plan states "Can user equipment be effectively integrated into a wide range of weapons' platforms and function effectively in the operational environments of those platforms?" The rationale for selecting this issue states, "This COI reflects the versatility necessary for the equipment to meet the unique requirements of each of the Services involved in this program." It is apparent that there is an operational suitability issue in the area of interoperability.

d. Suitability Parameters. (see 3.3.2)

Are parameters defined for each of the suitability COIs?

To determine that suitability COIs have been satisfied requires the identification of parameters that can be measured and will provide the insight required to resolve the COIs. The four Operational Test and Evaluation Agencies have published a memorandum of agreement on reliability, availability, and maintainability parameters that will be used in multi-Service OT&E. This list of parameters is an excellent starting point for identifying the parameters that need to be examined to satisfy the suitability COIs.

The suitability parameters to be used during the test will be as follows.

Reliability	Mean Time Between Operational Mission Failure Mean Time Between Unscheduled Maintenance
Maintainability	Maximum Time to Repair (90 percentile)
Availability	Operational Availability
Diagnostics	Probability of Correct Detection Mean Time to Fault Locate

3.3.1 TEST OBJECTIVES

Test objectives provide an overview of what will be tested during a particular OT phase, as well as identifying the information required to evaluate whether a specific characteristic of the system meets the requirements. For each critical suitability issue there should be an objective which supports the decisionmaking process. For each test or test phase, the objective should be supported by a well understood test hypothesis. The hypothesis may not necessarily be a "statistical hypothesis." It should focus on the decision that will be made as a consequence of the test results. If the test does not influence a decision, but only provides useful data or information, then it is more properly termed an experiment.

AREA OF RISK

Faulty evaluation criteria will lead to faulty assessments.

Test objectives may be poorly developed or ambiguous. As a result, the evaluation criteria may not be properly defined, which could result in the inability to adequately evaluate test results.

OUTLINE FOR REDUCING RISK

Suitability test objectives should be well defined and descriptive of what is to be tested, as well as what data are needed to assess whether or not the objective is met. Objectives should be traceable to the COI or other issue that the objective supports. Objectives should be developed for each COI, and based on operational requirements or some other quantitative or qualitative measure.

For each test or test phase, there should be a description that includes:

- a well-defined test objective (or test hypothesis),
- the sample size planned (test hours, repetitions, etc.),
- the scenario planned for the test,
- the number of test articles,
- configuration of test articles, and
- what is missing from the system or what the differences are between the test articles and the planned operational configuration.

This description could be presented in a summary matrix.

a. **Objectives Consistent with Test Rationale.**

Are the suitability objectives clearly defined?

Each suitability objective should be clearly defined and support a critical operational issue. The objective and scope should be consistent, with the objective being more specific about what operational characteristic must be assessed. The hypothesis for each test should be clear from the test plan discussion How will the test provide the data needed to prove or disprove the hypothesis? The measure of effectiveness to be used in evaluating the suitability objective also should be stated.

> An objective states "Evaluate the capability to maintain the 1-channel manpack at the organizational and intermediate levels " Organizational maintenance will be limited to built-in-test fault detection/isolation, and battery and antenna removal and replacement At the intermediate maintenance level, there will be technicians and test equipment to perform I-level maintenance. Mean repair time will be the measure of effectiveness for on- and off-equipment maintenance. Mean repair time is the average clock hours required to return the system to a serviceable condition excluding administrative and logistics delay times.

b. **Objective Traceability.**

Does each objective relate to a critical operational issue?

It is necessary to review each objective and compare these objectives to the list of critical operational issue(s) to ensure that each appropriate COI is addressed. This review should result in each objective being traceable to a critical operational issue in support of the decisionmaking process.

> The mean repair time measure of effectiveness (MOE) is traceable through the reliability and maintainability objectives and (finally) to the reliability, availability, and maintainability critical operational issue.

c. **Data Requirements.**

Does the objective describe what data will be collected?

Data elements necessary to evaluate the objective must be identified. The information may be collected either by test representatives or as a result of the normal course of events. How the information will be used, as well as any necessary adjustments to the data, should be discussed. (This discussion may be integrated into the OT&E plan, in an appendix, or in a separate Data Management Plan.)

> The mean time to repair includes time to access equipment, troubleshoot, repair, and check-out Data elements required to be documented include maintenance actions, task times required to perform the repair and/or servicing of the user equipment receivers, the contributions of receiver built-in test, and intermediate test set capability to diagnose correctly the unit under test. The test team will prepare these data for input into the R&M database

3.3.2 SUITABILITY PARAMETERS

The suitability parameters are standards by which operational suitability can be gauged and evaluated. They may be quantitative or qualitative, and measure a system's performance or a characteristic of the system that indicates how well the system performs or meets a requirement. Depending on the operational suitability element being examined, there may be multiple suitability parameters that will support the decisionmaking process, especially when the suitability element is complex (e.g., reliability or maintainability). In addition, suitability parameters used in OT&E should be representative of parameters actually used in the operating environment.

AREA OF RISK

A distorted view of the system's suitability may be presented.

Inappropriate or inadequate parameters may be selected. The appropriate number of parameters should be selected to provide decisionmakers with a complete picture of the system's suitability. In addition, the test should allow the collection of adequate data to establish confidence in the test results. In the early stages of development, the operational suitability elements may not be fully developed and the maturity level may be inadequate for operational testing.

OUTLINE FOR REDUCING RISK

Selected suitability parameters should be related to the system's operational requirements. This allows the test team to define testing that provides measures for the evaluation of operational suitability issues.

a. Suitability Parameters.

Are the suitability parameters representative of those used in the actual operating environment?

The suitability parameters in the OT&E plan should be consistent with the operational suitability issues as developed in the TEMP. Multiple parameters may be required to adequately evaluate some of the suitability elements. Additional measures may be necessary for each support level. Volume I of the Operational Suitability Guide discusses suggested suitability parameters.

> To measure the maintainability of support equipment and its capability to support the mission in the tactical environment, the parameters will be mean corrective maintenance time (MCMT) at each level of maintenance, maintenance man-hours required per hour of operation (MMH/OH), and distribution of the maintenance workload and required maintenance capability at each echelon of maintenance support. The need for special tools, test equipment, and skills will be examined.

b. Performance and Supportability Parameters.

Is there a sufficient mix of performance and supportability parameters to provide the decisionmakers with information about the total system, as well as the support environment?

The test should exercise the system in its intended missions, preferably its combat missions as well as secondary missions. The maintenance and support procedures used in the test should be representative of those that will be used when the system is fielded. The suitability parameters must be fully defined in the test plan or in reference documents (e.g., the TEMP). Failure definitions and scoring criteria also must be included in the appropriate test documents.

> One operational suitability parameter will be mean time between operational mission failures. Data collected as a result of satisfying this requirement also will be used to evaluate the mean time between corrective and preventive maintenance. In addition, maintenance publication procedures will be evaluated as to their adequacy to support preventive or corrective maintenance.

c. Data Elements.

Have required suitability data elements been identified that will allow for the evaluation of each suitability objective?

The test plan should identify the data elements required to support the objectives, measures of effectiveness, or measures of performance.

> When measuring the mean time to repair, the data elements that will be required to be collected include date and time that the item is received at the appropriate maintenance activity, date and time that a maintenance action begins or resumes, date and time that a maintenance action stops due to administrative procedures (e.g., maintenance action ceases due to end of shift or lack of a maintenance technician), and date and time that a maintenance action is completed.

3.4 TEST LIMITATIONS

Test constraints or limitations can prevent or degrade the assessment of operational suitability and the resolution of test issues. Each constraint or limitation will result in objectives, or portions of objectives, that cannot be fully assessed. Constraints or limitations can result from the system itself, the availability of resources, environmental conditions, or the time available to conduct the operational test.

AREA OF RISK

Critical operational suitability issues may go unresolved during OT&E.

Constraints or limitations may not be identified as early as possible. Alternatives may not be addressed in test planning. As a result, critical operational suitability issues may be unresolvable during OT&E. This situation can result in systems being approved for fielding with significant potential for operational problems. Limitations must be known and understood.

OUTLINE FOR REDUCING RISK

It is not feasible from both a cost and test realism standpoint to design a test that will address every conceivable issue that should be addressed. However, once constraints and limitations have been identified, test planning can be focused to reduce the likelihood of significant voids in evaluation information.

a. System Configuration.

Is the equipment to be tested representative of the planned production configuration? What will be the impact on operational suitability evaluation?

The test plan should indicate the configuration of the test articles and compare them to the planned production configuration. The system to be tested may represent only a portion of the system that is planned for eventual operation. If significant differences exist, then the potential limitations to the suitability testing should be highlighted. Differences in configuration may lead to differences in reliability levels, or an inability to evaluate other suitability elements such as support equipment or documentation.

> According to the system test plan, the space segment is not yet fully operational. Only five or six satellites may be available during short periods of the 6-hour window. A minimum of four satellites is required to obtain precise three-dimensional position fixes as required by some users. The assessments of system availability issues will be limited due to the limited total operating time.

b. **Test Resources.**

Are the necessary suitability test resources planned and scheduled?

The test plan and other program or test documentation should be reviewed to understand the suitability resources required to conduct the test. The personnel selected to conduct the test should be representative of the operators and maintainers of the system once it is fielded. The required level of training for operators and maintainers involved in the testing should be discussed. The personnel selected to operate the system should be representative (e.g., experience, education, and career fields) of those who will operate the system in the field.

> The test will be conducted using 10 receiver sets of the Block IA configuration. The personnel involved in the test will have received contractor-conducted "pilot" training courses that are planned for the operational users and maintainers at the first two levels of maintenance.

c. **Test Scenario.**

Have the appropriate geographical settings been selected for conducting the test? Where possible, have locations been selected that take advantage of natural settings, as well as weather conditions (e.g., high altitude areas with snow, as well as hot and dry locations)?

