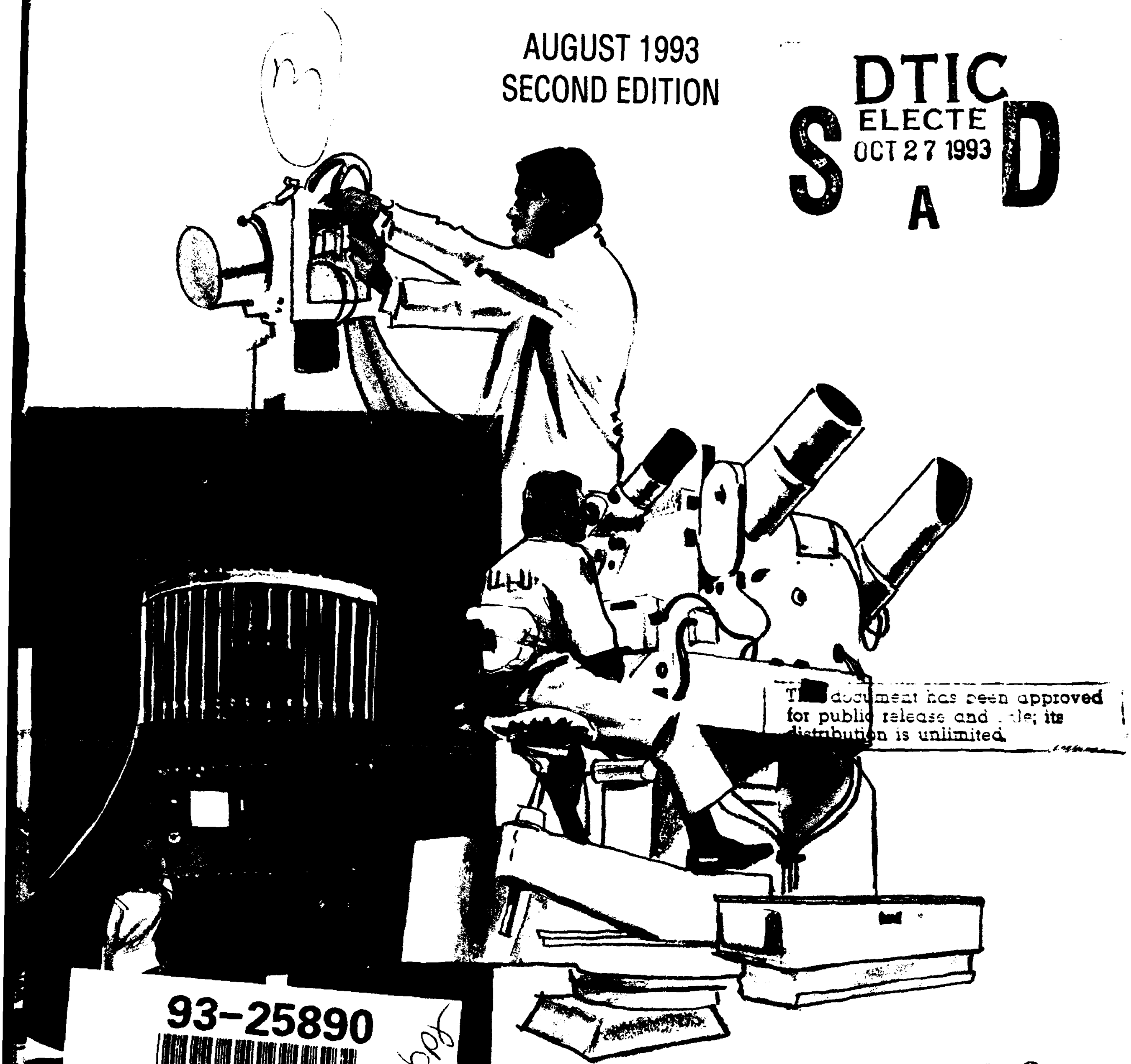
DEFENSE SYSTEMS MANAGEMENT COLLEGE

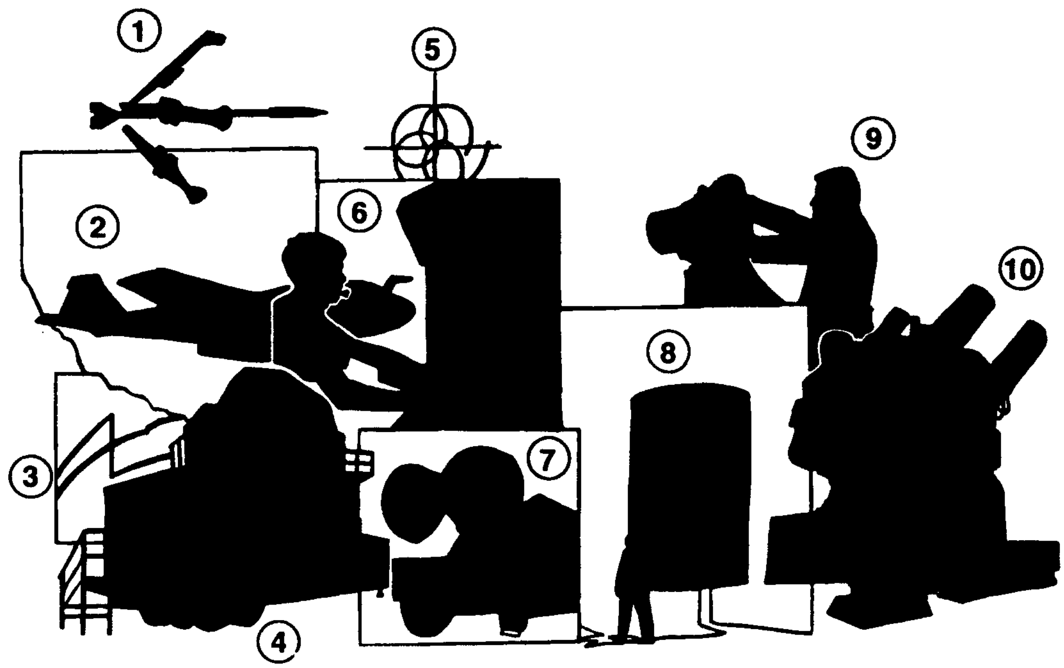
1

TEST AND EVALUATION MANAGEMENT GUIDE

AUGUST 1993
SECOND EDITION



24pp



1. Munitions testing in environmental chambers at CSTA
2. Hybrid laboratory simulation anechoic chamber (in-the-loop simulator and EW system test)
3. Data-reduction charts
4. AN/MPS-39 multiple-object tracking radar
5. Position attitude diagram
6. Testing of training device that simulates system operation under battlefield conditions
7. Mobile data telemetry receiver at CSTA
8. Acoustical vibration test (in a 5,000-cubic-foot reverberation chamber) on an MX missile stage
9. A 3,000-psig, flow interrupt, pressure impulse, and water-hammer test on a series of nuclear valves
10. EOTS-F Cinethrodolite with automatic infrared tracking system

TEST AND EVALUATION MANAGEMENT GUIDE

AUGUST 1993
SECOND EDITION

PUBLISHED BY THE
DEFENSE SYSTEMS MANAGEMENT COLLEGE PRESS
FORT BELVOIR, VA 22060-5426

For sale by the U.S. Government Printing Office
Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328
ISBN 0-16-041892-5

ACKNOWLEDGMENTS

I would like to acknowledge the valuable contributions, reviews, comments and technical assistance from LTC Charles K. Gailey, USA; Lt Col Dennis Van Liere, USAF; LTC Larry Damman, USA; and Mr. Paul Alfieri of the Test and Evaluation Department faculty. The cover and graphics were developed with the assistance of Mr. Greg Caruth and Mr. John Kerpchar; the editing was done by Mrs. Kay Sondheimer and Mrs. Esther Farria. Special thanks to Mrs. Peggy Claxton for her patient typing, retyping and editorial assistance.

FOREWORD

This book is one of a family of educational guides written from a Department of Defense perspective; i.e., non-Service peculiar. They are intended primarily for use in the courses at the Defense Systems Management College (DSMC) and secondarily as a desk reference for program and project management personnel. These guidebooks are written for current and potential Department of Defense (DOD) acquisition management personnel who are familiar with basic terms and definitions employed in program offices. They are designed to assist government and industry personnel in executing their management responsibilities relative to the acquisition and support of defense systems. They include:

a. Integrated Logistic Support (ILS) Guide (May 1986)
b. Mission Critical Computer Resources Management Guide
c. Systems Engineering Management Guide (January 1990)
d. Department of Defense Manufacturing Management Handbook for Program Managers (April 1989).

The Defense Systems Management College is the controlling agency for this book. Comments and recommendations relating to the text are solicited.

The introduction offers a perspective of test and evaluation (T&E) management activities during the system life cycle. Subsequent material in this book provides a guide for managing specific T&E events. The past several decades have seen the rise of large, highly-interactive defense systems that are often on the forward edge of technology. These systems have a natural evolutionary process, or life cycle, during which actions taken or avoided in early stages can mean the difference between success or failure.

The T&E process for DOD materiel acquisitions is a complex exercise of integrating the engineering development, data collection, and evaluations necessary to satisfy the decision-maker's information requirements on system performance. Poorly managed T&E does not support an informed decision process and can generate schedule slips or adverse media exposure, leading to intensive interest in a program's status by the Office of the Secretary of Defense (OSD) and/or the Congress.

The objective of a well-managed T&E program is to provide timely and accurate information. This guide has been developed to assist the acquisition community in obtaining a better understanding of who the decision-makers are and determining how and when to plan test and evaluation events.

John Claxton
Professor
Test and Evaluation Department

TABLE OF CONTENTS

Chapter 17 — Test and Evaluation Master Plan

Module V — Specialized Testing

Chapter 18 — Embedded Computer Systems Testing

Chapter 19 — Testing for Vulnerability and Lethality

Chapter 20 — Logistics Infrastructure T&E

Chapter 21 — EC/C3 Test and Evaluation

Chapter 22 — Multi-Service Tests

APPENDICES

LIST OF FIGURES

LIST OF TABLES

I

MODULE

Management of Test and Evaluation

Test and Evaluation is a management tool and an integral part of the development process. This module will address the policy structure and oversight mechanisms in place for test and evaluation.

1

IMPORTANCE OF TEST AND EVALUATION

1.1 INTRODUCTION

The test and evaluation (T&E) process is an integral part of the systems engineering process which identifies levels of performance and assists the developer in correcting deficiencies. It is also becoming a significant element in the decision-making process, providing data supportive of trade-off analysis, risk reduction and requirements refinement. Programmatic decisions on system performance maturity and readiness to advance to the next phase of development are becoming more dependent on demonstrated performance. The ultimate customer, the Service-member user, is concerned about neither unit cost nor production schedule. The issue of paramount importance is system performance; i.e., will it fulfill the mission. The test and evaluation process provides data to tell the user how well the system is performing during development and if it is ready for fielding. The program manager must balance the risks of cost, schedule and performance to keep the program on track to production and fielding. The responsibility of decision-making authorities centers on assessing risk trade-offs. This chapter describes how test and evaluation functions as a risk management tool. It also addresses the contribution T&E makes by providing empirical data before each milestone review.

1.2 TESTING AS A RISK MANAGEMENT TOOL

Correcting defects in weapons has been estimated to add from 10-30 percent to the cost of each item (Reference 107). Such costly redesign and modification efforts can be reduced if carefully planned and executed test and evaluation programs are used to detect and fix system deficiencies sufficiently early in the acquisition process (Figure 1-1). Fixes instituted during the Demonstration/Validation (DEM/VAL) Phase cost significantly less than those required in the Engineering and Manufacturing Development (EMD) Phase after most design decisions have been made.

In 1983, the Assistant Secretary of Defense made the following statement to the Senate Committee on Governmental Affairs regarding the importance of T&E:

> . . . the criterion should not be how quickly we can field any new weapon, but rather how quickly we can field a new weapon that works. The only weapons that would be significantly delayed would be the ones that operational testing shows to be unsuitable for combat, and I cannot believe that any of us would advocate saddling our fighting forces with any of those. In fact, the most likely effect of operational testing is not to delay, but to accelerate the development process. Trying to fix a faulty weapon after it's in the field — if it can still be fixed — is a far slower process than fixing the design before it goes into production.

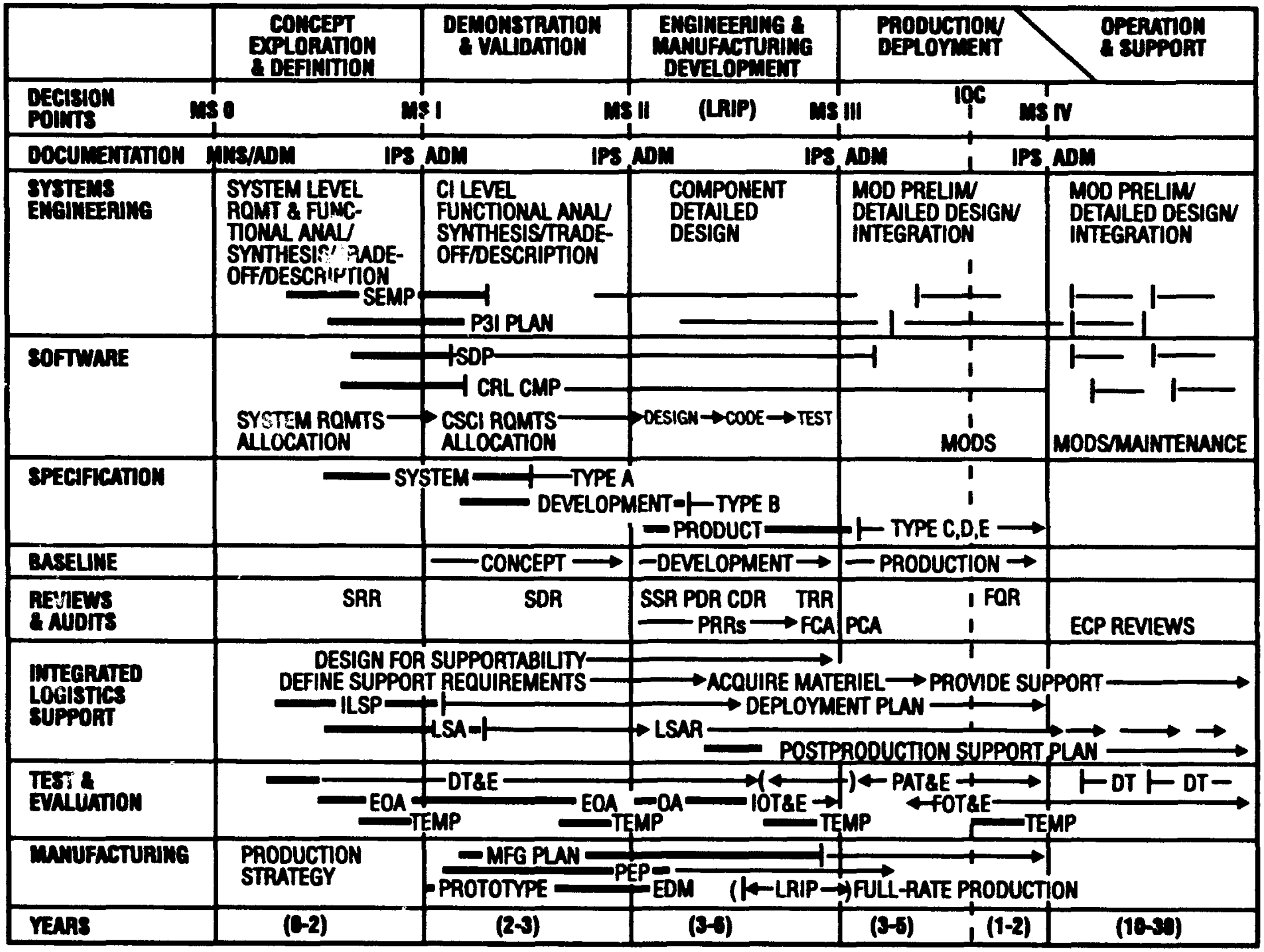


Figure 1-1. System Life-Cycle Technical Activities

Thus, T&E may reduce cost, schedule and technical risks. A third type of risk involved in the development and acquisition of new systems is technical risk. Test and evaluation of parts, components, subsystems and systems can also be used to estimate and manage this technical risk.

Test and evaluation results figure prominently in the decisions reached at design and milestone reviews. However, the fact that T&E results are required at major decision points does not presuppose that T&E results must always be favorable. The final decision responsibility lies with the executive who must examine the critical issues and weigh the facts. Only he can determine the weight and importance that is to be attributed to a system's diverse capabilities and shortcomings and the degree of risk he is willing to accept. The decision-making authority will be unable to make this judgment without a solid base of information provided by T&E. Figure 1-2 illustrates the life-cycle cost of the system and how decisions impact program expenditures.

A Defense Science Board 1983 Task Force focused on the reduction of risk in program acquisition (Reference 42). This group made the following observations:

• A poorly-designed product cannot be properly tested or produced;

• Control techniques needed to successfully complete the design, test and production of an item dictate the management system required;

• The industrial process of weapon system acquisition demands a better understanding and implementation of basic engineering and manufactuing disciplines;

• The industrial process is focused on the design, test and production of a product;

• The design, test and production processes are a continuum of interdependent disciplines. Failure to perform well in one area will result in failure to do well in all areas. When this happens, as it does too often, a high-risk program results with equipment fielded later and at far greater cost than planned.

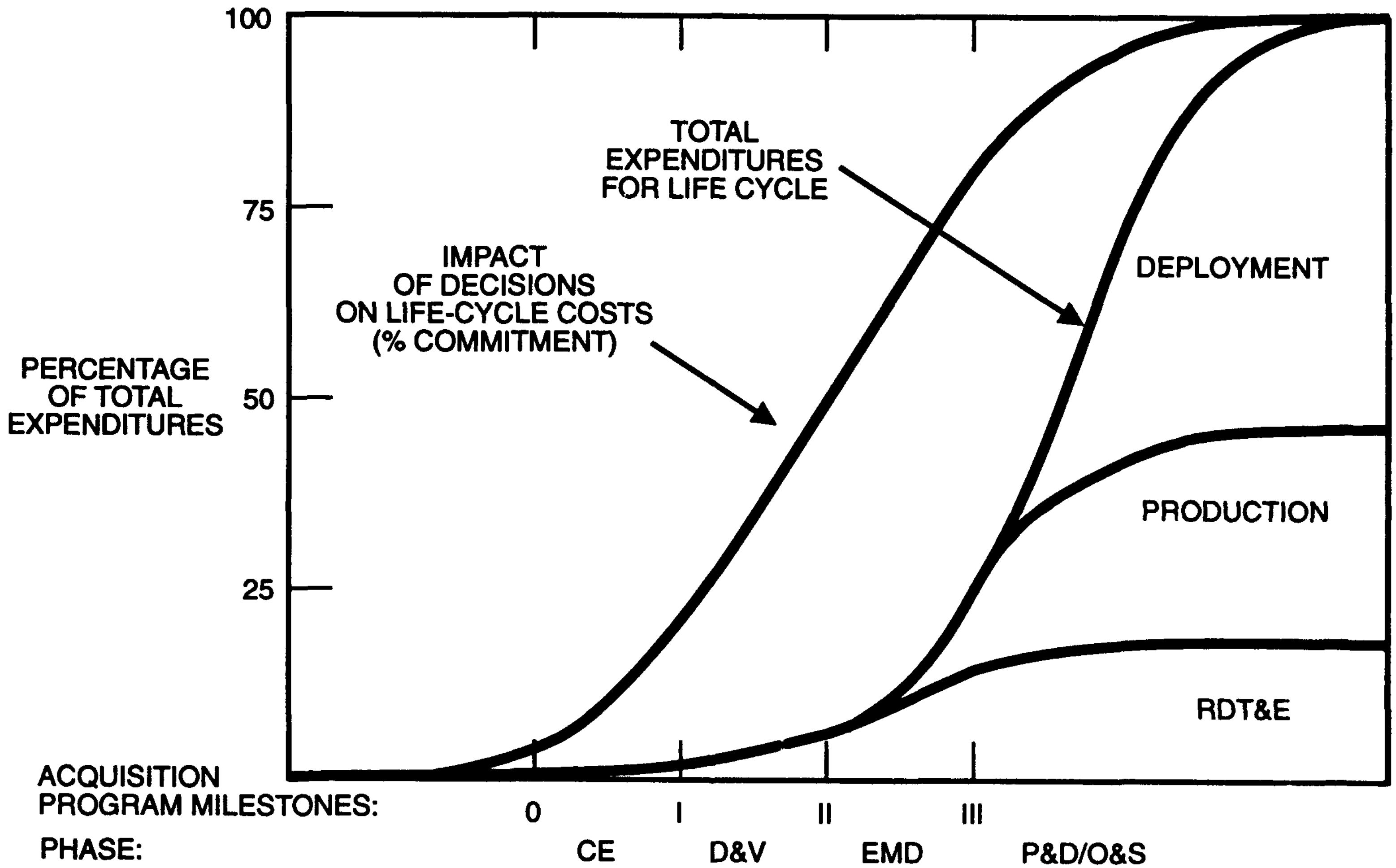
The Task Force developed a set of templates for use in establishing and maintaining low-risk programs. Each template describes an area of risk and then specifies technical methods for reducing that risk. Program managers and test managers may wish to consult these templates for guidance in reducing the risks frequently associated with test programs. Sample risk management templates are found in DOD 4245.7-M, "Transition from Development to Production."

1.3 THE T&E CONTRIBUTION AT MAJOR MILESTONES

Test and evaluation progress is monitored by the Office of the Secretary of Defense (OSD) throughout the acquisition process. Their oversight extends to the major materiel acquisitions or designated acquisitions, which are about five percent of all the acquisitions being managed within DOD. Test and evaluation officials within OSD render independent assessments to the Defense Acquisition Board, the Defense Acquisition Executive, and the Secretary of Defense at each major system milestone. These assessments are based on the following T&E information:

• The Test and Evaluation Master Plan (TEMP) and more detailed supporting documents developed by responsible Service activities;

• Service test agency reports and briefings;



Source: Defense Systems Management College

Figure 1-2. Life-Cycle-Cost Decision Impact and Expenditures

• Development test and evaluation data from sources such as Service program managers, laboratories, industry developers, studies and analyses.

At Milestone I, the OSD T&E assessment reflects an evaluation of system concepts and alternatives based on specific objectives and thresholds established in an approved preliminary TEMP. At Milestone II, it includes an assessment of previously established test plans and results. At Milestone III, reviews verify the operational effectiveness and suitability of major weapon systems.

A primary contribution made by T&E is the detection and reporting of deficiencies that may adversely impact the performance capability or availability/supportability of a system. A deficiency reporting process is used throughout the acquisition process to report, evaluate and track system deficiencies and to provide the impetus for corrective actions.

1.3.1 Test Contributions Prior to Milestone I

During the Concept Exploration and Definition Phase prior to Milestone I, laboratory testing, modeling and simulations are conducted by the contractor and the development agency to demonstrate and assess the capabilities of key subsystems and components. The test and simulation designs are based on the requirements documented in the Mission Need Statement. Studies, analyses, simulation and test data are used by the development agency to explore and evaluate alternative concept designs proposed to satisfy the requirements. Also during this period, the operational test agency (OTA) monitors concept exploration T&E to gather information for future T&E planning and to provide effectiveness and suitability input desired by the program manager. The OTA also conducts operational assessments, as feasible, to assess the operational impact of candidate technical approaches and to assist in selecting preferred alternative systems concepts.

Toward the end of the phase, the development agency prepares the Development Test and Evaluation (DT&E) System Concept Report. This report records and presents T&E results of system design(s) engineering and performance evaluations compared to stated requirements and concept specifications. This information is incorporated into the Program Manager's Status Briefing and the Integrated Program Summary, key documents that form the basis for the Milestone I decision to proceed to the DEM/VAL Phase.

1.3.2 Test Contributions Prior to Milestone II

During the DEM/VAL Phase prior to Milestone II, concepts approved for demonstration and validation form the baseline that is used for detailed test planning.

The development agency conducts development test and evaluation during the DEM/VAL Phase to assist with engineering design, system development and to verify attainment of technical performance specifications and program objectives. The DT&E includes T&E of components, subsystems and prototype development models. Test and evaluation of functional compatibility and interoperability with existing and planned equipment and systems is also included. During this phase of testing, adequate DT&E is accomplished to ensure engineering is reasonably complete (including survivability/vulnerability, compatibility, transportability, interoperability, reliability, maintainability, safety, human factors, and logistic supportability). Also, this phase confirms that all significant design problems have been identified and solutions to these problems are in hand.

The Service operational test and evaluation agency (OT&E) conducts early operational assessments to estimate the system's operational effectiveness and suitability; identifies needed modifications; and provides information on tactics, doctrine, organization and personnel requirements. The OT&E program is accomplished in an environment as operationally realistic as possible. Typical operational and support personnel are used to obtain a valid estimate of the user's capability to operate and maintain the system. The user of the system monitors T&E during the DEM/VAL Phase. Some of the most important products of user monitoring are the attainment of early orientation and advanced training, demonstrations of system performance, and valid operational testing (OT) assessments of system maintainability and supportability.

The development agency prepares a report on the results of demonstration and validation DT&E for review by the Service headquarters and the Service acquisition review council prior to system acquisition review by DOD. The report includes the results of testing and supporting information, conclusions and recommendations for full-scale development. At the same time, the OT&E agency prepares independent early operational assessments, which contain estimates of the system's operational effectiveness and suitability. The OT&E assessments provide a permanent record of OT&E accomplished, an audit trail of OT&E data, test results, conclusions and recommendations. This information is used to support the development of the Integrated Program Summary, which is prepared for Milestone II, and recommends which of the alternative systems studied in the DEM/VAL Phase will proceed into engineering and manufacturing development.

1.3.3 Test Contributions Prior to Milestone III

Prior to Milestone III, the objective of the EMD Phase is to design, fabricate and test a preproduction system that closely approximates the final product. Test and evaluation activities during this period yield much useful information. For example, data obtained during EMD test and evaluation is used to assist in evaluating the system's maintenance training requirements and the proposed training program. Test results generated during EMD test and evaluation also support the user in refining and updating tactics.

During the EMD Phase, development test and evaluation is conducted to satisfy the following objectives:

(1) As specified in program documents, assess the critical technical issues:

(a) Determine how well the development contract specifications have been met;

(b) Identify system technical deficiencies and appropriate corrective actions;

(c) Determine whether the system is compatible and interoperable with existing and planned equipment or systems;

(d) Estimate the reliability, maintainability and availability of the system after it is deployed;

(e) Determine whether the system is safe and ready for OT&E;

(f) Validate any configuration changes caused by correcting deficiencies, modifications or product improvements;

(g) Assess human factors and identify limiting factors;

(2) Assess the technical risk and evaluate the trade-offs among specifications, operational requirements, life-cycle costs and schedules;

(3) Assess the survivability, vulnerability and logistic supportability of the system;

(4) Verify the accuracy and completeness of the technical documentation developed to maintain and operate the weapons system;

(5) Gather information for training programs and technical training materials needed to support the weapon system;

(6) Provide information on environmental issues for use in preparing environmental impact assessments;

(7) Determine system performance limitations and safe operating parameters;

(8) Using Live Fire Test (LFT), evaluate vulnerability or lethality of a weapon system as appropriate and as required by law.

Initial OT&E is conducted prior to the production decision at Milestone III to:

(1) Estimate the operational effectiveness and suitability of the system;

(2) Identify operational deficiencies;

(3) Recommend and evaluate changes in production configuration;

(4) Provide information for developing and refining logistics support requirements for the system and training, tactics, techniques and doctrine;

(5) Provide information to refine operation and support (O&S) cost estimates and identify system characteristics or deficiencies that can significantly impact O&S costs;

(6) Determine whether the technical publications and support equipment are adequate; in the operational environment.

Thus, T&E activities intensify during the EMD Phase and make significant contributions to the overall acquisition decision process.

1.3.4 Test Contributions After The Production Decision

After Milestone III, when the production decision is made, T&E activities continue to provide important insights. Tests described in the TEMP and not completed during the EMD Phase are completed during the Production and Deployment Phase. The residual DT&E is usually limited to all-weather testing, correction of deficiencies and engineering modifications. System elements are integrated into the final operational configuration, and development testing is completed when the system performance requirements are met. During the Production Phase, government representatives normally monitor or conduct the production acceptance test and evaluation (PAT&E). Each system is verified by PAT&E for compliance with the requirements and specifications of the contract.

Postproduction testing requirements may result from an acquisition strategy calling for block changes to accommodate engineering changes or the use of preplanned product improvements (P^3I). This will allow parallel development of high-risk technology and modular insertion of system upgrades into production equipment. Technology breakthroughs and significant threat changes may require system modifications. The development of the modifications will require development testing; and, if system performance is significantly changed, operational testing may be appropriate.

Operational T&E activities continue after the production decision in the form of follow-on operational test and evaluation (FOT&E). The initial phase of FOT&E may be conducted by either the OT&E agency or user commands, depending on Service directives. It verifies the operational effectiveness and suitability of the production system and determines if deficiencies identified during the initial OT&E have been corrected. A second phase of FOT&E is conducted by the user to refine doctrine, tactics, techniques and training programs for the life of the system.

The OT&E agency prepares a final report at the conclusion of its management phase of FOT&E. This report records test results, describes the evaluation accomplished to satisfy critical issues and objectives established for FOT&E and documents its assessment of deficiencies resolved during EMD. Deficiencies that are not corrected are recorded with recommended corrective actions.

A final report on FOT&E is also prepared by the using command test team with emphasis on the operational utility of the system when operated, maintained and supported by operational personnel using the concepts specified for the system. Specific attention is devoted to the following:

(1) The degree to which the system accomplishes the mission when employed by operational personnel in a realistic scenario with the appropriate organization, doctrine, threat (including countermeasures and nuclear threats), environment and using tactics and techniques developed during earlier FOT&E;

(2) The degree to which the system can be placed in operational field use, with specific evaluations of availability, compatibility, transportability, interoperability, reliability, wartime usage rates, maintainability, safety, human factors, manpower supportability, logistics supportability and training requirements;

(3) The conditions under which the system was tested including the natural weather and climatic conditions, terrain effects, battlefield disturbances and enemy threat conditions;

(4) The ability of the system to perform its required functions for the duration of a specified mission profile;

(5) System weaknesses such as the vulnerability of the system to exploitation by countermeasures techniques and the practicality and probability of an adversary exploiting the susceptibility of a system in combat.

A specific evaluation of the manpower and logistics changes needed for the effective integration of the system into the user's inventory is also made. These assessments provide essential input for the later phases of the system acquisition cycle.

1.4 SUMMARY

"Risk management is the means by which the program areas of vulnerability and concern are identified and managed." (Reference 20). Test and evaluation is the discipline that helps to illuminate those areas of vulnerability. The importance of T&E in the acquisition process is summarized well in a December 1986 report produced by the General Accounting Office. While the following remarks focus on OT&E, they also serve to underscore the importance of the T&E process as a whole:

OT&E is the primary means of assessing weapon system performance. OT&E results are important in making key decisions in the acquisition process, especially the decision to proceed from development to production. OT&E results provide an indication of how well new systems will work and can be invaluable in identifying ineffective or unreliable systems before they are produced.

Starting production before adequate OT&E is completed has some risks. If adequate OT&E is not done and the weapon system does not perform satisfactorily in the field, significant changes may be required. Moreover, the changes will not be limited to a few developmental models, but may also be applied to items already produced and deployed. In extreme situations, DoD also risks (1) deploying systems which cannot adequately perform significant portions of their missions, thus degrading our deterrent/defensive capabilities and (2) endangering the safety of military personnel who operate and maintain the systems.