The selection of missions and test sites should be reviewed to ensure that capabilities to be tested can be supported, and environmental characteristics such as climate, terrain, and foliage are representative of the planned operational environment. This review should take into consideration funding constraints, site availability, and unique site capabilities.

> If the system is planned to be operated in geographic settings such as desert, tropical, populated, jungle, and arctic, then testing should be conducted in areas that are representative of these geographical settings.

3.5 TEST CONDUCT

The Test Conduct section of the OT&E plan should describe the general approach to be used to conduct the test, including testing for suitability. The description should include the scenario, environment, threat, tactics and doctrine to be used, and requirements to be met. It also should highlight constraints or limitations that could affect the realism of the test.

AREA OF RISK

Data that are essential for suitability evaluation may not be provided.

The planned test may not be realistic enough to provide the data that are essential for suitability evaluation. The data must be useful in evaluating the system's potential suitability for the intended operational environment. The OT&E test plan also may be of less detail than is needed to provide an understanding of the specific actions required during the conduct of the test.

OUTLINE FOR REDUCING RISK

The discussion of the suitability-related testing should be broad enough for the reader to understand the realism of the test. Test scenarios should be planned considering the operations that will be required of the system in combat, as well as the types of threat it is likely to encounter. Personnel operating and maintaining the system during testing should be representative of those who will operate the system once it is deployed. Personnel should be trained in advance of the test and be thoroughly familiar with the test plan.

a. Test Scenario. (see 3.5.1)

Are the factors and conditions of the test scenario representative of those that will be present in the actual operating environment?

The plan for operational suitability testing must take into consideration the range of environments (tactical, climatic, etc.) that the system will be exposed to when operationally employed. From an operational suitability standpoint, it is important to demonstrate the ability to support the system using the planned support structure in various operating conditions. If unrealistic test missions are used, then the suitability assessment can be optimistic or unfavorable to the system. The missions should be reviewed to ensure that the factors and conditions sufficiently portray the intended operating environment and tempo.

> The land navigation user equipment tests will be conducted along with some HELO tests. Missions will be conducted with representative operational user personnel. The aviation set will be tested in the UH-60A configuration Test events will be conducted in accordance with the operational mode summary and mission profile of representative operational users The support for the first two levels of maintenance will be in accordance with the planned support concept.

b. **Test Articles. (see 3.1)**

Does the OT&E plan include an adequate description of the number and configuration of the test articles?

On major weapons, the test article normally consists of "worked over" prototypes; therefore, test articles are a different configuration and some are without subsystems. Test article configuration should be examined to determine its effect on the test scenario, realism, and data collection. Test articles often are early prototypes and often will not be representative of the fielded system. There are examples where hundreds of modifications have been made to test article configuration before the systems are fielded. The Staff Assistant may wish to estimate the impact of these configurations on the system's evaluation.

> Prototypes of a new aircraft have been developed by the two competing contractors. Developmental testing has been progressive and both contractors consider their prototype aircraft to be "representative" of the system to be produced. Several subsystems (including the weapon release system and the on-board diagnostics system) are being developed and are expected to be in an operational configuration in the eighth production aircraft. An aircraft with all systems completely representative of the final configuration will be available when the tenth aircraft is delivered in another three years. A DAB to approve the beyond-limited-rate initial production is planned in 18 months. (A total of 50 production aircraft are planned.) Consequently, the OT community is faced with the dilemma of a "concurrent" program structure wherein a complete operationally representative system will not be available for actual field tests before the scheduled Milestone IIIB decision.

c. **Test Hours. (see 3.5.2)**

Are the planned test hours sufficient for an adequate operational suitability evaluation?

The number of test hours required to produce the necessary data to evaluate operational suitability is directly dependent on the system's characteristics and the mission to be performed. For high reliability systems, the number of test hours required to verify the level of reliability may be quite high. In addition, if few maintenance actions are to be performed during the OT activities, it will be difficult to provide a complete evaluation of all suitability elements. The OTA should justify the adequacy of the test hours available for the OT.

> The number of test hours scheduled for the system OT was limited due to restrictions on test funds and limited spare parts. Had the OT been conducted as planned, the spares and test hours may have been sufficient; however, because there was a test schedule slip, the hours were not sufficient to support the prolonged test schedule. The Test Director made a determined effort to execute the test as planned. He was willing to except test data, although some were of limited value. Hours became the main driving force in the test, compelling the Test Director to except a lot of unusable data. The SA should ensure that test hours, as with any test resource, are not the driving force in test execution. He should ensure that data collection, test realism, and, only then, schedules are the primary considerations in the OT&E.

3.5.1 TEST SCENARIO

A test scenario is a set of circumstances by which a system is tested in a representation of its intended operational and support environment. The test scenario should be based on mission scenarios, support concept, critical operational issues, objectives, and test limitations. Consideration must be given to terrain, weather, and other operational factors.

AREA OF RISK

The test scenario may fail to reveal operational suitability deficiencies.

The operational test scenario may lack the detail and realism needed to reflect the intended operational use of the system. In addition, it may not provide for adequate operation of the total system to exercise realistic demands on the support system.

OUTLINE FOR REDUCING RISK

Scenarios should be based on realistic factors and conditions that will be present in the intended operating environment. Mission roles and types should be used and varied in such a manner to ensure that test objectives and, ultimately, the critical suitability issues can be answered with an acceptable degree of confidence.

a. Test Scenario.

Is the test scenario representative of the planned operating scenario and broad enough to allow for sufficient data collection to analyze suitability issues?

The degree of realism of the test scenario and mission determines the extent to which the support system and operational suitability can be evaluated. If the missions used are unrealistically severe, then suitability may appear to be less than it really is. If the missions are less severe than the expected operational missions, then the evaluation of operational suitability elements may be optimistic compared to what is achievable in the operational environment.

> The system tests will be conducted with a cadre of command and control personnel, equipment, and sufficient user units and operators to bring the community up to 370 user units. There will be no opposing force except for the EW jammers during this period. The scenario calls for two days of scripted defense and two days of scripted attack over approximately the same area as the Phase I test. Tests will be conducted 12-hours a day, which will allow data processing overnight prior to continuing the next day of testing.

b. Test Events.

Do the test events represent the minimum number of mission types that can be expected to be performed with the system? Are affected operating activities identified and test events planned for each?

The planned testing should include a representative set of the planned missions. There should be planned test events for all system functional elements; they should involve all mission types. The events must be representative to ensure that suitability can be evaluated.

> As a minimum, 16 corridor and six linkup aviation mission types, under nap-of-the-earth benign conditions, will be accomplished using nine pre-surveyed waypoints. In addition, a minimum of six corridor and six linkup mission types, under nap-of-the-earth electromagnetic conditions using nine pre-surveyed waypoints, will be accomplished. All operational missions will include en-route navigation, present positioning data, and termination fixes.

c. Maintenance Concept.

Is the maintenance and support concept to be used during the testing an accurate representation of the operational environment?

Maintenance support to be provided during testing should be representative of how maintenance will be performed after the system is fielded. Operational suitability elements (e.g., test equipment, maintenance documentation, and man-machine interfaces) that impact the effectiveness and efficiency of maintenance operations should be included in the evaluation.

> The system has been designed to use the established five echelons of maintenance support. During testing, first echelon maintenance will be performed by the radio operator for all ground configurations and by the organizational maintenance activity of each user of the airborne radio set. Second echelon maintenance will be performed by the Electronics Maintenance Company and will consist of operating and performing checks to determine fault areas prior to evacuating to third echelon maintenance. Third echelon maintenance shall be performed by the contractor and will not be evaluated in this phase.

d. Test Environment.

Were all potential operating environments considered in the construction of the test environments?

The planned operating environment can be described in many ways, including weather and geographical conditions, electromagnetic conditions, and battlefield conditions including smoke, noise, and CBR. The operational testing usually is limited in the range of environments that can be addressed. Therefore, it is important to assure that those conditions that potentially could have adverse impacts on the operational suitability of the system are addressed. Adequate consideration of these factors should be included in the test scenarios.

> The Air Force testing of the GPS UE requires support from the Air Force Electronic Warfare Center to test the susceptibility of the 5-channel airborne set to jamming of the satellite downlink signal. The Air Force will provide the technical expertise on the selection and operation of an airborne jammer. The jammer will be operated in accordance with Warsaw Pact radio electronic combat doctrine during six bombing missions, and three level and three loft profiles.

3.5.2 TEST HOURS

The number of test hours planned for the system under test is directly influenced by the mission types and roles that the system must perform. The length of time required to complete a mission plus the number of missions in the scenario should be representative of the mission lengths and duration that will be required in the intended operating environment.

AREA OF RISK

Operating time will not provide representative data and thus will lead to faulty conclusions.

The number of test hours or the mission duration may not be adequate to provide the necessary data for evaluation. The amount of test data that is available for evaluation is directly tied to the number of test hours. To provide the decisionmaking authority with a realistic assessment of the operational suitability of the system, careful consideration must be given to mission length and duration.

OUTLINE FOR REDUCING RISK

The number of test hours should be sufficient to demonstrate the system's operational and physical characteristics that can impact the operational suitability of the system. The test hours should take into consideration the impact that a short test will have on statistical confidence, as well as the long term effects on operational suitability.

a. Hours to Demonstrate Characteristics.

Is the number of test hours sufficient to demonstrate the operational suitability characteristics?

Operational suitability characteristics, such as reliability and maintainability, require relatively lengthy test periods to adequately demonstrate their achieved levels. In addition, other operational suitability items, such as technical data, have many alternatives, sections, or paths that cannot be exercised in a limited period of operational testing. The test plan must deal with the need for lengthy test periods and realistic levels of test resources.