2

THE TEST AND EVALUATION PROCESS

2.1 INTRODUCTION

The fundamental purpose of test and evaluation (T&E) in a defense system's development and acquisition program is to identify the areas of risk to be reduced or eliminated. During the early phases of development, T&E is conducted to demonstrate the feasibility of conceptual approaches, evaluate design risk, identify design alternatives, compare and analyze trade-offs, and estimate satisfaction of operational requirements. As a system undergoes design and development, the emphasis in testing moves gradually from development test and evaluation (DT&E), which is concerned chiefly with attainment of engineering design goals, to operational test and evaluation (OT&E), which focuses on questions of operational effectiveness, suitability and supportability. Although there are usually separate development and operational test events, DT&E and OT&E are not necessarily serial phases in the evolution of a weapon system. Combined and concurrent development and operational testing is encouraged when appropriate (Reference 16).

Test and evaluation has its origins in the testing of hardware; this tradition is heavily embedded in its vocabulary and procedures. The advent of software-intensive systems has brought new challenges and new approaches to testing, which are discussed in Chapter 18 of this management guide. Remaining constant throughout the T&E process, whether testing hardware or software, is the need for thorough, logical, systematic and early test planning including feedback of well-documented and unbiased T&E results to system developers, users and decision-makers.

Test and evaluation has many useful functions and provides information to many customers. The T&E gives information to: developers for identifying and resolving technical difficulties; decision-makers responsible for procuring a new system and for the best use of limited resources; and to operational users for refining requirements and supporting development of effective tactics, doctrine and procedures.

2.2 DEFENSE SYSTEM ACQUISITION PROCESS

The defense system acquisition process was revised in 1991 to make it less costly, less time-consuming and more responsive to the needs of the operational community. As it is now structured, the defense system life cycle consists of the following four phases:

(1) Concept Exploration and Definition

(2) Demonstration and Validation

(3) Engineering and Manufacturing Development

(4) Production and Deployment/Operational and Support.

As Figure 2-1 shows, these phases are separated by key decision points when a milestone (MS) decision authority reviews a program and authorizes advancement to the next stage in the cycle. Thus T&E planning and test results play an important part in the milestone review process.

The following brief description of the defense system acquisition process shows how T&E fits within the context of the larger process. The description is based primarily upon information found in DOD Instruction 5000.2 (Reference 16).

2.2.1 Concept Exploration and Definition (Phase 0)

The defense system acquisition process begins with the submission of a Mission Need Statement. A Concept Exploration and Definition Phase follows the Milestone 0 decision during which alternative approaches for satisfying the *requirement* are investigated. The Concept Exploration and Definition Phase concludes with the Milestone I selection of a concept or concepts to enter a Demonstration and Validation Phase. The Milestone I decision establishes broad objectives for program cost, schedule, operational effectiveness and suitability. Key documents for the T&E manager at the time of the Milestone I review are the Acquisition Decision Memorandum (ADM) (exit criteria), Integrated Program Summary (IPS), Operational Requirement Document (ORD), Mission Need Statement (MNS), System Threat Assessment Report (STAR), the Test and Evaluation Master Plan (TEMP), and the Integrated Logistics Support Plan (ILSP)/Logistics Support Analysis (LSA). Additional program management documents prepared before Milestone I include: the Cost and Operational Effectiveness Analysis (COEA), Independent Cost Estimate, and Concept Baseline, which summarizes the weapon's functional specifications, performance parameters, and cost and schedule objectives.

2.2.2 Demonstration and Validation (Phase I)

After the Milestone I decision for a program start, the Demonstration and Validation Phase begins during which selected concepts, typically brassboard or early prototype, are refined through engineering and analysis. This phase ends with the Milestone II decision to either enter into engineering and manufacturing development (EMD) or terminate the program. The Milestone II decision establishes more specific cost, schedule, operational effectiveness and suitability objectives and thresholds. Documents interesting to the T&E manager at the time of the Milestone II review include the ADM (exit criteria), IPS, updated TEMP, COEA, updated ORD, Development Baseline, the early Operational Assessment and low-rate initial production (LRIP) guidance.

2.2.3 Engineering and Manufacturing Development (Phase II)

During the EMD Phase, the selected system and its principal items of support are fabricated as engineer development models. This phase ends with the Milestone III decision to enter full-rate production of the system. Key documents for the T&E manager at the time of the Milestone III review are the IPS, updated TEMP, Beyond LRIP Report, and Live Fire Test Report. The director of OT&E is required by law to document his assessment of the adequacy of OT&E and the reported operational effectiveness and suitability of the system. This is done in the Beyond LRIP Report. Also mandated by law is the requirement for the director of T&E to write the Live Fire Test Report prior to proceeding beyond LRIP.

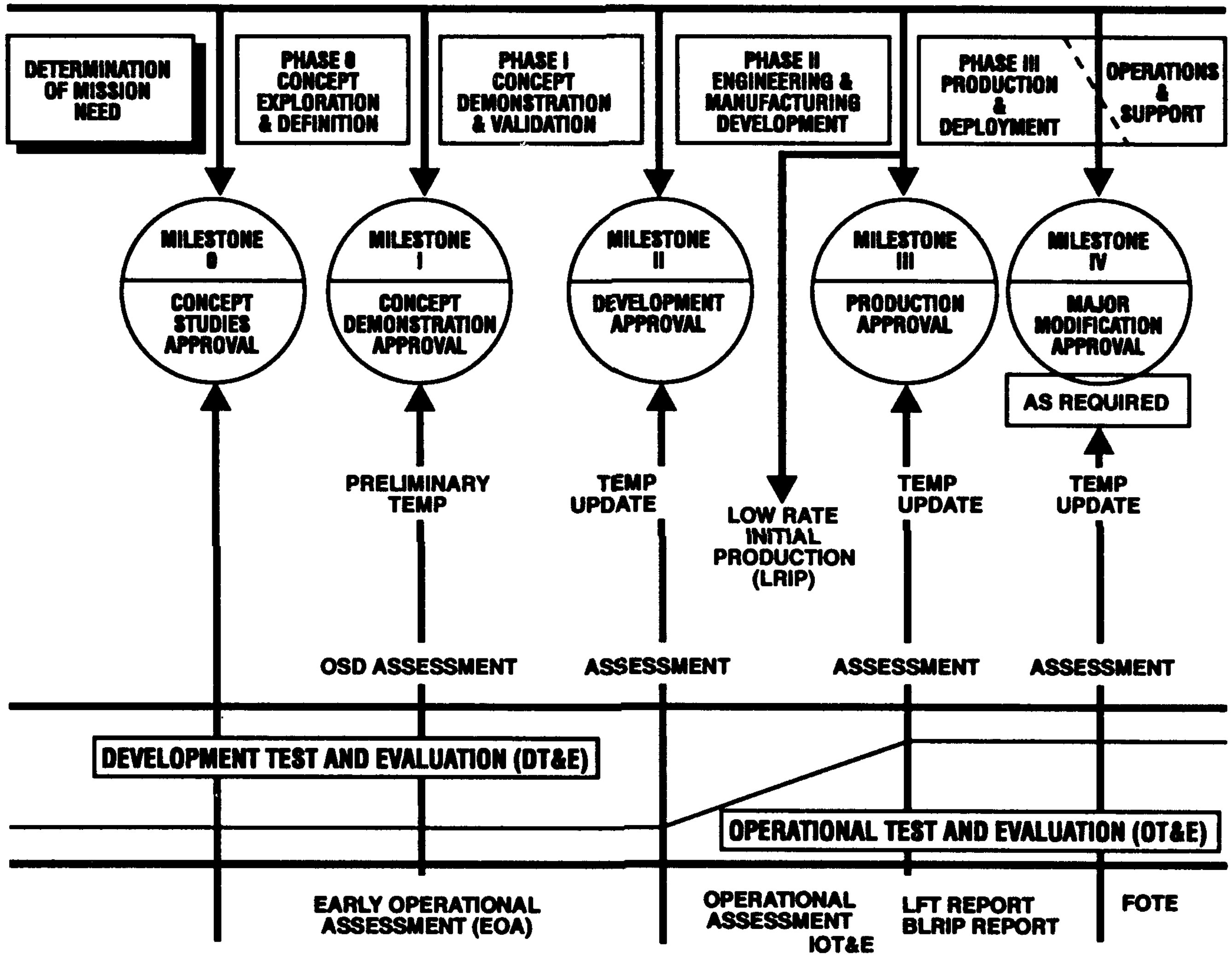


Figure 2-1. Testing and the Acquisition Process

2.2.4 Production/Deployment and Operations/Support (Phase III)

The Milestone III decision is followed by a Full-Rate Production/Deployment Phase. This phase ends with a Major Modification Approval (Milestone IV) to identify the actions and resources needed to achieve and maintain operational readiness and support objectives. To determine whether major upgrades are necessary or deficiencies warrant consideration of replacement, the Milestone IV decision encompasses a review of system operational effectiveness, suitability and readiness. In preparation for Milestones IV, the IPS, TEMP, and production baseline are updated to describe the program status, changes and issues.

2.3 T&E AND THE SYSTEMS ENGINEERING PROCESS

In the early 1970s, Department of Defense (DOD) test policy became more formalized and placed greater emphasis on T&E as a continuing function throughout the acquisition cycle. These policies stressed the use of T&E to reduce acquisition risk and provide early and continuing estimates of system operational effectiveness and operational suitability. To meet these objectives, appropriate test activities had to be fully integrated into the overall development process. From a systems engineering perspective, test planning, testing and analysis of test results are integral parts of the basic product definition process.

In MIL-STD-499, systems engineering is defined in the DOD context: "Systems engineering is an interdisciplinary approach to evolve and verify an integrated and optimally balanced set of product and process designs that satisfy user needs and provide information for management decision making." (Figure 2-2)

A system's life cycle begins with the user's needs, which are expressed as constraints, and the capability requirements needed to satisfy mission objectives. Systems engineering is essential in the earliest planning period, in conceiving the system concept and defining performance requirements for system elements. As the detailed design is prepared, systems engineers ensure balanced influence of all required design specialties, including "testability." They resolve interface problems, perform design reviews, perform trade-off analyses and assist in verifying performance.

The days when one or two individuals could design a complex system, especially a huge, modern-age weapon system, are past. Now systems are too complex for a small number of generalists to accommodate; they require too much in-depth knowledge over a broad range of areas and technical disciplines. System engineers coordinate the many specialized engineers involved in the concurrent engineering process and are responsible for the integration of the components into a system.

Through interdisciplinary integration systems engineering manages the progress of product definition from system level, to configuration-item level, detailed level, deficiency correction, and modifications/product improvements. Test results provide feedback to analyze the design progress toward performance goals. Tools of systems engineering include design reviews, configuration management, simulation, technical performance measurement, trade-off analysis and specifications.

What products are produced by systems engineering? It determines what specialists are required, what segments and nondevelopmental items are used, design performance limits, trade-off criteria, how to test, when to test, how to document

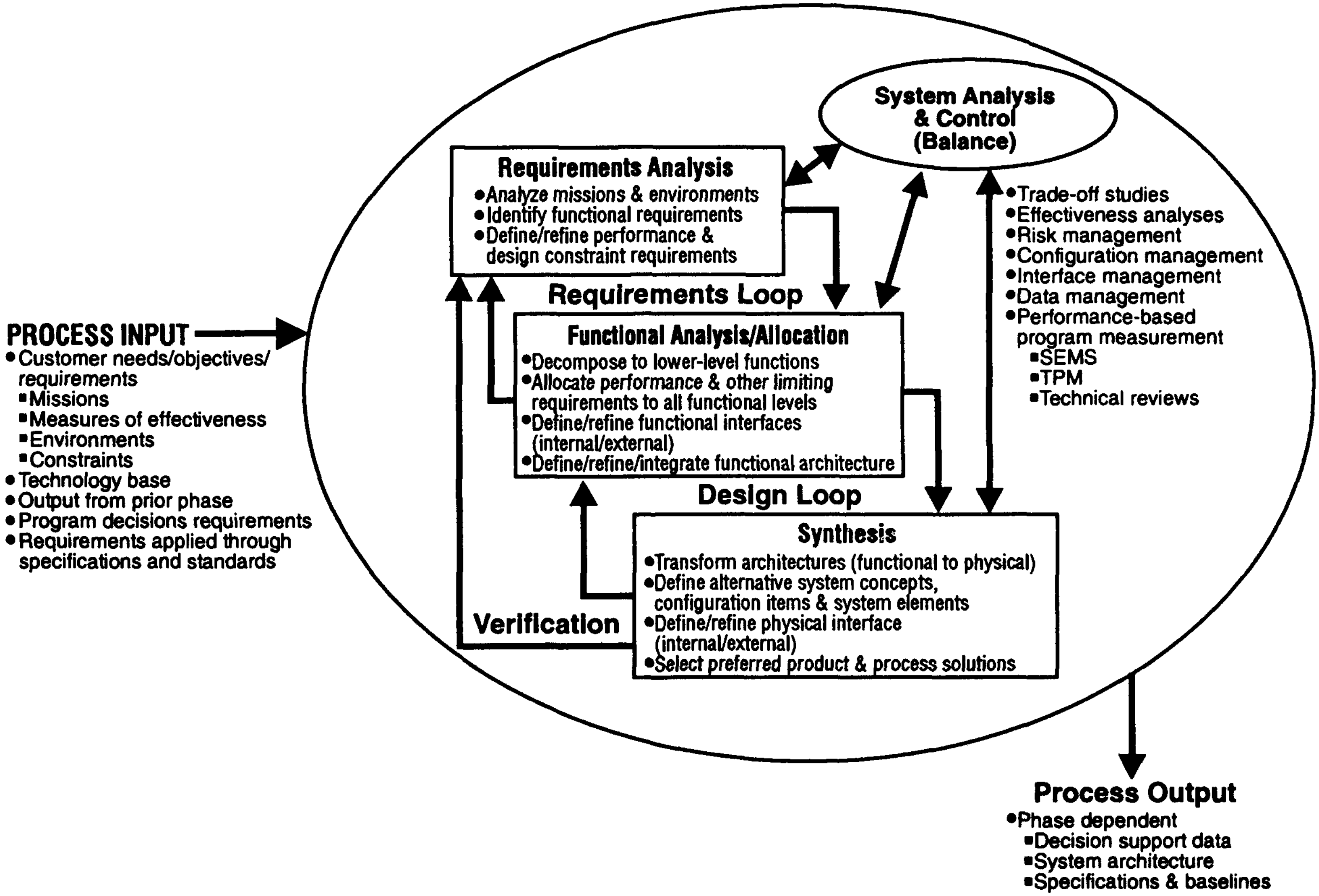


Figure 2-2. The Systems Engineering Process

(specifications), and what management controls to apply (technical performance measurement and design reviews).

Development testing (DT) and operational testing (OT) support the technical reviews used to monitor the systems engineering process. More information on the reviews is contained in Chapter 8.

2.3.1 The Systems Engineering Process

The systems engineering process is the iterative logical sequence of analysis, design, test and decision activities that transforms an operational need into the descriptions required for production and fielding of all operational and support system elements. This process consists of four activities. They include functional analysis, synthesis, evaluation and decision (trade-off) and description of system elements.

The functional analysis activity identifies what the system, component or part must do. It works from the top downward ensuring requirements traceability and revealing alternative concepts. This is done without assuming how functions will be accomplished. The product is a series of alternative Functional Flow Block Diagrams (FFBD). A functional analysis can be applied at every level of development. At the system level, it may be a contractor or Service effort. During the Concept Exploration Phase, developmental testers assist the functional analysis activity to help determine what each component's role will be as part of the system being developed.

The synthesis activity involves invention — conceiving ways to do each FFBD task — to answer the "how" question. Next, the physical interfaces implied by the "how" answers, are carefully identified (topological or temporal). The answers must reflect all technology selection factors. Synthesis tools include Requirements Allocation Sheets (RAS), which translate functional statements into design requirements and permit a long and complex interactive invention process with control, visibility and requirements traceability. Developmental testers conduct demonstration/validation testing to determine how the components will perform assigned functions to assist this synthesis activity.

The evaluation and decision activity is a trade-off of alternative approaches to "how." This activity is conducted in accordance with decision criteria set by higher-level technical requirements for such things as life-cycle costs, effectiveness, reliability, availability, maintainability, risk limits, schedule, etc. It is repeated at each level of development. The evaluation and decision activity is assisted by developmental testers during the later Demonstration and Validation Phase and the Engineering and Manufacturing Development Phase when competitive testing between alternative approaches is performed.