> A limited number of test hours will be available due to small numbers of host vehicles, limited satellite coverage, and a compressed test schedule To supplement the operational testing data, additional data will be available from a number of platforms not directly involved with this dedicated OT&E This additional operating and failure data will be used to expand the system suitability database. Careful scrutiny and evaluation of the supplemental test data will increase the confidence level of the suitability results.

b. **Long Term Effects.**

Is the amount of time dedicated to testing individual test articles sufficient to evaluate the long-term effects on the articles?

Test items may have risk areas that require knowledge of the effects of lengthy periods of operation. Although statistical methods sometimes allow the testing of a large number of articles for relatively short periods of time, there is still a possibility that unknown risks will be realized when the system is exposed to longer periods of operation. When the OT&E plan calls for short periods of testing, the effects and risks associated with use of the system for longer periods of time must be considered.

> A selected number of the test articles from the OT&E Phase IA will be provided to support the test The use condition will be similar to that employed in Phase IA. Failure data from this extended period of operational use will be evaluated to determine if any long-term failure modes are likely to exist. The results of this additional operation will be provided as an appendix to the OT&E report.

c. **Statistical Confidence.**

Is the test time sufficient to ensure any required test confidence?

The system's operational requirements for reliability and maintainability may include the need for confidence levels. Confidence levels also may be required as part of the reporting requirements. The length of the operational test should be reviewed to assure that the required level of confidence can be achieved. The plan should state any requirement for confidence levels to be calculated for operational suitability measures. DoD 3235.1-H may be used as a reference in assessing the confidence levels and test times included in the OT&E test plan.

> The mission reliability of the system shall be measured in terms of the mean time between operational mission failures (MTBOMF). The point estimate MTBOMF that results from the scoring conference data shall be used to compute an 80-percent lower one-sided confidence interval. The 80-percent confidence value shall exceed the requirement stated in the approved ROC. Operational mission failure will be determined by a formal scoring conference in accordance with the approved failure definition and scoring criteria.

3.6 DATA MANAGEMENT

This section describes in general terms how the testing organization plans to collect, organize, reduce, verify, manage, control, analyze, and store the data needed to perform the evaluation. Suitability evaluation requires data on many aspects of the system's operation and on the support elements. The data management plan must provide for the proper collection and control of the test data.

AREA OF RISK

Inappropriate or incorrect data may be used for critical evaluations.

Poor data management can result in ineffective and inefficient data collection, control and reduction, and data evaluation methods that are unable to support the decisionmaking process.

OUTLINE FOR REDUCING RISK

The suitability data gathered during testing must be properly managed to ensure credibility and validity of the conclusions about the system being tested. In determining the proper types and amounts of suitability data to be collected, and the way in which the data are controlled and evaluated, a well defined data management plan is required. This plan may be part of, or a supplement to, the test plan.

a. Quality Control.

Does the data management plan include quality control methods to ensure incorrect or inaccurate suitability data are eliminated?

The data management plan should make provisions for tracking the data from initial receipt until final input into the master database. The plan should include procedures for collection forms, data verification, data reduction, and entry and editing of the data.

> When suitability data discrepancies are discovered through regular quality control checks of the data, the data will be returned to a previous step in the data flow process for resolution Ultimate control of the data is the responsibility of the quality assurance test manager; however, the day-to-day quality control of the RAM data is the responsibility of the RAM data manager. Data collection personnel will perform quality control checks to identify missing or logically incorrect entries. Quality control during the data reduction process will be accomplished by both manual and automated methods. Data managers will ensure data reduction procedures are consistent across missions and that results are checked by someone other than the person doing the actual data reduction. Scrubbing, correlation, and merging of automated data from various sources with other data, both manual and automated, will be controlled by computer programs, when possible. The program(s) will be verified to be functionally correct prior to the pilot test. Data entry and edit routines will be controlled and monitored by computer programs. No data will be loaded into the master database until all data entry errors identified by quality control procedures have been corrected

b. Data Collection and Validation.

Have procedures and responsibilities been established for manual and automated suitability data collection? Does the test plan or data management plan describe the forms that will be used in the data collection process?

The test plan or data management plan should include procedures to control the data collection process for both manual and automated data. These procedures should address how the data are to be organized, reduced, verified, managed, controlled, and stored. The process to include or exclude data from the database should be described. How will "no tests" be identified? The data collection forms should have been developed and approved, and the responsibilities for data collection defined. Test personnel should be trained in manual and automatic data collection prior to the start of the test. Sections 4.2.c and 4.7 also discuss data collection and validation.

> A data collector will be assigned to each of the systems under test for the purpose of RAM data collection and to track equipment operating time and events at the time of a RAM incident. RAM data forms will be based on forms from the RAMES management information system, but tailored for use on the system. The RAM data forms also will be used to collect RAM data on the MX-379 test set and other support equipment RAM forms will be supplemented by narrative commentary on the acceptability of the technical documentation and special tools used for each maintenance action.

c. Data Processing and Analysis.

Are data processing procedures described in sufficient detail to allow assessment of the procedures?

The planning should include data handling procedures for both manual and automated suitability data. The procedures should include the necessary quality control for the data flow from the data collectors until it is finally entered into the master test or suitability database. Procedures for analyzing suitability data should be described in sufficient detail to provide an understanding of how the data will be reduced and used in critical evaluation. Equations and algorithms should be addressed.

> Data management personnel will log all manual suitability data forms provided by the data collector, review them for consistency and completeness, and add any additional information that may be needed for internal control. If there are problems, the forms will be returned to the data collection section for resolution. Once the forms are determined to be correct, the data will be entered into the microcomputer database. Automated quality control and edit routines will ensure data validity. The quicklook performance and RAM reports will be generated for analysis and validation by data managers.

THIS PAGE
INTENTIONALLY
LEFT BLANK

Chapter 4

OT&E OBSERVATION

In addition to reviewing test documentation such as the TEMPs, OT&E plans, and OT&E reports, the DOT&E Staff Assistants have an important responsibility in performing on-site observation of the actual conduct of the operational testing. (Title 10, U.S. Code, Section 138 states that the DOT&E may require that observers who he designates be present during the preparation for and the conduct of the testing part of a DoD OT&E.)

Test program reviews for programs assigned to an individual Staff Assistant should be planned to ensure that appropriate priority is assigned for the coverage required for specific tests. Considerations include which test events require a DOT&E presence, and the degree of on-site monitoring which will be required. Given that a particular event is to be covered, the level of involvement may vary considerably. The Staff Assistant may arrange for a team of DOT&E observers to be on-site for the duration of the test, or there may be only one DOT&E observer on-site during only one or more of the critical testing periods.

Once the particular test event is identified and the degree of DOT&E coverage has been decided, then preparation for the test observation is required. This Chapter provides guidance to assist in preparing for on-site visits, and it provides information on items to be considered while at the test site. Once at the site, there are general areas that need to be considered. The management of the test site should be examined, and there are specific items in the test events that need to be compared to the discussions in the test plan. Finally, documentation that is used during the operational test events, the personnel that are involved in the operation and the maintenance of the system under test, the data collection program, and the test scenarios also must be considered.

This material was prepared within the context of this Operational Suitability Guide and has as its focus the oversight of areas of testing that impact on the operational suitability results. However, most of the items in this Chapter apply equally well to the testing for the operational effectiveness objectives.

The templates that follow address suitability considerations for a number of the activities that are involved in observing the operational testing. These suitability considerations are:

4.1 PLANNING FOR TEST SITE VISITS

This section discusses considerations for the general "test observation" aspects of the OT&E oversight responsibilities. Observe the actual operational test activities is needed to provide for credible analyses in DOT&E-prepared Beyond-LRIP, annual, and other DOT&E reports. It also provides first-hand knowledge to support other DOT&E responsibilities such as review/concurrence on the Test and Evaluation Section of the Congressional Data Sheets and Selected Acquisition Reports (SARs).

AREA OF RISK

Poor preparation for on-site visits may result in invalid conclusions.

A prime contributor to this risk concerns the balance between the amount of time dedicated to day-to-day Pentagon responsibilities and time expended in the field. The ability to ensure accurate and credible DOT&E reports can be strengthened by first hand observation of test events. For the on-site visit and observation of the test activities to be effective, preparation must be made prior to arrival at the test site.

OUTLINE FOR REDUCING RISK

a. Awareness of Current OT Schedule.

Are current test activity schedules available in the DOT&E office?

Test schedules in the TEMP as well as detailed schedules for discrete test events are subject to long- and short-term revisions. Attendance at selected Test Plan Working Groups (TPWG) or Test Integration Working Groups (TIWG) will provide valuable insight into schedule variability, the status of system development, and the readiness for operational test activity. The Program Manager, the OTA, or the Service Staff's Program Element Monitor are sources of updated activity schedules. Informal dialogue with members of the test team will facilitate an ongoing cognizance of the program status, at times prior to notification from the Service staff. Schedule awareness should include significant pretest activity, e.g., Test Readiness Reviews, in addition to actual test conduct. DOT&E staff members can leverage their effectiveness by utilizing all sources of information including the operational test agencies, program offices, and using commands. These groups assist in establishing test objectives, identifying needed resources, providing forums for discussion and resolution of test and integration subjects/problems, and related activities. During attendance at meetings, there is the opportunity to gain better insight into underlying issues and the associated potential impacts to the operational suitability of the weapon system under test.

> Test schedules were changed due to the late arrival of the test article, time required to install and calibrate instrumentation, and time required to conduct pretests and training. After two weeks of test trials, the Data Authentication Group (DAG) recommended that early trials be rerun due to the poor data from those events. At this point, the test was 4-1/2 weeks behind the original schedule

b. Structured Planning for Each Test-Observation Trip.

Will the limited on-site time be used effectively by the test monitors?

When planning for a trip to observe test operations, an activity checklist is extremely useful. This checklist will assist in test observation activities and provide reminders of critical issues or risks identified in key source documents as well as those surfaced in prior OSD reviews and congressional interest items. This approach ensures a consistency of purpose and maximizes the benefit of the on-site activity. Field observations will include deviations from planned activities and/or discrepancies that could impact quality of test results.

> The Staff Assistant prepared a checklist of pertinent information prior to the scheduled visit to the test sites:
>
> | | |
> |---|---|
> | Configurations | Interagency Coordination |
> | Hardware | Lessons Learned |
> | Software | Analysis Procedures |
> | Critical Operational Issues | Management Plans |
> | Test Design/Methods | Policy Documents |
> | Data Management | Security Requirements |
> | Environmental Considerations | Simulation Certification |
> | Intelligence/Threat Information | Test Operations |

c. Preparation for Accurate On-site Observation Reporting.

Will the DOT&E observers recall and accurately report what actually happened during the monitored activity?

An accurate account of the monitored activity is crucial. Careful notes must be taken to assure that incorrect information or conflicting information is not reported. In some cases, it may be desirable to use a portable tape recorder/dictation machine as the test event progresses; this would be supplementary to voice-tapes (which may be recording radio transmissions, etc.) used by test participants. Particular attention should be given to recording the specifics of unusual events and those which caused deviation from planned scenarios. Test monitors should maintain a log, recording the conditions which existed during the events, or scenarios.

> The DOT&E Staff Assistant used a hand-held cassette recorder to record her observations during the testing. These tapes were reviewed each evening for major significant points to be included in the trip report. Upon return to the Pentagon, portions of the recordings were transcribed completely by the DOT&E administrative support staff. This method of recording test site observations proved to be more thorough than notes or personal memory.

4.2 GENERAL OBSERVATIONS AT THE SITE

The benefits of personal observations at the operational test site are many. The Staff Assistant (SA) may discover information that would not have been thought important enough to be included in the OT&E report, but that will be important in the deliberations on the meaning of the OT&E results. The SA must be sensitive to this type of information before going to the test site, and have a sense of priorities about what observations and conclusions are important.

AREA OF RISK

Unexpected items or unusual events may yield test results that differ from those expected when the test was planned.

Once the operational testing is initiated, unplanned situations will occur and the testing that was intended by the people who reviewed the plan may not be accomplished. This situation is due in part to the fact that operational testing can never be completely described in the written document. The readers of the test plan information all will bring different interpretations to what the written word means. Once the test begins, it is common for these differences to become known. When the DOT&E Staff Assistant visits the test site, it is essential to identify any significant difference between what was "thought to be" and "what actually is." This identification is not for the purpose of redirecting the testing, but to place critical test results within the proper context.

OUTLINE FOR REDUCING RISK

a. Monitors Must Watch for "Good Intentions" Deviations.

Are deviations from approved test documents occurring inadvertently or through overzealous efforts of test participants?

A part of the DOT&E responsibilities is to designate observers for the planning for, and conduct of, the testing portion of OT&E. Experience in participation in test reviews, readiness briefings, post-test debriefings or reviews also can be key elements of DOT&E monitoring activity. Prior experience can reveal that deviations have been highlighted during internal DoD reviews, as well as by agencies such as the General Accounting Office (GAO). DOT&E representation at test-related activities can identify impending or actual variances so that appropriate corrections can be completed thereby ensuring credibility of the test programs. During the on-site visit, the SA should search out deviations from the test plan caused by instrumentation, personnel problems, or schedules changes that may effect the validity of the test results.

> During an operational test, the Engagement Line-of-Site System became inoperative and estimates for making repairs in a reasonable time were not good Due to this failure, the Test Director elected not to run additional trial events of the system's ability to detect and track the threat system, even though the system's ability to detect and engage systems was a COI that was to be addressed during the test The Test Director stated that data would be obtained from models and simulations to answer these issues

b. Resolution of Uncertainty and Trip Report Preparation.

Are there "observations" that the Service Test Director will not agree with from an accuracy viewpoint?

Observations may be reviewed with the Service Test Director to assure that the test monitor has a complete understanding of the actual circumstances and events. However, test monitors will not allow such review to impact reporting of their observations or their consideration in subsequent DOT&E analyses. Upon completion of the monitoring of a specific test event, the Staff Assistant (SA) should coordinate with others who monitored other aspects of the same testing, and all significant observations should be consolidated. Prior to departing the test location, the SA should review his major observations with the Service Test Director, and any resulting conflicting conclusions should be highlighted in the SA's trip report. A database of significant findings from the monitored activities should be compiled by the SA and maintained for use in development of B-LRIP reports, annual reports, etc.

> From interviews at the test site, the SA determined that the mission pilots were flying practice "safety" sorties over the planned flight profiles on the day prior to the OT test events. These practice flights would have skewed the test data When the observation was discussed with the Test Director, it was discovered that while the "safety" flights were being made, the flights did not involve the pilots assigned to the OT test events.

c. Data Authentication and Validation.

Are data being validated and authenticated by an independent group?

When required by the test complexity or data volume, a Data Authentication Group (DAG) should be formed to perform data evaluation and certification of engineering analysis. Normally, the DAG is independent of the data management and quality control process and is not under the supervision of the data manager; it provides a level of quality assurance above that to be expected from the data management quality control function. Each DAG should be tailored to the unique requirements of the test. There should be a procedure that includes the DAG's concept of operations, organization, responsibilities, validation procedures, reports to be produced, and required schedules.

> There are critical windows in the data collection effort. Should the windows be passed without appropriate actions being taken by the test team, valuable data are lost forever. During the test of an aircraft, it was determined that much of the data collected from the MIL-STD-1553 data bus were of very limited value because the on-board recorders and playback systems had not been developed sufficiently to allow access to the event data. If the Test Director/chief of instrumentation/chief of data management had the playback machines developed and tested in sufficient time to be used to evaluate the data, valuable information could have been gained and test resources preserved.

4.3 TEST SITE MANAGEMENT

The proper management and control of the test site and the test site activities are essential to providing the discipline necessary for an effective operational test program.

AREA OF RISK

Poor management procedures or procedures that are not followed can lead to invalid test results.

Converting the OT&E plan into specific actions requires test site management and operating procedures. The Staff Assistant should assure that planning has been performed for adequate management of the test site and that procedures have been issued that formalize the planning.

OUTLINE FOR REDUCING RISK

a. Program Documents.

Have the pertinent program test documents been analyzed?

A prerequisite to effective monitoring of system testing, simulation activity, demonstrations, or other events is an in-depth understanding of existing guidance related to the event (derived from the TEMP, Operational Test Plan, and other documents). Pretest analyses assist in clarifying the important test parameters, developing linkages to tie field test data with modeling/simulation data, and in establishing the analytical structure to perform final analyses. An early item would be verifying OSD approval of the TEMP as well as the DOT&E approval of the operational testing of the system.

> Prior to leaving for the test site, the SA made copies of the pages from the unclassified test plan that describe the test events that were planned. The TEMP for the system had been approved six months earlier and agreed with the OT&E test plan that was reviewed three months earlier. Major suitability deficiencies from the previous test phase were summarized as low percentage of fault detection by the diagnostics subsystem, the immaturity of the maintenance software, and very high mean times to repair. The corrections to these deficiencies were to be verified during the test events to be observed.

b. Instrumentation Plans.

Have instrumentation plans been analyzed to determine their effect on the tactical scenario?

A prerequisite to effective monitoring of a system test is an understanding of the instrumentation that will be used to collect the data required to evaluate the system. If not carefully planned and managed, instrumentation will tend to "drive' the test. After review of the instrumentation plan, the SA should be satisfied with the proper balance between test realism and the requirements for instrumentation.

During a helicopter OT&E, there was a requirement for instrumentation to be calibrated before each test event. The calibration was accomplished by flying the helicopter close to the instruments within minutes of the start of each trial. This requirement provided a tactical advantage to the tested system's crew because they had an opportunity to reconnoiter the battlefield and it reduced the tactical realism of the test. As it turned out, the instrumentation that required this calibration was not required for the test and the events were evaluated without the use of these data.

c. Data Management Plans.

Have data management plans been analyzed to determine if the plans conflict with test realism?

Reliable operational testing requires test realism. Data collection may tend to conflict with test realism, therefore there must be a compromise which maximizes the combined issue of realistic testing and thorough data collection. The SA should ensure that data collection and instrumentation is not causing the test to be unrealistic. The goal of instrumented data collection should be clearly specified in the data management plan. Although a realistic percentage of the total data generated must be developed to ensure proper analysis can take place, as a minimum the plan should identify the critical items for collection, to include a minimum quantity of each. In some cases the test event may be a singular event, in which case it becomes a critical item to collect these data and redundancy should be planned to ensure adequate collection.

During the operational test of an aircraft, the data collection plan stated the requirement for an on-board tape recording system to capture MIL STD 1553 Databus information. The recorders required the crew to land aircraft every 25 minutes to change the tape in the recorder. The test trial scenarios were developed around this data collection requirement, which made it very difficult to execute the test trials and collect meaningful data from other data points. Test realism also was limited because of this data collection requirement.

d. Data Authentication Group.

Is there a group of professionals established to validate data?

The Test Director or Program Manager often will establish a Data Authentication Group (DAG) to verify and validate test data, in addition to assisting in data reduction, quality control, and the identifying anomalies in the system, instrumentation, and test data. The DAG must be independent of the system developer and data manager and should report directly to the Test Director. It is important that a standard operating procedure (SOP) be written for the DAG and that this SOP be closely followed.

During the second week of the vehicle test, the Data Authentication Group announced that the night navigation trials data that already had been collected were sufficient to provide a "high level of confidence" evaluation of the system's performance. The Test Director elected to terminate the scheduled additional night navigation trials and use those resources in another area where confidence levels in the data were not as high.