The final activity is a description of system elements. Developing as the result of previous activities and as the final system design is determined, this activity takes form when specifications are verified through testing and when reviewed in the Physical Configuration and Functional Configuration Audits. During the EMD Phase, operational testers assist in this activity. They conduct operational testing of the test items/systems to help determine the personnel, equipment, facilities, software and technical data requirements of the new system when used by typical military personnel. Figure 2-2, System Engineering Process, depicts the activities and their interactions.

2.3.2 The System Engineering Management Plan

The System Engineering Management Plan (SEMP) is a concise, top-level management

TESTING	DEFINED IN	PERFORMED AGAINST	PARTICIPANTS
SYSTEM ENGINEERING			
QUALIFICATION ACCEPTANCE	SOW CONTRACT SPECS WBS	MIL-STD/SPEC MISSION PROFILE/ ENVIRONMENT	PROJ ENGR DPML FUNC ENGR Kr/GOVT TEAM
TEST AND EVALUATION			
SYSTEM DT & OT	TEMP DT/OT PLANS	STAR/ORD OPS CONCEPT MAINT CONCEPT	DTE/DOTE DT/OT PERS USER/SUPT PERS TEST WORK GP

Figure 2-3. Systems Engineering and Test and Evaluation

plan for the integration of all system design activities. Its purpose is to make visible the organization; mechanisms for direction and control; and personnel for the attainment of cost, performance and schedule objectives. The SEMP defines and describes the type and degree of system engineering management, the system engineering process, and the integration of related engineering programs. The design evolution process, which is described in the SEMP, forms the basis for comprehensive test and evaluation planning.

The TEMP must be consistent with the SEMP. The testing program outlined in the TEMP must provide the technical performance measurements data required for all design decision points, audits and reviews that are a part of the system's engineering process outlined in the SEMP. The configuration management process outlined in the SEMP controls the baseline for the test programs and incorporates design modifications to the baseline determined to be necessary by T&E.

The TEMP and the SEMP must be traceable to each other. The system description in the TEMP must be traceable to systems engineering documentation such as the FFBDs, the RASs, and the Test Requirements Sheets (TRSs). Key functions and interfaces of the system with other systems must be described and correlated with the systems engineering documentation and the system specification (Type A). Operational and technical thresholds in the SEMP include specific performance requirements that become test planning limits. They must be traceable through the planned systems engineering documentation and can be correlated to the content of the Technical Performance Measurement (TPM) Program. Failure criteria for reliability thresholds during OT&E testing must be delineated and agreed upon by the program manager and the operational test director and reflected in the SEMP and the TEMP.

2.3.3 Technical Performance Measurement

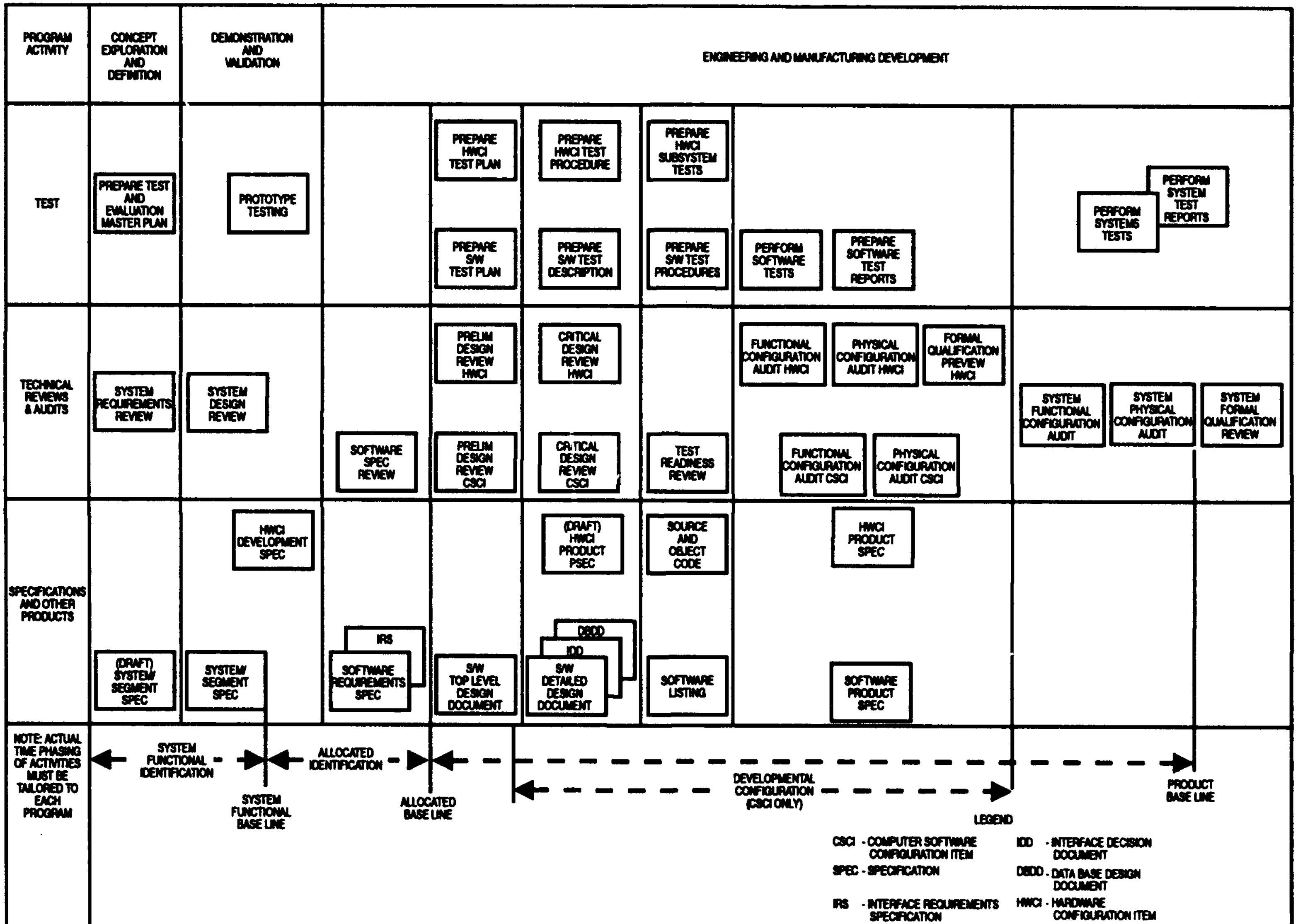
Technical performance measurement identifies critical technical parameters that are at risk during design. It tracks evaluation and test data, makes predictions about whether the parameter can achieve final technical success within the allocated resources, and assists in managing the technical program.

The TPM Program is an integral part of the T&E program. The TPM is defined as product design assessment and forms the backbone of the development testing program. It estimates, through engineering analyses and tests, the values of essential performance parameters of the current program design. It serves as a major input in the continuous overall evaluation of operational effectiveness and suitability. Design reviews are conducted to measure systems engineering progress. For more information, see Chapter 8. Figure 2-4 depicts the technical reviews that usually take place during the systems engineering process and the related specification documents.

2.3.4 Product Baselining and T&E

The systems engineering process establishes a product baseline throughout the acquisition cycle. This baseline can be modified with the results of engineering and testing. The testing used to prove the technical or development baseline is rarely the same as the operational testing or production baseline.

Related to the product baseline is the process of configuration management. Configuration management benefits the test and evaluation community in two ways.



Source: MIL STD 1521 B

Figure 2-4. Design Reviews

Through configuration management, the baseline product to be used for testing is determined. Also, changes that occur to the baseline as a result of testing and design reviews are incorporated into the test article before the new phase of testing (to prevent retest of a bad design).

2.4 DEFINITIONS

Test and evaluation is the deliberate and rational generation of data, which concerns the nature of the emerging system, and the creation of information useful to the technical and managerial personnel controlling its development. In the broad sense, T&E may be defined as all physical testing, modeling, simulation, experimentation and related analyses performed during research, development, introduction and employment of a weapon system or subsystem. *The Glossary: Defense Acquisition Acro*nyms *and Terms,* produced by the Defense Systems Management College, September 1991, defines "Test" and "Test and Evaluation" as follows:

> A "test" is any program or procedure which is designed to obtain, verify, or provide data for the evaluation of: research and development (other than laboratory experiments); progress in accomplishing development objectives; or performance and operational capability of systems, subsystems, components, and equipment items.

> "Test and Evaluation" is the process by which a system or components are compared against requirements and specifications through testing. The results are evaluated to assess progress of design, performance, supportability, etc. There are three types of T&E- Development (DT&E), Operational (OT&E), and Production Acceptance (PAT&E) — occurring during the acquisition cycle. DT&E is conducted to assist the engineering design and development process and to verify attainment of technical performance specifications and objectives. OT&E is conducted to estimate a system's operational effectiveness and suitability, identify needed modifications, and provide information on tactics, doctrine, organization, and personnel requirements. PAT&E is conducted on production items to demonstrate that those items meet the requirements and specifications of the procuring contracts or agreements. OT&E is further subdivided into two phases- Initial Operational (IOT&E) and Follow-on Operational (FOT&E). IOT&E must be conducted before the production decision (MS III) to provide a credible estimate of operational effectiveness and suitability. Therefore, IOT&E must be conducted on a system as close to a production configuration as possible, in an operationally realistic environment, by typical user personnel. FOT&E is conducted on the deployed system to determine if operational effectiveness and suitability is, in fact, being attained.

2.5 SUMMARY

Test and evaluation is an engineering tool used to reduce risk throughout the defense system acquisition cycle. This cycle consists of five phases separated by discrete milestones. The T&E results are used to support the design reviews that form an important part of the system engineering process used by system developers and to aid in the milestone decision process used by senior decision authorities in the Department of Defense.

3

T&E POLICY STRUCTURE AND OVERSIGHT MECHANISM

3.1 INTRODUCTION

This chapter provides an overview of the policy and organizations that govern the conduct of test and evaluation (T&E) activities within the Department of Defense (DOD) and discusses congressional legislation and activities for compliance by the DOD. It outlines the responsibilities of DOD test organizations, at the Office of the Secretary of Defense (OSD) and Service levels, and describes related T&E policy.

3.2 THE CONGRESS

The Congress has shown a long-standing interest in influencing the DOD acquisition process. During the early 1970s, in response to urging by the Congress and recommendations by a Presidential Blue Ribbon Panel on Defense Management, the Deputy Secretary of Defense, David Packard, promulgated a package of policy initiatives that established the Defense Systems Acquisition Review Council (DSARC). The DSARC was organized to resolve acquisition issues, whenever possible, and to provide recommendations to the Secretary of Defense (SECDEF) on the acquisition of major weapon systems. Also, as a result of the Congressional Directives, the Army and Air Force established independent operational test agencies. The Navy Operational Test and Evaluation Force was established in the late 1960s. In 1983, similar concerns led the Congress to direct the establishment of the independent Office of Director, Operational Test and Evaluation (DOT&E), within the OSD. In 1985 a report released by the President's Blue Ribbon Commission on Defense Management, chaired by David Packard, made significant recommendations on the management and oversight of DOD's acquisition process and, specifically, test and evaluation. All the Commission's recommendations have not been implemented, and the full impact of these recommendations is not yet realized. In FY 1987 the Defense Authorization Act required live fire testing of weapon systems before the Production Phase begins.

The DOD is required to provide to the Congress the following reports on test and evaluation:

- Congressional Data Sheets (CDS). The CDS are annual reports on each major system acquisition. They must be updated before the contract is awarded and when procurement of the system is requested in the fiscal year. The CDS describe the development test and evaluation (DT&E) and operational test and evaluation (OT&E) to be performed and system characteristics.

- Selected Acquisition Report (SAR). The SAR describes the system characteristics required and outlines significant progress and problems encountered. It lists tests completed and issues identified during testing.

• Annual System Operational Test Report. The Annual Systems Operational Test Report is provided by the DOT&E to the SECDEF and the committees on Armed Services and Appropriations. The report provides a narrative and resource summary of OT&E and OT&E-related issues, activities, and assessments.

• Beyond Low-Rate Initial Production (LRIP) Report. Before proceeding beyond LRIP for each major system acquisition program, the Director, Operational Test and Evaluation, must report to the SECDEF and the Congress. This report addresses the adequacy of OT&E and whether the T&E results confirm that the tested item or component is effective and suitable for combat.

3.3 OSD OVERSIGHT STRUCTURE

The DOD organization for the oversight of T&E is illustrated in Figure 3-1. In the OSD, T&E oversight is performed by two primary offices: the Director, Test and Evaluation (DTE) and the Director Operational Test and Evaluation (DOT&E). The management of acquisition programs in OSD is performed by the Defense Acquisition Executive (DAE), who uses the Defense Acquisition Board (DAB) and subcommittees to process information for decisions. The Under Secretary of Defense for Acquisition (USD(A)) uses the DAB and its committees to provide the senior-level decision process for the acquisition of weapon systems.

3.3.1 Defense Acquisition Executive (DAE)

The DAE position, established in September 1986, is held by the USD(A). The responsibilities include "establishing policies for acquisition (including procurement, research and development, logistics, development testing, and contracts administration) for all elements of the Department of Defense." His charter includes the authority over the Service and defense agencies on policy, procedure and execution of the acquisition process.

3.3.2 Defense Acquisition Board (DAB)

The DAB is the primary forum used by OSD to provide advice, assistance and recommendations, and to resolve issues regarding all operating and policy aspects of the DOD acquisition system. The DAB is the senior management acquisition board chaired by the DAE and attended by the Vice Chairman of Joint Chiefs of Staff, Principal Under Secretary of Defense for Acquisition and the component acquisition executives. The DAB conducts business through working committees (DODI 5000.2).

3.3.3 Defense Planning Resources Board (DPRB)

The *DPRB* was established by the SECDEF in 1979 to advise the SECDEF on policy, planning, program and budget issues. The DPRB is chaired by the Deputy Secretary of Defense and is responsible for the management and oversight of all aspects of the DOD planning, programming and budgeting process. It oversees the budget reviews process and, therefore, has a major impact on test and evaluation resources.

3.3.4 Director Test and Evaluation (DTE)

The DTE serves as the principal staff assistant and advisor to the USD(A) for T&E matters. He has authority and responsibility for all DT&E conducted on designated major programs. The DTE organization is illustrated in Figure 3-1.