4.4 COMPARISON TO THE TEST PLAN

The actual test will differ from the planned test due to a host of reasons. As with most plans, their usefulness ends with execution. However, test planning documentation has an added value to the test community in that it serves as a check sheet for the operational tester during execution phase. The SA should determine the degree of variation from the plans that the tester has been required to take in order to conduct the test. These unplanned variations often are weak links in the process and should be evaluated carefully.

AREA OF RISK

Actual test is different from planned activities.

During the actual test conduct, there may be room for on-the-scene decisions that result in the test being significantly different from that envisioned when the test plan was written and approved. These undefined areas need to be examined during test site visits and an understanding gained of how the items will effect the resulting test data and evaluation.

OUTLINE FOR REDUCING RISK

a. System Configuration.

Is the system configuration to be tested the same as that identified in the test plan?

In some situations, the system configuration planned for use during OT either is not available or it is delivered to the test site with some unexpected deviation. In limited cases, there is a deviation in the number of test articles. The limited assets available early in the development program may result in a different number being available at the test site from the number that was planned. The Staff Assistant should determine what the system configuration deviations are and evaluate their impact on test results.

> The system configuration of the three prototype helicopters provided for the OT were different from one another and significantly different from the planned fielded system. Maintenance personnel training was difficult due to these differences. (More than 1600 modifications were made to the prototypes before a fielded system was produced.) Significant contractor support was required to maintain the systems during the test. Many prototype subsystems were of such early configuration that their performance was difficult to measure.

b. Test Limitations.

Will test limitations significantly impact the test results?

The SA can review all listed test limitations to ensure that they do not impact on the ability of the test to meet the stated objectives. The test limitations should be clearly explained and it should be determined if the limitations can be avoided or resolved by the responsible organization.

The testing of a combat earth mover was limited by the types of soil and weather conditions to which the system would be exposed at the selected OT location. After evaluating the system's performance at the test site (clay/sandy soil), it was determined that it would not be possible to estimate the system's earth-moving capabilities in rocky and other types of soil. Additional OT at other test sites was required to provide the necessary evaluation report.

c. Unusual Pretest Procedures.

Are the test articles or support equipment subjected to unusual pretest maintenance?

During most OT periods, there are test events spaced throughout the testing period with periods of minimal activity in between. During the on-site observation, the range of activities that occur during the non-test periods should be assessed. Are the test systems subjected to unusual pretest maintenance? Are the test articles to be included in the testing selected from a "pool" of available assets? If so, is the selection representative of what the operational commander would do, or is the selection likely to skew the resulting test data? While there may be motivation to maximize the use of scarce test range time, the test team must guard against the use of unusual pretest maintenance that skews the reliability or maintainability data.

During the operational test and evaluation of a mortar system, it was determined that the ammunition casings had been machined rather than cast, as is normally the case. By machining the casings, the weapon proved to be very accurate although the ammunition was very expensive. The mortar is an area weapon and the normal method of manufacturing ammunition is by casting. When cast ammunition was obtained and fired, the mortar did not meet accuracy or dispersion pattern requirements at extended ranges.

d. Maintenance Activities Reflect Operational Concepts.

Is troubleshooting or system repair activity performed by representative personnel?

In the case of a major defense acquisition program, no person employed by the contractor for the system being tested may be involved in the conduct of the (initial) OT&E that will satisfy the B-LRIP reporting requirements of Public Law. In addition to pretest maintenance activities, post-test involvement of such contractors could be a cause for concern (an example being if contractor personnel became overzealous in "Scoring Conferences" by defending the cause/chargeability of a system anomaly).

During the operational pretest, the system's contractor monitored the missile system's failures and made corrections on spare computer circuit boards. Before the start of the OT, the contractor inserted the modified boards into the test articles. The substitution of modified circuit boards disabled all on-board test instrumentation and caused four missile shots to be made without data being collected. The uncoordinated modification of the test articles during pretest caused extensive delay and cost to the test program.

4.5 DOCUMENTATION

The documentation at the test site should include the operator and maintenance instructions and various supporting manuals, as well as manuals relating to the system software. The role of the documentation is to provide for system operation and maintenance that is consistent and that can serve as a basis for evaluation of the system.

AREA OF RISK

System documentation is not evaluated, nor does it provide the foundation for realistic system operation or maintenance.

When the operational testers are not using the intended system documentation for operation or maintenance procedures, then there is little assurance that the data from such a test are representative of what will occur in operational use. The system documentation should provide a foundation of system operation that renders the test data usable in operational evaluation. On-site observation of testing should include an effort to assure that the documentation is being properly used by the test site personnel.

OUTLINE FOR REDUCING RISK

a. Source and Status of Documentation.

Is system operation and maintenance documentation in a mature state to allow use during testing?

System operation and maintenance documentation normally is developed either by the system's developing contractor or another contractor. If development is not begun early, the documentation usually is in draft stage at the beginning of OT&E. Often document developers will attempt to modify the documents during pretests and tester training. These modifications usually result in significant changes in the system's operations or maintenance. The SA should review the system's operations and maintenance documents and determine their state of readiness for the OT&E; this includes the assessment of the effect of early drafts and/or poorly presented and partial documents on the ability of the crew and maintenance personnel to perform during the test.

> During the training phase for a weapon system test, it was determined that operation and maintenance documents were in an early state of development. The Test Director requested the system's contractor to extend the crew and maintenance training to compensate for this lack of documentation; the draft documents were used in training and during the operational test. During the test, many crew/maintenance errors could not be classified as to whether they were caused by improper training, poor reference documentation, or personnel errors.

b. Use of Maintenance and Operating Documents.

Are test players using the operational and maintenance documents to assist in the performance of their duties?

A good indication of the level of completion and the confidence that the troops have in the documentation is their decision to use the documents during the test. When it is determined that the documents are not being used, it may be desirable to ask the test player's opinion about the documents. When the documents are available, the SA should determine if they are presented as being complete, and if they are considered accurate by the test personnel.

> Just prior to the operational test, it was determined that the system's maintenance documentation was in an early draft stage. The Training/Logistics Command personnel were planning to use the OT&E to evaluate the maintenance documentation and make the necessary changes before going back to the contractor with final revisions prior to publication. The personnel had developed sampling techniques such as questionnaires, survey forms, and observation techniques to be used during the OT&E. The Staff Assistants should determine if other segments of the Service (i e , training, logistics, doctrine, or tactics) are using the OT to gather information and if their methods might adversely affect the test realism, data collection/evaluation, or test operations. He also should determine the status of the system's manuals prior to the start of OT.

4.6 TEST PERSONNEL

The observation of the operational test gives the Staff Assistant a sense of the skill level and qualifications of the personnel. Direct contact with members of the units or test team also will provide information about how representative of typical user troops the operations and maintenance personnel are.

AREA OF RISK

The use of unrepresentative personnel may result in invalid data.

If the personnel who are operating the system during OT are more highly skilled than will be the planned operational troops, the system will perform better than might be expected in the operational units. Likewise, if the maintenance personnel are unusually highly skilled, the suitability of the system may be optimistically reported. In the visit to the test site, the Staff Assistant needs to ascertain if the skill levels involved will effect the test data.

OUTLINE FOR REDUCING RISK

a. Tester Training.

Are test personal properly trained?

Training plans and certification plans for test personnel should be developed and published early in the Full-Scale Development Phase. Errors by test personnel usually are expensive and often cloud the reason for test failure. The SA should identify the training performed prior to the start of test and determine if additional training was given to test personnel as a result of the pretest events.

> During the pretest trials for the OT of a helicopter, it was determined that the test system crews had received additional training on aircraft subsystems, whereas the crews for the baseline system had not received this additional training. The subsystems involved basically were the same; therefore it was decided that the baseline crews should receive the same blocks of instruction as did the test system's crews. When comparing a new system with a baseline system, training becomes very important and must be carefully monitored.

b. Use of Contractors.

Are the system's developing contractors participating in the OT&E?

Personnel from the developing contractor are prohibited from participating in the conduct of operational testing. During very early test phases, the contractor may have a role in some of the suitability areas, e.g., performing second or third level maintenance, or assisting in maintenance that is specific to the developmental systems. If there is some level of contractor involvement, the Staff Assistant must examine the way in which this involvement might be influencing the OT data results. Are the Service maintenance personnel "consulting" with the contractor during routine maintenance?

During the OT of an aircraft, the contractor and subcontractors were allowed to move into a motor pool complex near the test site. It was very difficult to keep the contractors away from the test systems. On one occasion, a contractor was observed pushing a military maintenance person out of the way so that he could make adjustments on the aircraft engine.

c. Use of "Gold Crews."

Are the crews involved in the OT typical of those who would be expected to operate and maintain the system once fielded?

The term "Gold Crews" is used to identify a situation where the personnel who are operating or maintaining the system are of a higher skill level than the crews who are expected to operate or maintain the system once it is fielded. Is the skill level being used of such a degree that it skews the OT data and the evaluation results? In other cases, the skill levels of maintenance personnel in the field units are primarily mid-level. In this case, the maintenance personnel in the test should not (all) be "highly" skilled.

An aircraft test team was formed with maintainers fresh from the operating units of predecessor systems. While at the test site, the SA determined that while these troops were representative of the operating units, over 75 percent of them were assigned the 7-level skill code. In the operating units, only 25-30 percent of the personnel are 7-levels. The evaluation of the OT maintenance data must consider that some of the maintenance actions that were performed by 7-level personnel would be performed by 5-levels in the operating units.

d. Stress Level of Personnel During Test.

Will the test crews be placed under the operational stress that would be expected under combat conditions?

The stress levels that are present during the operational use of combat systems cannot be duplicated during operational testing. While this statement is true, it should not be taken as a rationale for avoiding the replication of realism--enemy fire can be simulated and unexpected events can be entered into "free-play" portions of tests. The human factors evaluation of stress levels in combat requires that consideration be given to all of the items that are unknowns during the actual combat use of the system.

The troops used in the OT of a shoulder-fired missile were being evaluated for human factors, portability of the missile, firing effects, ability to guide the missile, etc. The crews selected for this portion of the evaluation should not have more experience firing the predecessor system than will the typical user troops. If the user troops only have experience with missile simulators, then the crews used in the OT should only have an experience level that is representative. Stress can be introduced into the OT by having the user troops operate for long periods, similar to the projected operational use, and at the projected pace, or tempo.

4.7 TEST DATA

The observation of the operational test site provides the Staff Assistant with an understanding of the test organization's approach to data collection, reduction, and reporting. The data collection plan will include the methods for collecting, marking, handling, storage, and disposition of test data, as well as the training required for the data collectors.

AREA OF RISK

Improper data management may result in valuable test resources being lost or invalid conclusions being reached.

A poorly written or executed data collection plan may result in the wrong data being collected, the data not being clearly marked, the loss of data due to improper storage and distribution, or the loss of data during the reduction process.

OUTLINE FOR REDUCING RISK

a. Data Collection Concept.

Is the data collection concept for the systems to be tested rational and executable?

The data collection concept should provide pertinent information, including a list of key personnel, resources, and procedures. The concept should indicate the degree to which instrumented and manual data collection techniques will be used, with a discussion of why each technique is used. The concept should discuss the questionnaires and structured interview forms that will be used to support the qualitative measures. There should also be a discussion of how and when opinions, interviews, and observations of the players, controller, data collector, and test directorate personnel will be gathered.

> The data collection concept for a missile did not specify the forms or formats to be used for the collection of test player comments. During the field assembly of the missile, it was evident that troops could not assemble the missile with the equipment provided. The Test Director halted the test and had the troops interviewed to clearly identify the problem. After the interviews were conducted, it was determined that various interviewers had placed undue emphasis on specific aspects of the problem and had ignored other areas.

b. Data Collection and Processing System.

Is there a collection and processing system in place and do the test personnel understand what they must do to use the system?

The collection and processing system should describe how the data are to be organized, reduced, verified, managed, controlled, and stored. The test plan for the tested system will list each data requirement and its means of collection. The SA should consult the test plan to determine if collection and processing systems are arranged according to plans and if there are data requirements that will not be collected because of the process. The SA also may determine whether the data

are properly organized and stored, according to collection source. The data flow diagram (in the data collection plan) is a good source for information.

> During the test of a armored vehicle, the test team did not correctly execute the data collection plan This resulted in a disorganized data collection activity with piles of data forms, questionnaires, and other documents not being properly recorded, filed, and stored. By the third week of testing, the Chief of Data Collection was unable to determine, from the documents, if trials were conducted at day or night, what the specific weather conditions were, or the degree of slope that the vehicle was operating on during the trial. As a result of this improper data management, an additional week of testing was required to fill in blanks in the data matrix.

c. Quality Control of Test Data.

Are the test data being independently validated in a timely manner?

There is a need to quickly validate the test data to determine if additional test trials are required to answer operational test issues. The SA should determine if the data collections are adequate to support the OT&E report. He should ensure that there are checks and procedures in place at the test site to preclude or detect and correct errors made in data collection, data entry, and data reduction. The procedure should also outline the process for making required corrections or changes in test data and how the audit trail for those corrections will be maintained.

> During the pretest, it was determined that data collected procedures were not validated; it was debated whether the test should begin before corrections were made. The Test Director was unsure if critical data were being accurately collected during the pretest. However, because of the resources involved, it was decided to start the test and complete validation during the early stages. Critical data were lost from the trials performed in the early days of the test. The test was stopped after one week to fully install the data collection procedures.

d. Data Integrity.

What are the established procedures for ensuring the maintenance of data integrity?

The data collection plan will discuss the process through which each set of collected data is to pass before reaching the storage medium which supports the Test Director and the evaluator. Data flow diagrams identify where data are combined with other data, and where they are processed, scored, reorganized, validated, or otherwise manipulated. The data collection plan should describe the data manipulation process at each step, along with rules and procedures for manipulation. The SA may wish to review the data collection plans and to include the flow diagrams to ensure that careful consideration has been given to data integrity.

> Control of the data was lost during the transfer from the test site, through a support contractor, to the test agency. It was determined that an audit trail was not established during the planning for data collection, transfer, and storage. This situation allowed data to be reorganized and manipulated without consideration of the effec s of these activities on their evaluation. As a result, some data could not be used because their authenticity could not be assured.

4.8 TEST SCENARIOS

While visiting the test site, the Staff Assistant should examine and record the details of the test scenario that is being used. The comparison to the scenario that was described in the test plan is an important part of the on-site assessment. The test scenarios should provide realism, as well as an opportunity for test personnel to collect data on the system's effectiveness and suitability. Operational test scenarios normally are well planned, but due to location, instrumentation, and data collection requirements, they may not be well executed.

AREA OF RISK

Operational control often is lost and trials become force-on-force tactical maneuvers.

The Staff Assistant may determine if operational control of the tactical forces is maintained to ensure that test objectives are being met. Scenarios that are designed to support the evaluator's requirements must be closely followed to ensure that data collection is accomplished. Tester control, in a force-on-force operational test, is not always easy because tactical units tend to perform in direct proportion to their level of training and motivation. During initial trials, troops will be highly motivated and tend toward being uncontrollable in their zeal to "win the war." However, as the test wears on, test personnel will become bored. Therefore, test scenarios have to be carefully monitored to ensure that troops react with enough "gusto" to provide realism.

OUTLINE FOR REDUCING RISK

a. Operational Uncertainty.

Does the test plan dictate that a level of operational uncertainty be maintained during the test?

To conduct realistic operational tests on some systems requires that actions by the threat systems be uncertain to the player personnel. This can mean uncertainty of time, location, type of action, direction, etc. The SA needs to assess the "reasonableness" of the threat scenarios.

> The enemy attack forces were scheduled to initiate a new test event each day after the forces reached the exercise area. As a result, the players had a rough idea of when the attack would begin. After three days, this scheduling became obvious and the Test Director took control of initiating the attack forces.

b. Approved Doctrine and Tactics Used During the Test.

Are the tactics and doctrine used by the test system and the threat systems approved by the responsible agency?

There usually is considerable disagreement about how the enemy will fight a new weapon system. For example, in a high-intensity conflict, will the enemy attack helicopter be deployed nap-of-the-earth, low-level, or at altitude? The doctrine and tactics for engagements of these helicopters will depend on an assessment of the most probable methods of deployment. The SA

should determine if the Service has clearly established doctrine and tactics, based on the approved STAR, for the deployment of the tested system. He should determine if the test scenarios follow the approved doctrine and tactics.

> During the testing of an attack aircraft, threat crew members decided to use tactics that did not conform to approved doctrine. These tactics caused wide divergence in the planned scenarios and several days of testing were disrupted while the crews were properly briefed. Proper tactics and doctrine were included in the test events thereafter.

c. Tactical Operations Center Activites.

Is there a tactical operational center located in the field which is operated by trained tactical operations personnel?

To maintain control of tactical forces during a force-on-force operational test, it is important to have a fully manned tactical operations center located in the field with the operating forces. This center should be manned by fully trained operations personnel who are detailed from a tactical unit. The forces must be carefully briefed on the scenarios and the objectives of the operational test and evaluation. The SA should ensure that these personnel understand that good control of the maneuvering forces must be maintained in order for the test personnel to collect the data required to evaluate the system being tested.

> Prior to an OT event that included an armor engagement, the tanks being tested where driven for a number of hours to simulate a cross-country movement; as a result, the systems in the tanks were in a condition similar to what they might reflect in actual use. In this situation, to go directly from maintenance facilities to the engagement areas was not realistic.

d. Testing Terrain and Climatic Conditions.

Are environmental factors for the OT realistic for the system's mission profile?

Because of the availability of operational test sites, it is very difficult to find the desired terrain and climatic conditions in order for OT to be conducted under conditions that are fully representative of the system's intended deployed operational mission profile. Also, tradeoffs will have to be made because of safety and operational constraints. The SA should determine if the site selected for the test is representative of the terrain and climatic conditions required to provide realistic testing.

> During the Phase X OT of an Army vehicle system, heavy rains inundated a number of test site areas. A test event was added to demonstrate the vehicle recovery capability and the associated training, equipment, and procedures.

THIS PAGE
INTENTIONALLY
LEFT BLANK

Chapter 5
OT&E REPORT

The OT&E report can be viewed as the most important document in the OT&E process. While it is true that the TEMP provides the foundation for all of the operational test and evaluation and the test plan provides the details for the information in the TEMP, the OT&E report is the product of all the test planning, management, and actual conduct of the T&E. The OT&E report is the document that provides the T&E results to the decisionmakers. It further serves to complete the process and realize the purpose of operational test and evaluation, and to provide information upon which to base decisions. Without clear and complete reporting in the test report, test planning and test conduct may have little impact on the acquisition program decision process.

There are a number of different OT&E reports; quick look reports, informal reports, briefing reports, and formal reports for different acquisition phases. Each of these OT&E reports must be considered for its individual purpose, and the phase the acquisition program is in. The report should show the audit trail through the program milestones and present the program's progress. It should summarize the planned testing, what happened during the test, and the test results.

In reviewing the OT&E report, the DOT&E Staff Assistant has two objectives. First, he should ensure that the report accurately reflects the test and evaluation that took place, and that it contains current, complete, and accurate data. Second, he should review the test and evaluation results on the system, including the issues and deficiencies, and recommend a DOT&E position to be reported to the appropriate decisionmaking forum.