3.3.4.1 Duties of the DTE

Within the acquisition community, the DTE:

• Serves as the focal point for coordina-

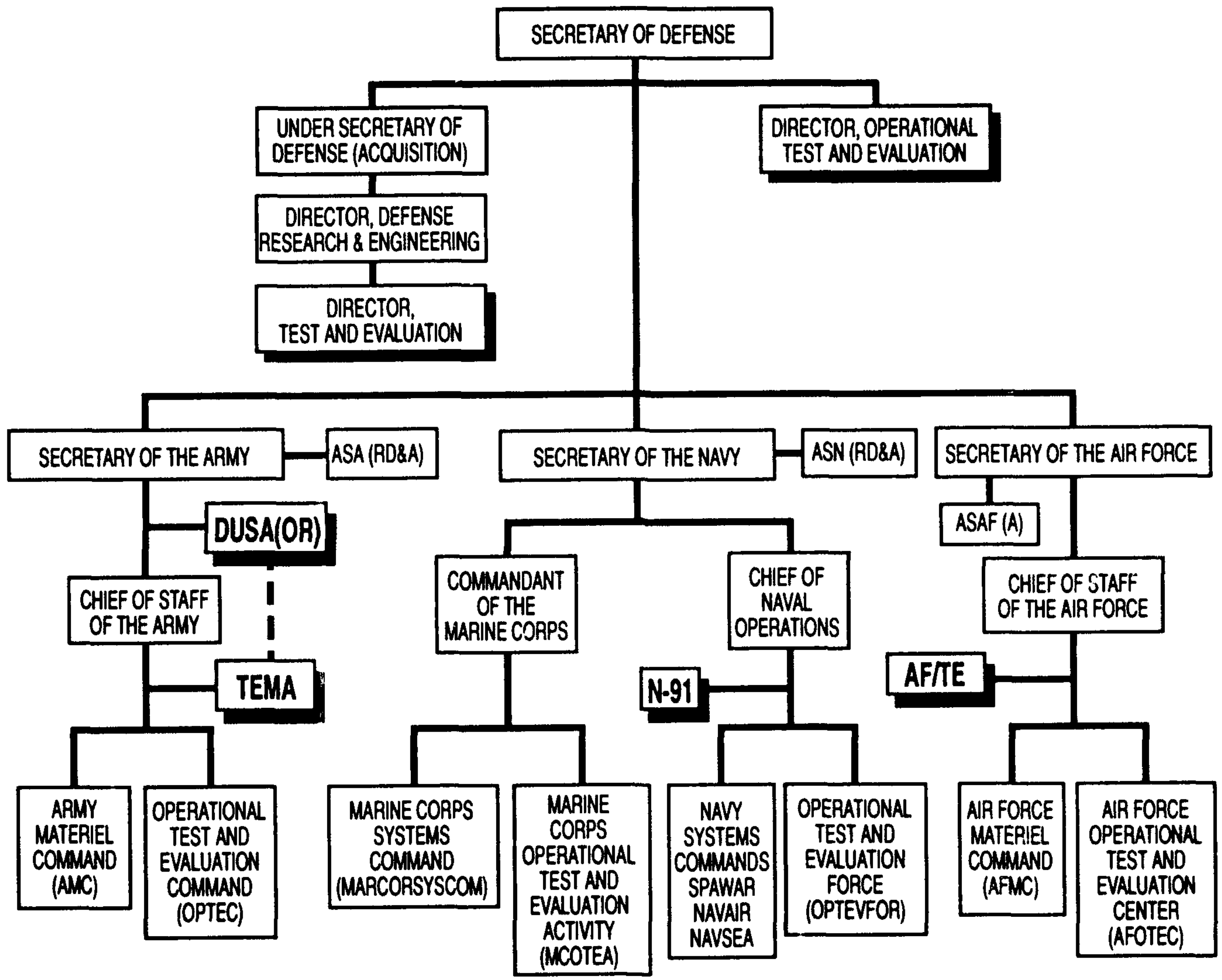


Figure 3-1. DOD Test and Evaluation Organization

tion of all major program test and evaluation master plans (TEMPs). Signs for approval of DT&E portion of TEMPs;

• Reviews major defense acquisition program documentation for DT&E implications and resource requirements to provide comments to the USD(A), DAE or DAB;

• Observes DT&E to ensure adequacy of testing and to assess test results;

• Provides the DAE and DAB with a technical assessment of development T&E conducted on a weapon system;

• Provides advice and makes recommendations to the SECDEF and issues guidance to the component acquisition executives with respect to DT&E;

• Performs the administrative processing of nominations and charters for joint development test programs;

• Provides management oversight of the Major Range and Test Facility Base (MRTFB);

• Administers the Foreign Weapons Evaluation Program and NATO Comparative Test Program;

• Confirms, with advice from the Assistant to the Secretary of Defense (Atomic Energy), that nuclear survivability and hardness objectives have been addressed during DT&E;

• Responsible for the Live Fire Test Program.

3.3.4.2 DTE and Service Reports

During the testing of major and designated weapon systems, the DTE and Services interaction includes the following reporting requirements:

• A TEMP (either initial or updated, as appropriate) must be provided for consideration and approval before each milestone review, starting with Milestone (MS) I.

• A significant T&E Event Report must be provided to the DTE within 24 hours of the test event.

• An End-of-Test Phase Report must be provided to the DTE listing the T&E results, conclusions and recommendations prior to a milestone decision or the final decision to proceed beyond LRIP.

3.3.5 Director Operational Test and Evaluation (DOT&E)

As illustrated in Figure 3-1, the director reports directly to the SECDEF and has special reporting requirements to the Congress. The DOT&E's responsibility to the Congress is to provide an unbiased window of insight into the operational effectiveness and suitability of new weapon systems.

3.3.5.1 Duties and Functions of the DOT&E

The specific duties of DOT&E are outlined in DOD Directive 5141.2. The functions of the office include:

• Obtaining reports, information, advice and assistance as necessary to carry out assigned functions (DOT&E has access to all records and data in DOD on acquisition programs);

• Signing the TEMPs for approval of OT&E and approving the OT&E funding for major systems acquisition;

• Approving test plans on all major systems prior to system starting operational testing (approval in writing required before operational testing may begin);

• Providing observers during preparation and conduct of OT&E;

• Analyzing results of OT&E conducted for each major or designated defense acquisition program and submitting a report to the SECDEF and the Congress on the adequacy of the operational test and evaluation performed;

• A final decision to proceed with a major program beyond LRIP cannot be made until DOT&E has reported (Beyond LRIP Report) to the SECDEF and to congressional Committees on Armed Services and Appropriations on the adequacy of T&E and whether the results confirm the system's operational effectiveness and suitability.

3.3.5.2 DOT&E and Service Interactions

For DOD and DOT&E-designated acquisition programs, the Service provides the DOT&E the following:

• A draft copy of the Operational Test Plan for review;

• Significant Test Plan changes;

• The final Service IOT&E report must be submitted to DOT&E before the DAB Milestone III review.

3.4 SERVICE T&E MANAGEMENT STRUCTURES

3.4.1 Army T&E Organizational Relationship

The Army management structure for T&E is illustrated in Figure 3-2.

3.4.1.1 Army Acquisition Executive

The Under Secretary of the Army is the Army Acquisition Executive (AAE). The AAE is responsible for all acquisition T&E (operational and developmental tests) planning, programming, budgeting, and developmental testing policy and oversight. The AAE performs these duties with the assistance of the Assistant Secretary of the Army, Research, Development, and Acquisition (ASA/RDA). As illustrated in Figure 3-2, the ASA/RDA is organized to provide technical assessments and program evaluations. He resolves acquisition issues whenever possible and recommends acquisition of weapon systems to the AAE. The Deputy Under Secretary of the Army for Operations Research (DUSA(OR)) is chartered to supervise all Army T&E policy and has oversight for all Army T&E.

3.4.1.2 Army Technical Testers and Evaluators

The U.S. Army Materiel Command (AMC) is responsible for the management of DT&E. The Test and Evaluation Command (TECOM) has the *primary responsibility* for conducting technical tests for the Army and under certain conditions Army Materiel Systems Analysis Agency (AMSAA) conducts the evaluation. The TECOM is responsible for:

• Planning, executing and reporting the results of technical tests. Technical tests include development tests, technical feasibility tests, production qualification tests, joint tests and contractor/foreign tests;

• Providing test facilities and technical expertise in support of the T&E life cycle;

• Maintaining the Army's Major Range and Test Facility Base;

• Maintaining the Army's T&E data base;

• Researching, developing and acquiring instrumentation and developing new and improved test methodology;

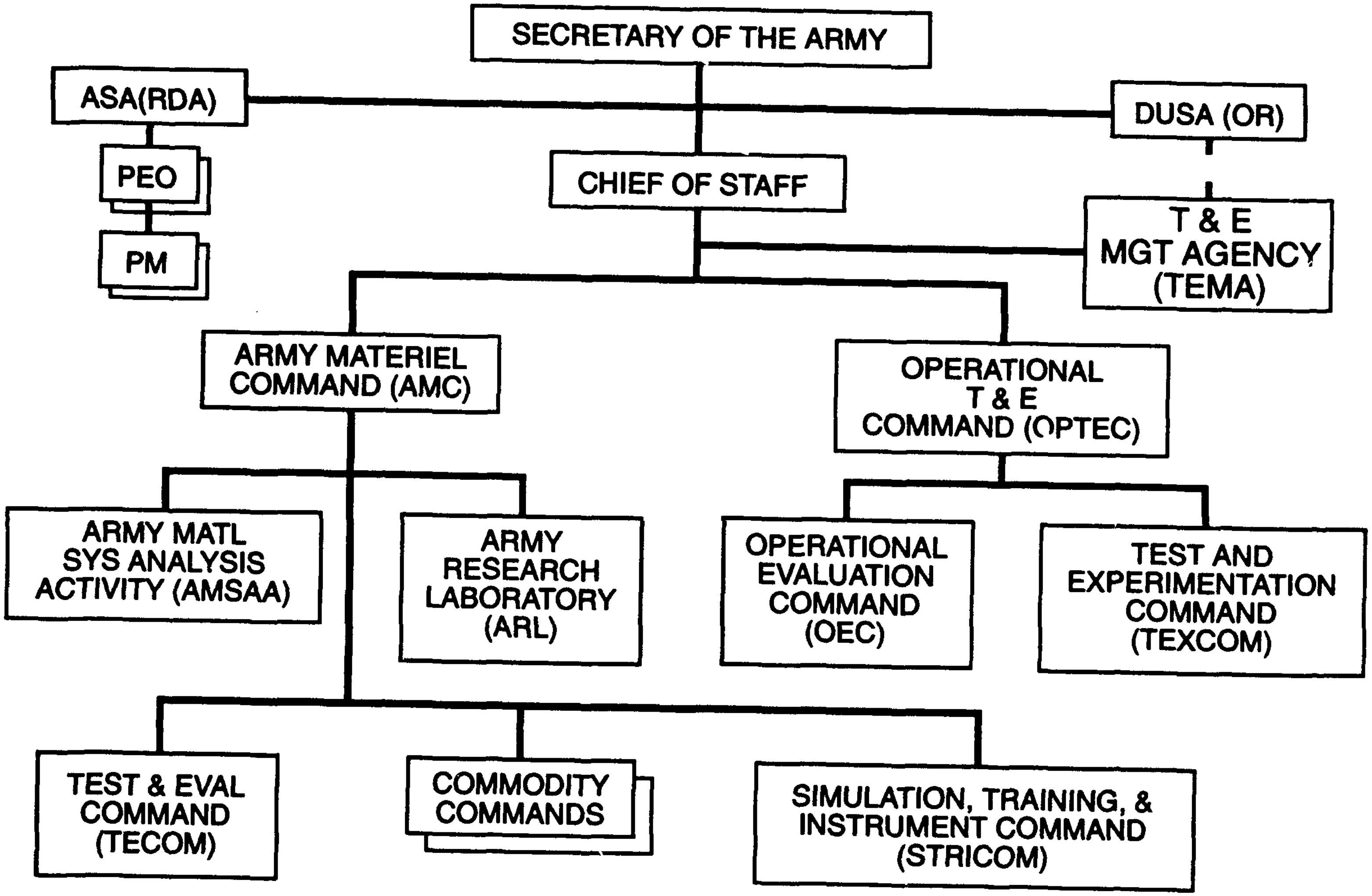


Figure 3-2. Army T&E Organization

- Providing safety confirmations.

3.4.1.3 Army Operational Test and Evaluation Command

- The Army Operational Test and Evaluation Command (OPTEC) is responsible for the management of operational testing as well as the management of joint user testing. The OPTEC is an independent agency reporting directly to the Army Vice Chief of Staff.

- The OPTEC combines the evaluation function performed by the Operational Evaluation Command (OEC) and the operational testing function performed by the Test and Experimentation Command (TEXCOM).

- The U.S Army Forces Command (FORSCOM) supports testing by providing user troops and facilities as needed.

3.4.2 Navy T&E Organizational Relationship

The organizational structure for T&E in the Navy is illustrated in Figure 3-3. Within the Navy Secretariat, the Secretary of the Navy has assigned general and specific research, development, test and evaluation (RDT&E) responsibilities to the Assistant Secretary of the Navy (Research, Development and Acquisition) and to the Chief of Naval Operations (CNO). The CNO has responsibility for ensuring the adequacy of the Navy's overall test and evaluation program. The T&E policy and guidance are exercised through the Directorate of Navy; T&E and Technology Requirements (N-91) staff support is provided by the Test and Evaluation Division (N-912) which has cognizance over planning, conducting and reporting all T&E associated with development of systems.

3.4.2.2 Navy DT&E Organizations

The Navy's senior systems development authority is divided among the commanders of the system commands with NAVAIR developing and performing DT&E on aircraft, NAVSEA developing and performing DT&E on ships and SPAWAR developing and performing DT&E on all other systems. System acquisition is controlled by a chartered program manager or by the commander of a systems command. In both cases, the designated developing agency is responsible for DT&E and for the coordination of all test and evaluation planning in the TEMP. Developing Agencies (DAs) are responsible for:

- Developing test issues based on the thresholds established by the user in the Operational Requirements Document;

- Identifying the testing facilities and resources required to conduct the DT&E;

- Developing the DT&E test reports and quick-look reports.

3.4.2.3 Navy Operational Test and Evaluation Force

The Commander Operational Test and Evaluation Force (COMOPTEVFOR) commands the Navy's independent operational test and evaluation activity and reports directly to the CNO. The functions of the COMOPTEVFOR include:

- Establishing early liaison with the DA to ensure an understanding of the test requirements and plans;

- Reviewing acquisition program documentation to ensure that documents are adequate to support a meaningful T&E program;

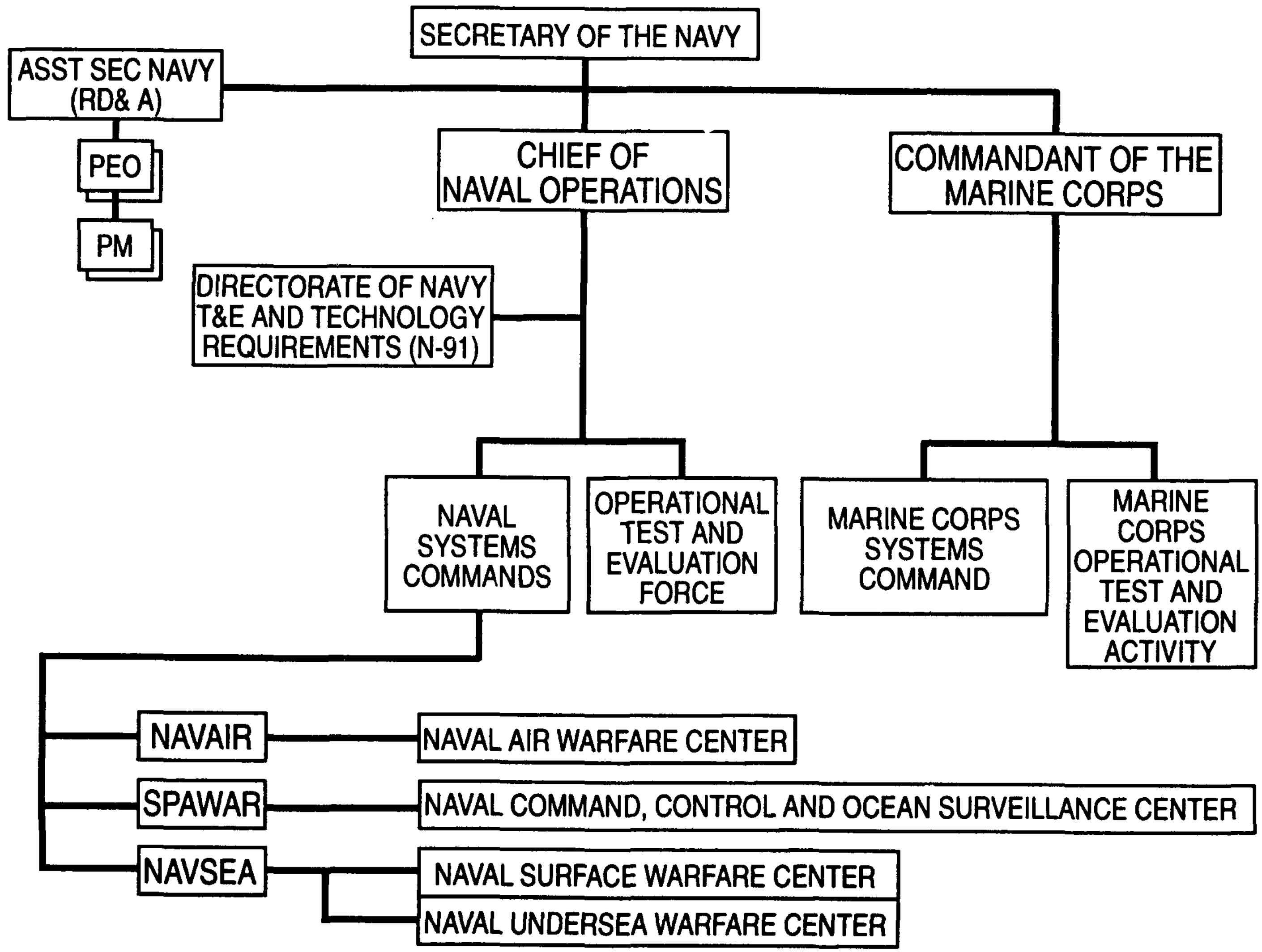


Figure 3-3. Navy T&E Organization

• Planning and conducting realistic OT&E;

• Developing tactics and procedures for the employment of systems that undergo OT&E (as directed by the CNO);

• Providing recommendations to the CNO for the development of new capabilities or the upgrade of ranges;

• Also reporting directly to the CNO, the President of the Board of Inspection and Survey (PRESINSURV) is responsible for conducting acceptance trials of new ships and aircraft acquisitions and is the primary Navy authority for production acceptance T&E of these systems;

• Conducting OT&E on aviation systems in conjunction with Marine Corps Operational Test and Evaluation Activity (MCOTEA).