There is no standard Department of Defense format for an OT&E report; each Service has developed its own approach to documenting the required information. Within the Army, there are two separate reports for OT&E activity. The test report (TR) is the primary record of the operational test and the findings and facts that resulted from the testing. The independent evaluation report (IER) documents the evaluation of the test data. These two documents together provide the composite information that is provided in the OT&E reports of the other Services.

From an operational suitability standpoint, the operational test and evaluation report is the concluding document in the process of determining if a system is suitable for operational use. The report should highlight the background of the test, i.e., the set of circumstances and decisions that determined what testing was conducted. Most operational testing has limitations, so it is important that the limitations that affect the suitability results are identified and documented in the OT&E report. The framework for reporting the testing results is the evaluation criteria that were developed prior to the actual testing. These criteria include the suitability critical issues and the associated characteristics, parameters, and thresholds. The test that was conducted should be summarized and any unexpected changes from the testing that was planned should be identified and documented. Finally, the report should discuss the test data and results, as well as provide an evaluation of the system and conclusions about the system's operational suitability.

Table 5-1 OT&E Report Formats

OPERATIONAL SUITABILITY GUIDE	ARMY TR (DAP 71-3)	ARMY IER (DAP 71-3)	AIR FORCE (AFOTECR 55-1)	NAVY (OTD GUIDE)	MARINE CORPS (MCOTEA)
5.1 Background Issues 5.2 Limitations 5.3 Evaluation Criteria	1.0 Introduction	1.0 General	I. Purpose and Background	1. Purpose 2. Equipment Description 3. Background 4. Scope 7. Critical Operational Issues	1. Introduction 2. System Description 3. Objectives and Issues
5.4 Summary of Test Conduct 5.5 Test Results	2.0 Test Results		II. OT&E Description III. Operational Effectiveness and Suitability	5. Project Operations 6. Results	4. Evaluation
5.6 Conclusions	* Conclusions Are Provided In Some Reports	2.0 Findings of Evaluation	V. Summary of Conclusions and Recommendations	9. Conclusions 10. Recommendations	
	Appendices	Appendices	IV. Service Reports	8. Operational Considerations	

The templates that follow address each of the major content areas included within OT&E reports. These areas are:

5.1 Background Issues
5.2 Limitations
5.3 Evaluation Criteria
5.4 Summary of Test Conduct
5.5 Test Results
5.6 Conclusions

Table 5-1 was prepared from the Services' policies on OT&E. It shows how the structure of this operational suitability guidance document relates to the major sections of the respective Services' OT&E reports.

5.1 BACKGROUND ISSUES

The background issues include those major events and information that are needed to place the OT&E results into proper perspective. Items in this category include program direction on the conduct of the OT&E, and the results and limitations of earlier operational testing as well as significant operational issues.

AREA OF RISK

Unclear statements of issues may result in an incomplete or biased view of what the test results mean.

If the background is not summarized adequately, the testing and T&E results may not be placed in the proper perspective. The T&E activity on a program is a continuum of activity that includes periodic reports. To gain maximum benefit from these reports, the context of the testing, the meaning of the evaluation, and full consideration of the results is essential. If the purpose of the testing and the background are not described adequately, the reader may formulate an incomplete or biased view of what the test results mean.

OUTLINE FOR REDUCING RISK

a. Previous Test Phases.

Are the important operational suitability aspects of the previous test phases summarized?

The previous test phases provide the fundamental starting point for the test that is being reported on. The summary of these earlier test periods should include the important areas that were satisfactorily tested and also summarize the significant results. It should include which operational suitability areas were tested well and had good results, which areas were not included in the test, and which areas were tested but did not meet their requirements.

> The previous system tests were summarized as they apply to the major RAM characteristics. The reliability chart showed the stated quantitative requirement (in mean time between operational mission failure - MTBOMF) for each of the three configurations. The test results for each of the completed OT phase were listed. For reliability, the point estimate value and the value at the 80 percent lower confidence limit were shown in the chart.

b. Previous Management Decisions.

Did previous management decisions drive the operational suitability test planning?

Are the decisions that directed the test and the issues involved in those decisions described in enough detail to permit the reader to understand the context of the decision? Management direction that resulted from previous testing may highlight critical or risk areas; understanding these concerns can assist in placing the operational suitability results into proper perspective.

> Management review of the previous testing results focused attention on the progress in the diagnostics area. The maturity level of the built-in test during the previous test phase was less that expected. This area was highlighted by the Acquisition Decision Memorandum as an area of risk that should received additional attention during the next OT&E phase.

c. Operational Suitability Characteristics.

Are suitability characteristics included in the system summary?

The system summary within the test report should be a relatively short discussion of what was tested, emphasizing the mission or function of the system. Any system attributes, i.e., new technology or capability, high risk areas, etc., that are the reasons for critical operational issues (COIs) should definitely be included. The support concept also should be summarized, since this is needed for assessing how complete the operational suitability portion of the OT&E was.

> The system summary includes an outline of the maintenance concept planned for the system. The summary outlines how the system will be supported at each level of maintenance and what maintenance will be performed at each level.

d. System Differences.

How was the tested system different than the planned operational system? What are the resulting implications for suitability testing?

The systems under test, particularly in the early stages of OT&E, may be significantly different from the planned operational system. These early systems generally will not be from the production line that is planned for the full rate production and therefore will not have the full benefit of stable production processes. The system differences should be discussed as they relate to limitations on the suitability testing and the need for any additional suitability testing in a later phase.

> The test report indicated that the operational software that was available during the test was not the version planned for the production systems. Because of this difference, there were some diagnostic and maintenance operational capabilities that were not verifiable during the operational suitability testing.

5.2 LIMITATIONS

All major tests have limitations. These may include limits in the operating environment, the length of the test period, or the type or number of test units, personnel, or supporting devices. The list of limitations highlights the potential risks involved in assuming that test results are totally indicative of what might be expected in the operating environment.

AREA OF RISK

Insufficient identification of limitations may result in inaccurate assessment of the OT&E results.

Limitations that occur during the actual testing may not be highlighted in the test report. Such limitations could result from the lack of realism of the test environment, the test duration, the number of test articles, or the availability of test hours. The test articles might not be of the latest configuration or capability. Without visibility of the limitations, suitability implications may be overlooked. The risk in this area is that limitations are not identified sufficiently for the decisionmaker to know how to assess the OT&E results.

OUTLINE FOR REDUCING RISK

a. Operational Suitability Test Limitations.

Was there a summary of all known test limitations that would affect suitability?

Operational suitability limitations may include many areas. Limited number of test hours can be a limitation on evaluating the level of reliability and maintainability. Others may be the result of the logistics support during the test period being not representative of the planned operational support. Test equipment may not be available. All operational suitability limitations should be summarized to permit the reader to place the test results into perspective.

> There was an inadequate number of test hours on selected host vehicles. A complete reliability evaluation of the system is dependent on sufficient test hours to provide statistical confidence in the results. Lack of test hours and a compressed test period will reduce the statistical confidence in the OT&E results.

b. Additional Limitations During Testing.

Were there any limitations that developed during the conduct of the actual testing? Are they presented and their impact on operational suitability discussed?

Additional limitations beyond those in the test plan may surface once the test has started. Specific test articles may not be available as planned. Other program priorities may dictate a revision to the activity that was planned in the OT&E test plan. The test equipment planned for the test site may not be available. Other items, such as technical documentation, may be delayed and not included in the operational suitability evaluation.

> The software for testing the radar system components on the XX-537 test station was not available at the test site as planned. As a result, the second-level maintenance for the radar was not included in the test.

c. Environmental Differences.

Were there significant differences between the test environment and the expected operational environment?

Test limitations could result for significant environmental differences. These differences could be major factors in assessing the acceptability of the operational suitability results. The test report should identify the major conditions that were different. This could include the ratio of support assets to test articles, the skill levels of the maintenance personnel, or the depth of training that the maintenance personnel received.

> The maintenance team that supported the system during the testing was generally of a higher skill level than the planned maintenance organization. Analysis of specific task times was performed to determine the effect of this skill difference and the resulting impact on the measured mean time to repair.

d. Suitability Elements Not in the Test.

What suitability elements were not able to be evaluated during the test?

In most OT&E phases it is common for some support elements to not be present during the testing. An example is second-level test equipment that is not yet developed to the stage where it is ready for the OT&E environment. Other elements may have items substituted that are significantly different, such as factory test equipment used instead of the planned operational test equipment. The effect of these situations must be considered to identify the areas of suitability that are not yet evaluated. These risk areas should be considered by the decisionmakers and included in future phases of test and evaluation.

> For the test period being reported, the second level of maintenance was performed by the hardware contractor at his facility and therefore was not part of the evaluation.

5.3 EVALUATION CRITERIA

The evaluation criteria should be derived from the requirements of the user organizations. There also may be qualitative criteria for some of the operational suitability elements.

AREA OF RISK

If the reported results are not related to the user's stated requirements, then the test and evaluation report may not be meaningful.

The proper evaluation of suitability test or analysis results is dependent on the pre-established criteria. These criteria must be related to the user's stated requirements. The test and evaluation report may not be meaningful if the results are not related to the user's stated requirements.

OUTLINE FOR REDUCING RISK

a. Criteria.

Were there established criteria for each of the operational suitability characteristics?

To be meaningful, each operational suitability characteristic should have evaluation criteria drawn from the user's stated requirements and stated in the OT&E test plan. Was enough detail provided to define the criteria? For example, is there some indication of how failures are defined? Was a reference listed that included the failure definition?

> The reliability criterion is stated in terms of mean miles between unscheduled maintenance actions (MMBUMA). The system shall achieve 200 MMBUMA.

b. Source of the Operational Suitability Criteria.

What is the source of the quantitative and qualitative criteria for the operational suitability elements?