3.4.3 Air Force Organizational Relationships

3.4.3.1 Air Force Acquisition Executive

The Assistant Secretary of the Air Force for Acquisition (ASAF/A) is the senior-level authority for research, development and acquisition within the Air Force. As illustrated in Figure 3-4, he is an advisor to the Secretary of the Air Force and interfaces directly with the DTE and DOT&E. He receives DT&E and OT&E results as a part of the acquisition decision process. The ASAF/A has, within his structure, a military deputy (acquisition) who is the Air Force primary staff officer with responsibility for R&D and acquisition. He is the chief advocate of Air Force acquisition programs and develops the RDT&E budget. Air Force policy and oversight for T&E is provided by a staff element under the Chief of Staff, Test and Evaluation (AF/TE). They process test documentation for DT&E and OT&E and manage the review of the TEMP.

3.4.3.2 Air Force DT&E Organization

The Air Force Materiel Command (AFMC) is the primary DT&E and acquisition manager. The AFMC performs all levels of research; develops weapons systems, support systems and equipment; and conducts all DT&E. The acquisition program managers are under the Commander, AFMC. Within the AFMC, there are major product divisions, test centers and laboratories as well as missile, aircraft and munitions test ranges.

Once the weapon system is fielded, AMC retains management responsibility for developing and testing system improvements, enhancements or upgrades.

3.4.3.3 Air Force OT&E Organizations

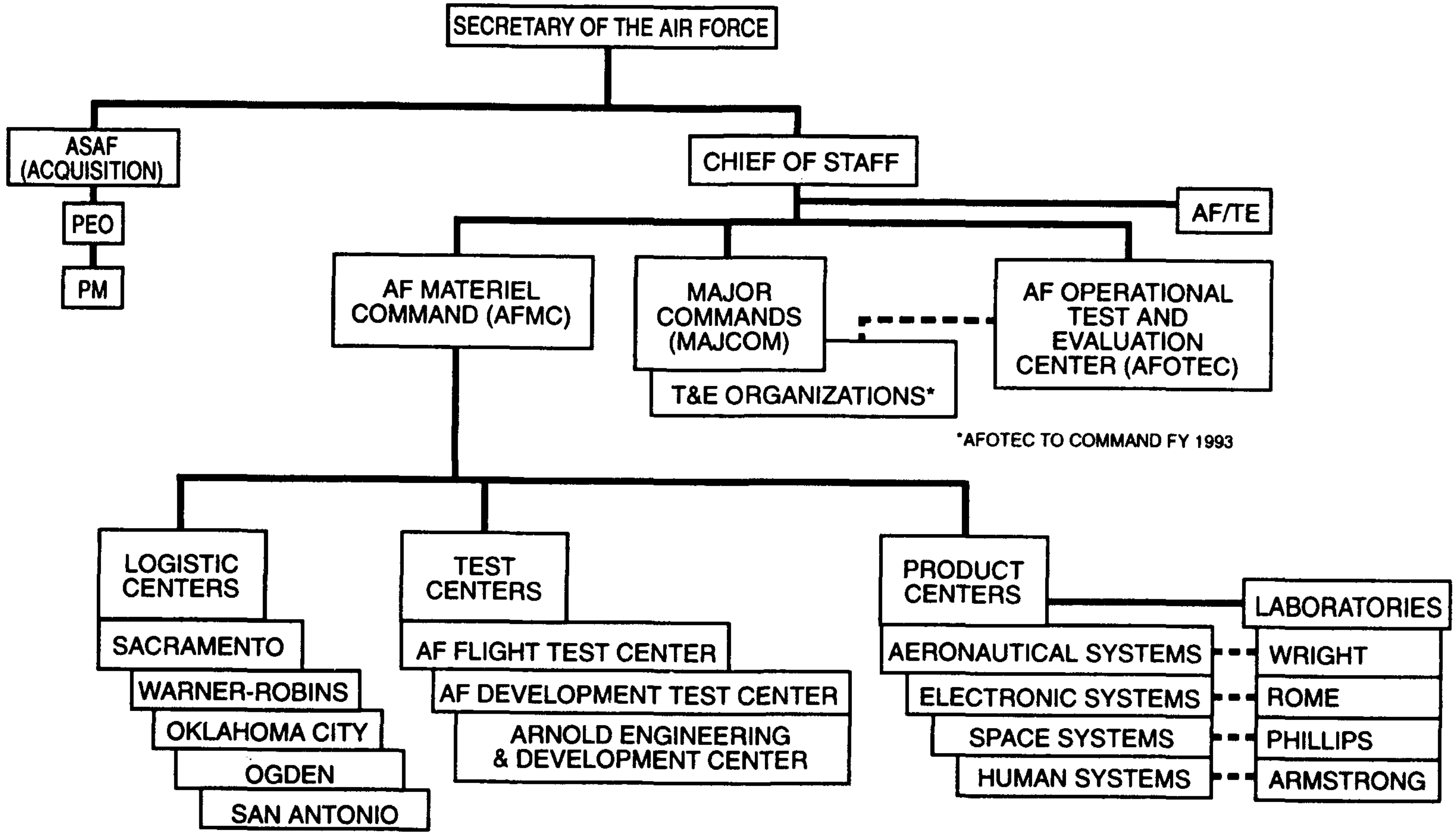
The AF/TE is responsible for supporting and coordinating the OT&E activities of the Air Force Operational Test and Evaluation Center (AFOTEC).

The Commander, Air Force Operational Test and Evaluation Center, is responsible to the Secretary of the Air Force and the Chief of Staff for the independent test and evaluation of all major and nonmajor systems acquisitions. He is supported by the operational commands and others in planning and conducting OT&E.

The AFOTEC develops operational requirements, employment concepts, tactics, maintenance concepts, training requirements and conducts OT&E. The operational commands provide operational concepts, personnel and resources to assist AFOTEC in performing OT&E.

3.4.4 Marine Corps Organizational Relationship

3.4.4.1 Marine Corps Acquisition Executive



*AFOTEC TO COMMAND FY 1993

Figure 3-4. Air Force T&E Organization

The Deputy Chief of Staff for Research and Development (DCS/R&D), Headquarters Marine Corps, directs the total Marine Corps RDT&E effort to support the acquisition of new systems. His position within the General Staff is analogous to that of the Director, T&E, Tech/N-91 in the Navy structure. The DCS/R&D also reports directly to the Assistant Secretary of the Navy/Research, Engineering and Science (ASN/RE&S) in the Navy Secretariat. Figure 3-3, illustrates the Marine Corps organization for T&E management.

3.4.4.2 Marine Corps DT&E Organizations

The Commanding General, Marine Corps Systems Command (CG MCSC), is the Marine Corps materiel developing agent directly interfaces with the Navy Systems Commands. The CG MCSC implements policies, procedures and requirements for DT&E of all systems acquired by the Marine Corps. The Marine Corps also uses DT&E and OT&E performed by other Services, which may develop systems of interest to the Corps.

3.4.4.3 Marine Corps Operational Test and Evaluation Agency

The MCOTEA is the independent OT&E activity maintained by the Marine Corps. Its function is analogous to that performed by OPTEVFOR in the Navy. The CG MCSC provides direct assistance to MCOTEA in the planning, conduct and reporting of OT&E. The Fleet Marine Force performs troop test and evaluation of materiel development in an operational environment.

3.5 SUMMARY

An increased emphasis on test and evaluation has placed greater demands on the OSD and DOD components to carefully structure organizations and resources to ensure maximum effectiveness. Renewed interest by the Congress on testing as a way of assessing systems utility and effectiveness and a recent report by the President's Blue Ribbon Panel on Acquisition Management have resulted in major reorganizations within the Services. These reorganizations will be ongoing for several years to improve the program management and test and evaluation of acquisition systems.

4

PROGRAM OFFICE RESPONSIBILITIES FOR TEST AND EVALUATION

4.1 INTRODUCTION

In government program management offices (PMOs), there should be an element dedicated to management of test and evaluation. This element would have the overall test program responsibility for all phases of the acquisition process. In the PMO, the Deputy for Test and Evaluation (T&E) would be responsible for defining the scope and concept of the test program, establishing the overall program test objectives and managing test program funds and coordination. The Deputy for T&E should provide test directors (such as a joint test director) as required, and coordinate the test resources, facilities and their support required for each phase of testing. In addition, he or a member of his staff, will be responsible for managing the Test and Evaluation Master Plan (TEMP) and planning and managing the special test programs required for the program. The Deputy for T&E will also review, evaluate, approve and release for distribution contractor-prepared test plans and reports and review and coordinate all appropriate government test plans. After the system is produced, he will be responsible for supporting production acceptance testing and the test portions of preplanned product improvement (P^3I) upgrades or enhancements to the weapon system/acquisition. If the program is large enough, the Deputy for T&E will be responsible for all T&E direction and guidance for that program.

4.2 RELATIONSHIP TO THE PROGRAM MANAGER

The program manager (PM) is ultimately responsible for all aspects of the system development, including testing. The Deputy for T&E is normally authorized by the PM to conduct all duties in the area of test and evaluation. The input of the Deputy for T&E to the contract, engineering specifications, budget, program schedule, etc., is essential for the PM to manage the program efficiently.

4.3 EARLY PROGRAM STAGES

In the early stages of the program, the Deputy for T&E writes the test sections of the Request for Proposal (RFP). Although the ultimate responsibility for the RFP is between the PM and the primary contracting officer (PCO), the Deputy for T&E is responsible for creating several sections. These sections include the test schedule, test program funding (projections), test data requirements for the program (test reports, plans, procedures, quick-look reports, etc.), the test section of the Statement of Work (SOW), the Acquisition Plan, Information for Proposal Preparation (IFPP), and (if a joint acquisition program) the Joint Operational Requirements Document (JORD).

4.3.1 Memorandums

Early in the program, another task of the Deputy for T&E is the arrangement of any Memorandums of Agreement or Understanding (MOA/MOU) between Services, NATO countries, test organizations, etc., which outline the responsibilities of each organization. The RFP outlines contractor/government obligations and arrangements on the access and use of test facilities (contractor- and government-owned).

4.3.2 Test Data Management

The Deputy for T&E may have approval authority for all contractor-created test plans, procedures and reports. He must have access to all contractor testing and test results, and he is responsible for disseminating the results to government agencies that need this data. Additionally, the Deputy creates report formats and time lines for contractor submittal, government approval, etc.

The data requirements for the entire test program are outlined in the Contract Data Requirements List (CDRL). The Deputy for T&E should review the Acquisition Management Systems and Data Requirements Control List (AMSDL), DOD 5010.12-L, for relevant test data item descriptions (DIDs). (Examples can be found in Appendix C.) The Deputy for T&E provides input to this section of the RFP early in the program. He ensures that his office and all associated test organizations requiring the information receive the test documentation on time. Usually, the contractor sends the data packages directly to the Deputy for T&E, who, in turn, has a distribution list trimmed to the minimum number of copies for agencies needing that information to perform their mission and oversight responsibilities. It is important for the Deputy for T&E to use an integrated test program and request contractor test plans and procedures well in advance of the actual tests performance to ensure his office has time to approve the procedures or implement modifications. Conversely, he must receive the test results and reports on time to enable him, the PM and higher authorities to make program decisions. Further, the data received should be tailored to provide the minimum information and copies needed. The Deputy for T&E must be aware that data requirements in excess of the minimum needed will lead to an unacceptable increase in overall program cost. For data that is needed quickly and informally (at least initially), the Deputy for T&E can request Quick-Look Reports that give test results immediately after test performance. The Deputy for T&E is also responsible for coordinating with the contractor on all report formats (the in-house contractor format is acceptable in most cases).

4.3.3 Test Schedule Formulation

A very important task the Deputy for T&E has during the creation of the RFP is the test program schedule. Initially, the PM will need contractor predictions of the hardware (and software in some cases) availability dates for models, prototypes, mockups, full-scale models, etc., once the contract is awarded. The Deputy for T&E uses this information to create a realistic front-end schedule of the in-house testing the contractor will conduct before government testing (development testing (DT) and operational testing (OT)). Then, a "strawman" schedule is developed upon which the government DT and OT schedules can be formulated and contractor support requirements determined. The Deputy for T&E can use past experience in testing similar weapon systems/acquisition items or contract test organizations that have the required experience to complete the entire test schedule. Since the test schedule is a critical contractual item, contractor input is very important. The test schedule will normally become an item for negotiation once the RFP is released and the contractor's proposal is received. Attention must be given to ensuring the test schedule is not too success-oriented so test failures will not result in serious program delays for either the government test agencies or the contractor.

4.3.4 Programmatic Environment Analysis

The PMO personnel should be sensitive to the potential environmental consequences of system materials, operations and disposal requirements. Public laws (Title 40, Code of Federal Regulations, Parts 1500-1508; National Environmental Policy Act (NEPA) Regulations; Executive Order 12114, Environmental Effects Abroad of Major Federal Actions; DOD Instruction 5000.2, part 3; etc.) require analysis of hazardous materials and appropriate mitigation measures during each acquisition phase. As stated in DOD Instruction 5000.2, part 6-I, "Emphasis shall be on reduced use of hazardous materials in processes and products rather than simply managing the hazardous waste created."

Litigations resulting in personal fines and imprisonment successfully executed against government employees have raised the environmental awareness at test ranges and facilities. Environmental Impact Statements (supported by long, thorough studies and public testimony) or Environmental Analysis and Assessments (DOD 5000.2-M, 4-F) are generally required before any system testing can be initiated. A summary of the environmental analysis appears in the Integrated Program Summary (IPS) and is updated for each milestone decision point.

4.4 PMO/CONTRACTOR TEST MANAGEMENT

The PMO will, in most cases, have a contractor test section counterpart. With this counterpart, the Deputy for T&E works out the detailed test planning, creation of schedules, etc., for the entire test program. The PMO uses input from all sources (contracts, development test agencies, operational test agencies, higher headquarters, etc.) to formulate the test program's length, scope and necessary details. The Deputy for T&E ensures that the RFP reflects the test program envisioned and the contractor's role in the acquisition. He also ensures the RFP includes provisions for government attendance at contractor's tests and that all contractor test results are provided to the government.

After the RFP has been issued and the contractor has responded, the proposal is reviewed by the PMO. The Deputy for T&E is responsible for performing a technical evaluation on the test portions of the proposal. In this technical evaluation, he compares the proposal to the SOW, test schedule, IFPP, etc., and reviews the contractor's cost of each testing item. This is an iterative process of refining, clarifying and modifying that will ensure the final contract between the PMO and the prime contractor (subcontractors) contains all test-related tasks and is priced within scope of the proposed test program. Once technical agreement on the contractor's technical approach is reached, the Deputy for T&E is responsible for giving inputs to the government contracting officer during contract negotiations. The contracting officer-requested contract deliverables are assigned contract line item numbers (CLINs), which are created by the Deputy for T&E. This will ensure the contractor delivers the required performances at specified intervals during the life of the contract. Usually, there will be separate contracts for development and production of the acquisition item. For each type of contract, the Deputy for T&E has the responsibility to provide the PCO and PM with the test and evaluation input.

4.5 TEST PLANNING WORKING GROUPS

Before the final version of the RFP is created, the Deputy for T&E will form a test planning working group. This group includes the operational test agency, development test agency, organizations that may be jointly acquiring the same system, the test supporting agencies, operational users, and any other organizations that will be involved in the test program by providing test support or by conducting, evaluating or reporting on testing. In later meetings, the contractor participates in this test planning group; however, the contractor may not be selected by the time the first meetings are held.

The purposes of these meetings are to review and assist in the development of early test documentation, the TEMP, and to agree on basic test program schedules, scope, support, etc. The TEMP serves as the top-level test management document for the acquisition program, being updated as the changing program dictates.

4.6 TEST PROGRAM FUNDING/ BUDGETING

The PMO must identify funds for testing very early so that test resources can be obtained. The Deputy for T&E uses the acquisition schedule, TEMP and other program and test documentation to identify test resource requirements. He coordinates these requirements with the contractor and government organizations that have the test facilities to ensure their availability for testing. The Deputy for T&E ensures that test costs include contractor and government test costs. The contractor's test costs are normally outlined adequately in his proposal; however, the government test ranges, instrumentation and test-support resource costs must be determined by other means. Usually, the Deputy for T&E contacts the test organization and outlines his test program requirements (Uniform Document System); and the test organization sends the program office an estimate of the test program costs. He then obtains cost estimates from all test sources he anticipates using and supplies this information to the PM. The Deputy for T&E must also ensure that any program funding reductions are not absorbed entirely by the test program. Some cutbacks may be necessary and allowable; but the test program must supply the PM, other defense decision-making authorities, and the Congress with enough information to make program milestone decisions.

4.7 TECHNICAL REVIEWS, DESIGN REVIEWS AND AUDITS

The role of the Deputy for T&E changes slightly during the contractor's technical reviews, design reviews, physical and functional configuration audits, etc. Usually he plans, directs or monitors government testing; however, in the reviews and audits, he examines the contractor's approach to the test problem and evaluates the validity of the process and the accuracy of the contractor's results. Using his experience and background in test and evaluation, he assesses whether the contractor did enough or too little testing; whether the tests were biased in any way; and if they followed a logical progression using the minimum of time, effort and funds. If the Deputy for T&E finds any discrepancies, he must inform the contractor, the PM and the PCO to validate his conclusions before effecting corrections. Each type of review or audit will have a different focus/orientation, but the Deputy for T&E will always be concerned with the testing process and how it is carried out. After each review, the Deputy for T&E should always document his observations for future reference.

4.8 CONTRACTOR TESTING

The Deputy for T&E is responsible for ensuring that contractor-conducted tests are monitored. He must also be given access to all contractor internal data, test results and test reports related to his acquisition program. Usually, the contract requires that government representatives be informed ahead of time of any (significant or otherwise) testing the contractor conducts so the government can arrange to witness certain testing or receive results of the tests. Further, the contractor's internal data should be available as a contract provision. The Deputy for T&E must ensure that government test personnel (development test and evaluation/operational test and evaluation) have access to contractor test results. It would be desirable to have all testers observe some contractor tests to help develop confidence in the results and identify areas of risk.

4.9 SPECIFICATIONS

Within the program office, the engineering section is usually tasked to create the preliminary specifications for release of the RFP. The contractor is then tasked with creating the specification documentation called out by the contract, which will be delivered once the item/system design is formalized for production. The Deputy for T&E performs an important function in specification formulation by reviewing the specifications to determine if performance parameters are testable; if current, state-of-the-art technology can determine (during the DT&E test phase) if the performance specifications are being met by the acquisition item; or if the specified parameters are too "tight." A specification is too "tight" if the requirements are impossible to meet or demonstrate, or if the specification has no impact on the form, fit or function of the end-item, the system it will become a part of or the system with which it will interact. He must determine if test objectives can be adequately formulated from those specifications that will provide thresholds of performance, minimum and maximum standards and reasonable operating conditions for the end-item's final mission and operating environment. The specifications shape the development test and evaluation (DT&E) testing scenario, test ranges, test support, targets, etc., and are very important to the Deputy for T&E.

4.10 INDEPENDENT EVALUATION AGENCIES

The PMO Deputy for T&E does not have direct control over government-owned test resources, test facilities, test ranges, test personnel, etc. Therefore, he must depend on those test organizations controlling them and stay involved with the test agency activities. The amount of involvement depends on the item being tested; its complexity, cost and characteristics; the length of time for testing; amount of test funds; etc. Usually, the "nuts and bolts" detailed test plans and procedures are written by the test organizations controlling the test resources with input and guidance from the Program Office Deputy for T&E. The Deputy for T&E is responsible for ensuring that the tests are performed using test objectives based on the specifications and that the requirements of timeliness, accuracy and minimal costs are met by the test program design. During the testing, the Deputy for T&E monitors test results. The test agencies submit a copy of their report to the

Program Office at the end of testing, usually to the Office of the Deputy for T&E.

4.11 SUMMARY

Staffing requirements in the PMO vary with the program phase and the T&E workload. Test and evaluation expertise is essential in the early planning stages but can be provided through matrix support. The Deputy for T&E may be subordinate to the chief engineer in early phases but should become a separate staff element after Milestone (MS) II. Changing of critical players can destroy established working relationships and abrogate prior agreements if continuity is not maintained. The PMO management of T&E must provide for an integrated focus and a smooth transition from one staff-support mode to the next.

5

TEST-RELATED DOCUMENTATION

5.1 INTRODUCTION

During the course of a defense acquisition program, many documents are developed that have significance for those responsible for testing and evaluating the system. This chapter is designed to provide background on some of these documents.

As Figure 5-1 shows, test-related documentation spans a broad range of materials. It includes requirements documentation such as the Mission Need Statement (MNS); program decision documentation such as the Integrated Program Summary (IPS) and Acquisition Decision Memorandum (ADM); and program management documentation such as the Acquisition Strategy, Baseline documentation, the System Engineering Management Plan (SEMP), the Integrated Logistics Support Plan (ILSP) and the Test and Evaluation Master Plan (TEMP). Of importance to the PM and to test and evaluation (T&E) managers are additional test program documents such as specific test designs, test plans, outline test plans/test program outlines, evaluation plans and test reports. This chapter concludes with a description of the End-of-Test Phase and beyond Low-Rate Initial Production (LRIP) Reports, two special-purpose T&E status reports that are used to support the milestone decision process.

5.2 REQUIREMENTS DOCUMENTATION

5.2.1 Continuing Mission Area Analyses

As indicated in DODD 5000.1, the Services are required to conduct continuing mission analyses of their assigned areas of responsibility. These Mission Area Analyses (MAA) may result in recommendations to initiate new acquisition programs to reduce or eliminate operational deficiencies. If a need cannot be met through changes in tactics, strategy, doctrine or training and a materiel solution is required, the needed capability is described in a document known as an Operational Requirement Document (ORD). When the cost of a proposed acquisition program is estimated to exceed $200 million for research, development, test and evaluation or $1 billion for procurement (FY 1980$), it is considered a major program and requires an MNS. The MAA is completed at the beginning of a program and reviewed to evaluate system modifications periodically.

5.2.2 Mission Need Statement (MNS)

The MNS is a short, nonsystem-specific statement of operational capability need prepared by any DOD component focusing on a specific mission area need or deficiency. Service validation and, for those potential Acquisition Category (ACAT) I Programs, review and validation by the Joint Requirements Oversight Council (JROC) results in forwarding of the MNS to the milestone (MS) decision authority for MS 0 consideration. The five-page document's content and format, as described in PT 2, DOD 5000.2-M, includes:

- Identification of the applicable Defense Planning Guidance Element;

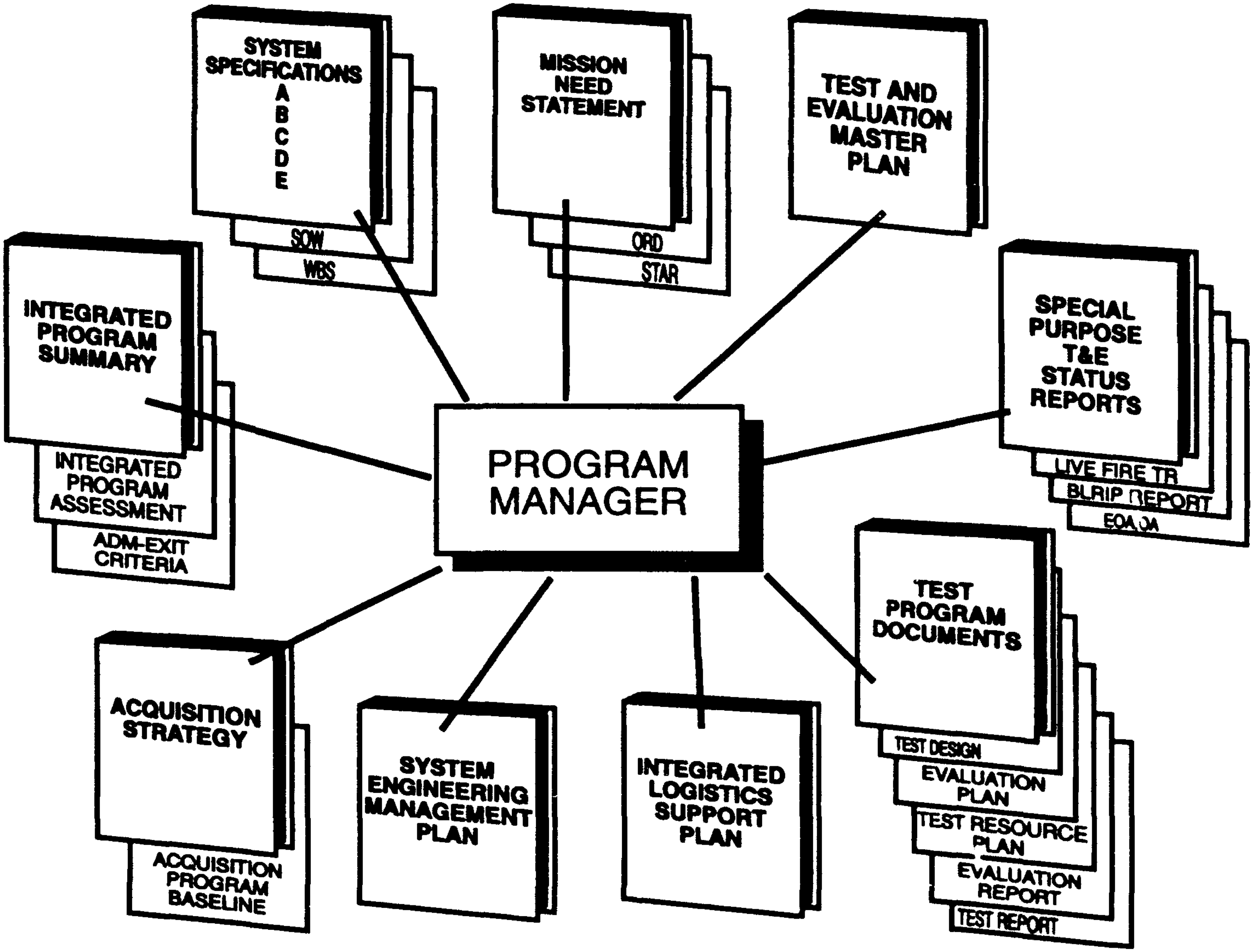


Figure 5-1. Test-Related Documentation

• Mission and threat analyses — need defined in terms of mission, objectives and general capabilities;

• Nonmateriel alternatives— tactics, doctrine, organization and training;

• Potential materiel alternatives — NDI, allied, inter-Service, and new;

• Constraints by infrastructure, treaties and environments.

The MNS and other requirements documents are of particular value to the tester since they form the basis for the initial identification of critical issues that will be addressed in the test program.

5.2.3 Operational Requirements Document (ORD)

The ORD is first prepared for MS I by the user or a user's representative and is approved by the Service Chief or a designated representative. At MS II, the updated ORD should contain "thresholds and objectives for more detailed and refined performance capabilities and characteristics based on the results of trade-off studies and testing conducted during Phase I, Demonstration and Validation." The ORD is a translation of the MNS into user requirements and each concept considered at MS I will have a tailored ORD. Objectives and thresholds for various system performance parameters outlined in the ORD will also be found in baseline documents, the TEMP and program specifications.

5.2.4 System Threat Assessment Report (STAR)

A STAR is prepared by the DOD Component Intelligence Command or Agency, and ACAT ID STARs are validated by the Defense Intelligence Agency. The STAR will contain a "concise description of the projected future operational threat environment, the system-specific threat, the reactive threat that could affect program decisions, and when appropriate, the results of interactive analysis obtained by the Service Program Manager when evaluating the program against the threat." Threat projections start at the initial operating capacity (IOC) and extend over the system's expected operational life. The STAR provides the basis for the test design of threat scenarios and the acquisition of appropriate threat equipment or surrogates. It provides threat data for development test and evaluation (DT&E) and operational test and evaluation (OT&E). Vulnerability and lethality analyses during live fire testing of ACAT I and II systems are contingent on valid threat descriptions. A summary of the STAR are included in the TEMP.

5.3 PROGRAM DECISION DOCUMENTATION

5.3.1 Acquisition Decision Memorandum (ADM)

Under Secretary of Defense for Acquisition (USD(A)) decisions at major acquisition program (ACAT ID) milestones are recorded in a document known as an Acquisition Decision Memorandum (ADM). The ADM documents a USD(A) decision on an MNS at MS 0 and on the IPS at MSs I, II and III. In conjunction with an ADM and its included exit criteria for the next phase, the IPS is a primary program guidance document providing goals/thresholds for systems performance. (PT 13, DODI 5000.2)

5.3.2 Integrated Program Summary (IPS)

The IPS, prepared by the program executive officer (PEO) with PM support, provides a comprehensive summary of pro-