When the criteria in the operational suitability are stated in the test report, they give the reader a framework for evaluation of the test results. To provide a complete picture, the test report should identify what the sources were for the evaluation criteria. This is particularly important in the statement of any qualitative criteria. Most quantitative criteria can be traced to requirements documents, but this is not always true of the qualitative criteria. (OPTEVFOR does not use criteria on the qualitative measures.)

> The system reliability, maintainability, and availability requirements were specified in the July 1987 revision of the system Required Operational Capability (ROC).

c. Quantitative Reliability, Availability, and Maintainability (RAM) Measures.

Are the suitability characteristics for RAM stated in quantitative terms?

For most systems, operational suitability characteristics such as reliability, availability, and maintainability (RAM) can be expressed with quantitative parameters. The evaluation criteria for these areas therefore should be stated quantitatively and used in evaluating the test results. The definition of the appropriate terminology is an essential part of understanding these numerical values.

> The fixed-installation communication set was required to have a mean time between operational mission failure (MTBOMF) of greater than 500 hours.

d. Confidence Limits.

Are there statements of confidence limits for the RAM quantitative criteria?

The measurement of statistical parameters such as reliability can require considerable test time. If limited test time is available, as is often the case, then the measured value is expressed with a degree of statistical confidence. Background information on determining confidence levels from test data is discussed in DoD 3235.1-H, "Test and Evaluation of System Reliability, Availability, and Maintainability--A Primer." This document discusses the mathematics of test statistics, but does not aid in deciding what level of confidence is needed. (While OPTEVFOR uses confidence calculations in test design, these factors usually are not mentioned in their test reports. The other OTAs use confidence levels when they are considered appropriate.)

> The mission reliability of the system shall be measured in terms of the mean time between operational mission failures (MTBOMF). The point estimate MTBOMF that results from the scoring conference data shall be used to compute an 80 percent lower one-sided confidence interval. The 80 percent confidence value shall exceed the requirement stated in the approved ROC.

5.4 SUMMARY OF TEST CONDUCT

The OT&E report must summarize the actual suitability testing that was performed. This summary must include an adequate description of the operational suitability testing and present all significant changes from information that was presented in the OT&E plan.

AREA OF RISK

Omitting critical information may alter the meaning of the test results.

A change in the planned suitability test activity may result in significant changes to the meaning of the test results. The test report should present the description of what was actually done and what the results were. Some of the information on deficiencies may be omitted from the summary in an effort to condense the information or to protect the results from further scrutiny. The DOT&E Staff Assistant must have enough visibility into the actual testing to know if the summary is complete and accurately reflects the suitability T&E that was performed. The risk in the summary is that some critical information will not be included.

OUTLINE FOR REDUCING RISK

a. Comparison to Test Plan.

Was the operational suitability test conducted consistent with the suitability test planning?

The discussion should highlight any differences between the suitability test as planned and as conducted. Any differences most likely are at the detailed level, thus, some detailed discussion is required to point out the differences and any impact these differences may have on the test results.

When operational testing is performed, there often are changes required because of conditions that were not foreseen when the OT&E plan was prepared. These changes must be summarized in the OT&E report. They should be discussed in enough detail to permit the reader to assess what impact these items had on the test and the test results.

> During test phase II, the XYZ test set was not available as had been planned. As a result, the reliability of the test set, the compatibility of the test set with the system, and the testability of the system were not evaluated as had been planned during test phase II.

b. Unexpected Limitations.

Were there any unexpected limitations to the operational suitability portion of the test?

Unexpected test limitations may result from the differences discussed above, but they also may result from unexpected suitability-related factors, such as weather and support personnel.

> The results of the operational testing showed that the maintenance training that was conducted prior to the start of the test was inadequate in the area of radar fault isolation and troubleshooting. This deficiency precluded a complete evaluation of maintainability, maintenance documentation, and the diagnostics system for the radar.

c. Summary of the Test Performed.

Is the summary of the test that was performed complete and unambiguous?

Based upon "first-hand" knowledge of the test conduct, the Staff Assistant should assess the summary in the OT&E report. Different authors will describe suitability-related events differently. Identical summary descriptions are unlikely, but the objective is to have a discussion that includes a fair and accurate summary of the significant aspects of the suitability testing that was conducted. The facts concerning what was done and what occurred during the suitability testing should be clearly stated. The activities should be related to the test planning or other reference documents.

> The search radar system was flown on 43 sorties. The Type I mission profile was flown as described in the system OT&E test plan. The total flying hours on the system were 168 hours. The system experienced three mission failures during the test period. The point estimate MTBOMF was 56 hours.

5.5 TEST RESULTS

The summary of the test results is the major section of the OT&E test report. The information here should support the suitability conclusions reached and provide a basis for the readers to form judgments that agree with the major conclusions. The level of detail that is provided should permit the readers to integrate independent thoughts with the details of the suitability test results.

AREA OF RISK

Poor presentation of important suitability factors may lead to incorrect conclusions.

The test results must be discussed in enough detail to support any conclusions that the system meets or does not meet its suitability criteria for the phase being assessed. If adequate detail is not presented, the major conclusions may not be accepted and may not be supported by the reader. The major findings should be highlighted in a manner that gives the reader insight into the important suitability results that evolved from the test and the evaluation of the test data.

OUTLINE FOR REDUCING RISK

a. Presentation of Suitability Results.

Are the major suitability findings presented in an understandable way?

The major operational suitability findings should be in areas related to the suitability critical operational issues (COIs). There may be important operational suitability findings in other areas as well. The DOT&E Staff Assistant needs to review the findings and ensure that all important areas are addressed.

> Overall Evaluation: The operational suitability of the system is satisfactory. Reliability and maintainability exhibited in OT&E exceeded the stated requirements.

b. Projection Methods.

Was the method used to project RAM values from the test results to mature values examined and validated?

If the test report states values for the mature projection of reliability and maintainability (R&M), is the method of projection described in enough detail to evaluate its applicability? Some operational testing agencies never use projections, others use them often. There are several projection methods available for use. When they are used, the validity of the methods and applicability of the method to the situation at hand must be assessed. The use of a questionable projection method may cause the decisionmaker to place little value on the OT&E results. A projection should never be reported as a test result--the test result should be an observed value.

> The reliability test results were projected to the mature system values using methods from Military Handbook 189. The applicability of the method to the system was validated by using growth experience with similar systems, and by reviewing the reliability growth plans presented by the program manager.

c. Confidence Levels.

Are confidence levels stated for the quantitative R&M values that resulted from the tests?

The use of limited test data to provide measures of reliability and maintainability always includes room for statistical error in the quantitative estimate. The likelihood of the error is shown by the statistical confidence associated with the test measurement. For each appropriate measurement, a statistical confidence should be stated. (While OPTEVFOR uses confidence calculations in both test design and in evaluation, the confidence values are not included in OPTEVFOR test reports.)

> Reliability: An 80 percent lower one-sided confidence interval was calculated for the point estimate value of the MTBOMF. Operational mission failures were determined by a formal scoring conference in accordance with the approved failure definition and scoring criteria. The 80 percent confidence value exceeded the requirement stated in the approved ROC.

d. Impact of Suitability Results.

Is the impact or consequence of the major suitability findings stated?

When operational suitability results are compared to their criteria, the report should state that for this particular suitability area, the system is acceptable, marginal, unacceptable, or make some other judgment. The impact or consequence of these conditions on the system should be included in the test report. The DOT&E Staff Assistant should have an understanding of the situation and its impact on the system's operational capability.

> The stability of the ammunition trailer is marginal. The implication of this deficiency is that the trailer will be unable to carry a full load of ammunition and travel at normal speeds on an unpaved road without tipping over.

5.6 CONCLUSIONS

The conclusions that are reached as a result of the operational testing should be clearly stated in the OT&E report. In this section, there should be a clear statement of whether the system was considered operationally suitable or not.

AREA OF RISK

Poorly stated or omitted conclusions can result in erroneous management decisions.

Not all test reports contain conclusions on the acceptability of the system. Conclusions should be stated that indicate the major points that are drawn from the suitability test results and the evaluation of the data against test criteria. Was the test passed and what did the test prove? What areas are acceptable and which are not? What is the meaning of the deficient areas relative to the progress of the system toward its suitability for operational use?

OUTLINE FOR REDUCING RISK

a. Operational Suitability Conclusion.

Does the report state whether the system is considered operationally suitable or not? Is this operational suitability assessment a composite view of the suitability elements?

How does the report deal with the situation where some of the elements are satisfactory and some are not satisfactory? Are the unsatisfactory items highlighted? Those items that are not within the user's stated needs or that do not meet the OT&E criteria must be assessed for impact on the system's overall suitability for fielding. A single conclusion in the operational suitability area may be difficult. This requires the combination of the many suitability elements, some of which may be good, some marginal, and some not evaluated, into a single judgment. The suitability test results must be evaluated within the context of the planned operational use of the system and a judgment made as to whether the system meets the user's needs. Suitability COIs are used to focus attention on areas that should receive increased weighting in any total assessment.

> Although the system meets its operational mission performance and the RAM requirements, it is not suitable for fielding without major changes to the planned logistic support. The system is judged as not operationally suitable because of deficiencies in second-level support equipment and the quantity of the planned spares.

b. Operational Effects of Any Adverse Test Results.

What is the operational implication of the operational suitability deficiencies?

When an operational suitability deficiency is highlighted in the test report conclusions, the report should include the implication of this deficiency on the operational use of the system. How is the system limited in its use? What missions or uses are degraded by the deficiency? What additional resources are required to compensate for the existence of the deficiency?

> The automatic diagnostics system for the weapon system was deficient in the area of fault isolation. The evaluation criterion for fault isolation was:
>
> Isolate faults to one SRU -- 90 percent; two SRUs -- 95 percent.
>
> Faults were isolated to one SRU -- 72 percent; two SRUs -- 84 percent.
>
> Additional unit-level spares will be required to support the initial operating units until the diagnostics deficiencies are corrected.