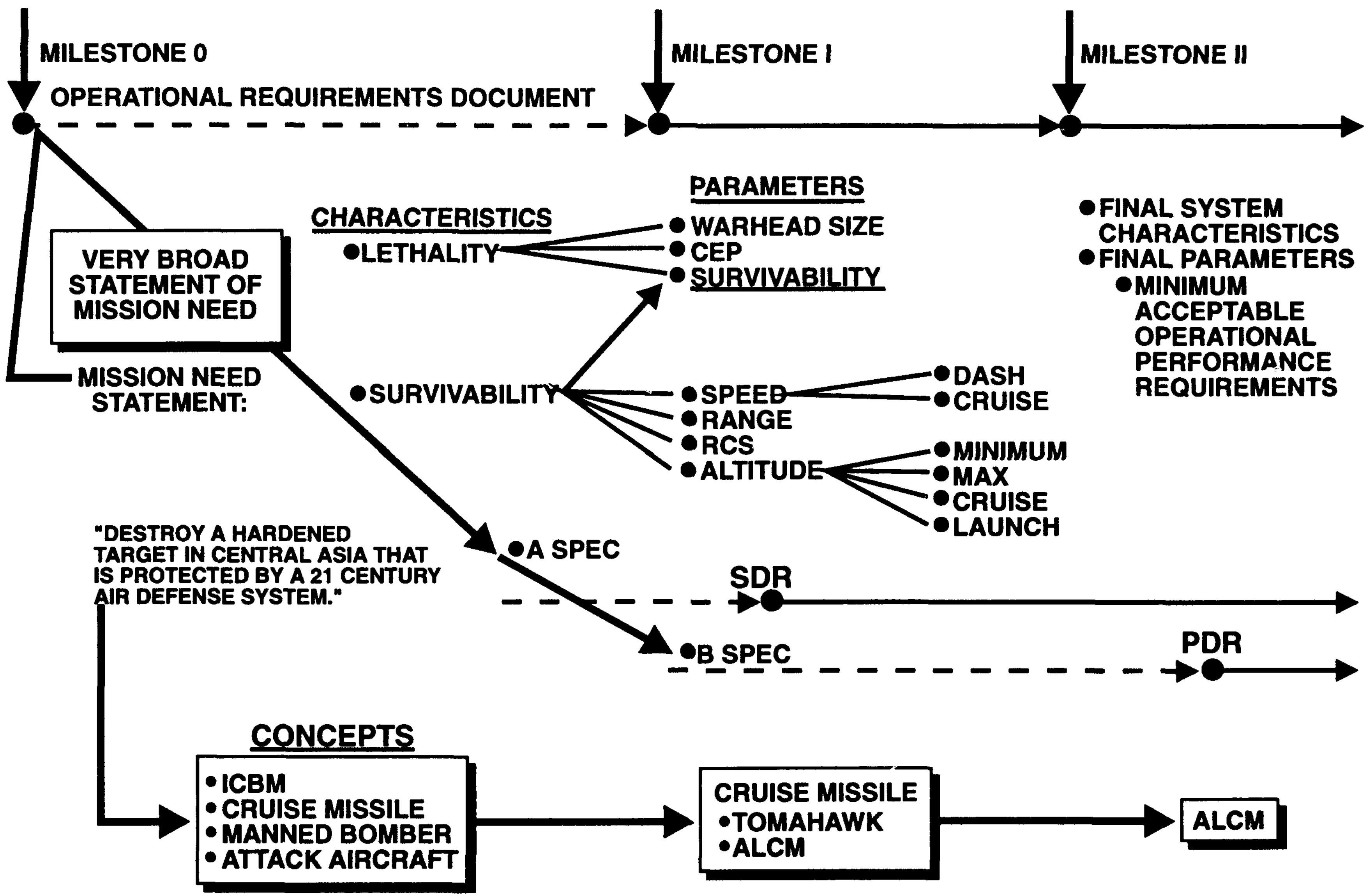


Figure 5-2. Requirements Definition Process

gram structure, status, assessment, plans and recommendations by the PM and the PEO. It addresses current program status, future plans and RISK areas. It provides for the establishment of explicit program cost, schedule and performance objectives, thresholds included in the acquisition program baseline and the next phase's proposed exit criteria. At MS II, the IPS provides information on LRIP quantities (total production, initial operational test and evaluation (IOT&E) assets) along with T&E events leading up to the start of LRIP. The format and content of the IPS are provided in PT 4, DOD 5000.2-M. Test managers will find that IPS information helps in defining program scope and identifies key areas of technological RISK that will influence the test program structure and duration. Test planning documents, especially the TEMP, should be compatible with the annexes to the IPS.

Annex A—Program structure vs. the TEMP integrated program summary, PT II

Annex B—Program life-cycle cost-estimate summary vs. funding profile TEMP, PT II and PT V

Annex C—Acquisition Strategy Report vs. PT III, DT&E and PT IV, OT&E

Annex D—RISK Assessment vs. PT III, DT&E

Annex E—Environmental Analysis vs. PT III, DT&E and PT IV, OT&E.

5.3.3 Cost and Operational Effectiveness Analysis (COEA)

A COEA is normally prepared by a Service agency other than the program management office. The COEA aids decision-makers by examining the relative advantages and disadvantages of program alternatives, providing the rationales for each option. The guidance in DODI 5000.2, part 4-E, was tied more directly to test and evaluation measures of effectiveness (MOE) by a memo issued by the USD(A) in March 1992 and cosigned by the Assistant Secretary of Defense for Program Analysis and Evaluation (ASD(PAE)) and the Director, Operational Test and Evaluation (DOT&E). It stated:

The DOD component in the process of performing a MSI COEA, should identify the MOE's [Measures of Effectiveness] to be used in the COEA and show how these MOE's are derived from the MNS. Each COEA should include MOE's reflecting operational utility that can be tested.

> The TEMP should document how the COEA MOE's and related measures of performance (MOP) will be addressed in test and evaluation.
>
> In particular, the MOE's, MOP's and criteria in the ORD, the COEA, the TEMP and the APB, should be consistent.
>
> In assessing the possible impact of test limitations, the DOD component responsible for the COEA should explain in a quantitative evaluation how and to what extent COEA results would be expected to vary as a result of test limitations.

The driving factor behind this linkage is the decision-maker's reluctance to accept modeling or simulation projections for system performance in the future without actual test data that validates COEA results.

5.4 PROGRAM MANAGEMENT DOCUMENTATION

5.4.1 Acquisition Strategy

An event-based acquisition strategy must be formulated at the start of a development program (MS I). Event-driven acquisition strategy explicitly links program decisions to demonstrated accomplishments in development, testing and initial production. The strategy constitutes a broad set of concepts that provide direction and control for the overall development and production effort. The acquisition strategy is updated at each MS decision point throughout the life of a program. The level of detail reflected in the acquisition strategy can be expected to increase as a program matures. The acquisition strategy serves as a conceptual basis for formulating functional plans such as the System Engineering Management Plan, Integrated Logistics Support Plan, and the Test and Evaluation Master Plan.

It is important that T&E interests be represented as the acquisition strategy is formulated because the acquisition strategy should:

- Provide an overview of the T&E planned for the program, ensuring that adequate T&E is conducted prior to the production decision;

- Discuss plans for providing adequate quantities of test hardware;

- Describe levels of concurrence and combined development test/operational test (DT/OT).

5.4.2 Baseline Documentation

The acquisition program baseline will initially be developed by the PM at MS I (concept) and revised for each subsequent milestone (MS II — Development, MS III — Production). Baseline parameters represent the cost, schedule and performance objectives and thresholds for the system in a production configuration. Each baseline influences the T&E activities in the succeeding phases. Guidance on the formulation of baselines is found in PT 11, DODI 5000.2; and the format is in PT 14, DOD 5000.2-M. A requirement that baselines incorporate the contract specifications applicable to each baseline parameter was issued by the USD(A) in March 1991. Performance demonstrated during T&E of production systems must meet or exceed the thresholds. The thresholds establish deviation limits (actual or anticipated breach triggers reports — PT 19, DOD 5000.2-M) for parameters beyond which the PM may not trade off cost, schedule or performance without authorization by the MS decision authority. Baseline and test documentation must reflect the same expectations for system performance.

5.4.3 Integrated Logistics Support Plan

Integrated logistics support (ILS) is a composite of all support considerations necessary to ensure the effective and economical support of a system at all levels of maintenance for its programmed life cycle (Reference 64). The ILSP describes the overall ILS program and includes ILS requirements, tasks and milestones for the current and succeeding phases of the program. The ILSP serves as the source document for ILS testing requirements.

Standards and procedures for logistic support analysis (LSA) are documented in MIL-STD-1388-1A. This standard requires that T&E programs be planned to serve the following three logistics supportability objectives:

(1) Provide measured data for input into system-level estimates of readiness, operational costs and logistics support resource requirements;

(2) Expose supportability problems so they can be corrected prior to deployment;

(3) Demonstrate contractor compliance with quantitative supportability — related design requirements.

Development of an effective T&E program requires close coordination of efforts among all system engineering disciplines, especially those involved in logistics support analyses. The ILSP should be drafted before MS I to provide a skeletal framework for logistics support analysis, to identify initial logistics testing requirements that can be used as input to the TEMP and to provide test feedback to support ILS development.

5.5 TEST PROGRAM DOCUMENTATION

5.5.1 Test and Evaluation Master Plan

The Test and Evaluation Master Plan (TEMP) is the basic planning document for all T&E related to a DOD system acquisition. It is prepared by the program management office with the operational test information provided by the Service Operational Test Organization. It is used by Office of the Secretary of Defense (OSD) and the Services for planning, reviewing and approving T&E programs and provides the basis and authority for all other detailed T&E planning documents. The TEMP identifies critical technical characteristics and critical operational issues (COI); and it describes the objectives, responsibilities, resources and schedules for all completed and planned T&E. The TEMP is required by DODI 5000.2 (see Chapter 17 for more information regarding the TEMP); guidelines for its preparation are found in DOD 5000.2-M, PT 7.

5.5.2 Evaluation Plan

Evaluation planning is usually included within the test plan. Evaluation planning considers the evaluation and analysis techniques that will be required once the test data has been collected and processed. Evaluation is linked closely to the test design, especially the statistical models on which the test design is built.

The Army requires the development of a TEMP, with one part dedicated to the evaluation being conducted by a technical independent evaluator or an operational independent evaluator.

The objective of the Army's emphasis on evaluation is to "address the issues; describe the evaluation of issues which require data from sources other than test; state the technical or operational issues and criteria; identify data sources; state the approach to the independent evaluation; specify the analytical plan and identify program constraints." (Reference 54)

Evaluation plans are prepared for all systems in development by the operational evaluators during concept exploration and in coordination with the system developer. The Army Master Evaluation Plan becomes an annex to the TEMP and is updated when the TEMP is revised. It identifies each evaluation issue and the methodology to be used to assess it and specifies requirements for exchange of information between the development/operational testers and materiel developers.

5.5.3 Test Design

Test designers need to ensure that the test is constructed to provide useful information in all areas/aspects that will lead to an assessment of the system performance. For example, a complicated, even ingenious, test that does not provide the information required by the decision-makers is, in many respects, a failed endeavor. Therefore, part of the process of developing a test concept

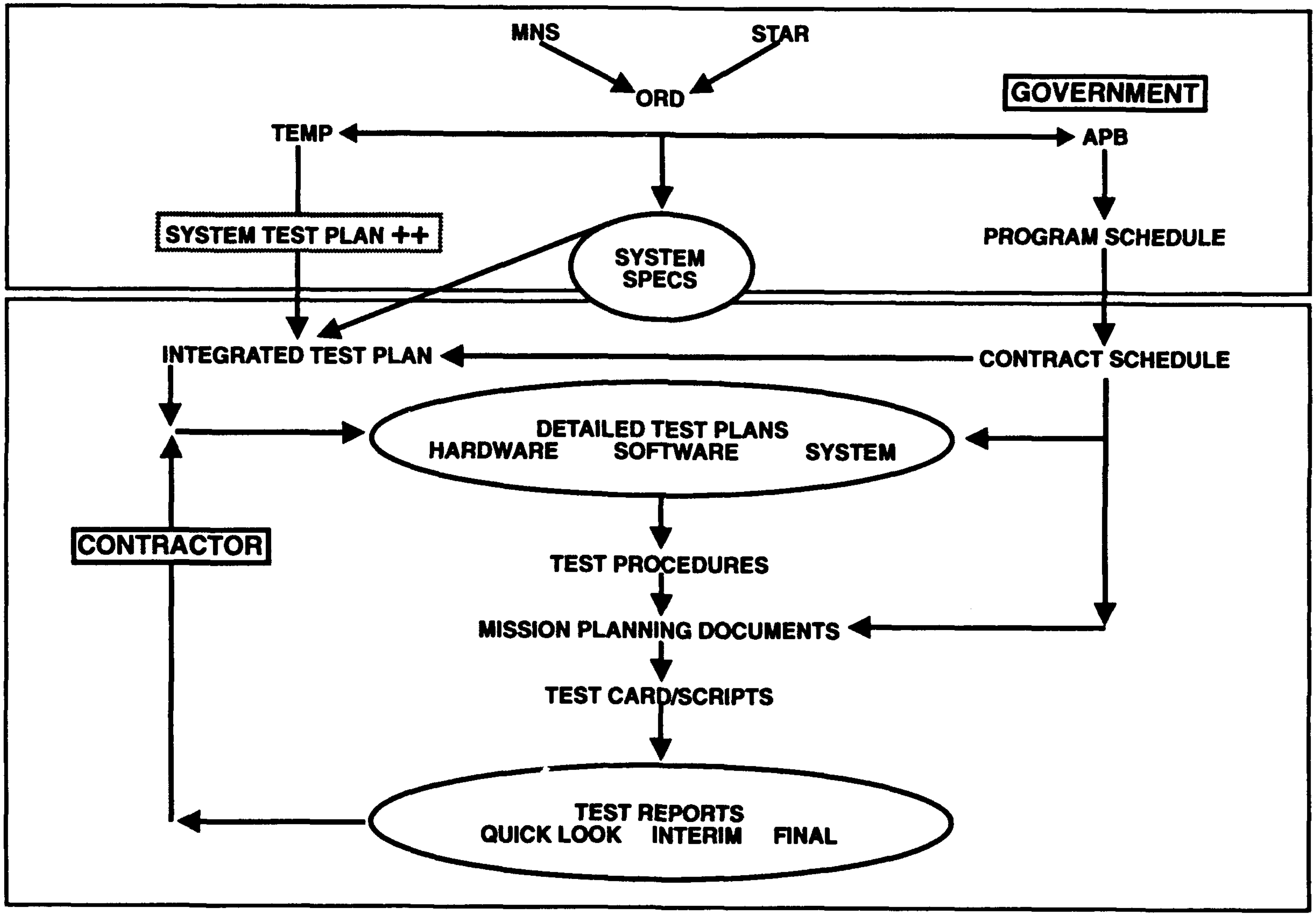


Figure 5-3. Test Program Documentation

or test design (the distinction between these vary from organization to organization) should be to consider whether the test will provide the information required by the decision-makers. In other words, "Are we testing the right things in the right way...and are our evaluations meaningful?"

The test design is statistical and analytical in nature and should perform the following functions:

(1) Structure and organize the approach to testing in terms of specific test objectives;

(2) Identify key measures of effectiveness (MOEs) and measures of performance (MOPs);

(3) Identify the required data and demonstrate how the data will be gathered, stored, analyzed and used to evaluate MOEs;

(4) Indicate what part modeling and simulation will play in meeting test objectives;

(5) Identify the number and type of test events and required resources.

The test design may serve as a foundation for the more-detailed test plan and specifies the test objectives, events, instrumentation, methodology, data requirements, data management needs and analysis requirements.

5.5.4 Test Plan

The test plan is the vehicle that translates a test concept and statistical/analytical test design into concrete resources, procedures and responsibilities. The size and complexity of a test program and its associated test plan are determined by the nature of the system being tested and the type of testing that is to be accomplished. Some major weapons systems may require large numbers of separate tests to satisfy test objectives and, thus, require a multivolume test plan; other testing may be well-defined by a relatively brief test plan. The test plan also provides a description of the equipment configuration and known limitations to the scope of testing. The type of information typically included in a test plan is shown in Table 5-1.

5.5.5 Outline Test Plan/Test Program Outline

The Army's Outline Test Plan (OTP) and Air Force's Test Program Outline (TPO) are essential test planning documents. They are formal resource documents specifying the resources required to support the test. Since the OTP or TPO provides the basis for fiscal programming and coordinating the necessary resources, it is important that these documents be developed in advance and kept current to reflect maturing resource requirements as the test program develops. The Navy makes extensive use of the TEMP to document T&E resource requirements. Each Service has periodic meetings designed to review resource requirements and resolve problems with test support.

5.5.6 Test Reports

5.5.6.1 Quick-Look Reports

Quick-look analyses are expeditious analyses performed during testing using limited amounts of the data base. Such analyses often are used to assist in managing test operations. Quick-look reports are used occasionally to inform higher authorities of test results. Quick-look reports may have associated briefings that present T&E results and substantiate conclusions or recommendations. Quick-look reports may be

Table 5-1. Sample Test Plan Contents

PRELIMINARY PAGES

i. Title page
ii. Abstract
iii. Table of Contents
iv. Terms and Abbreviations
v. Related Documents*

*The actual number of these pages will be determined by the length of preliminary elements (e.g., Table of Contents, Terms and Abbreviations, etc.).

MAIN BODY

1. Introduction
2. Test Purpose and Objectives
3. Concept of Test Operations
4. Method of Accomplishment
5. Test Schedule
6. Test Management and Organization
7. Responsibilities/Support
8. Personnel
9. Required Test Reports
10. Safety
11. Security
12. Information
13. Environmental Protection

ANNEXES

A. Test Design
B. Data Requirements
C. Instrumentation Plan
D. Logistics Support Requirements
E. Reliability and Maintainability Data Plan
F. Intelligence/Threat Information
G-Z. As Required

1, 2, 3, etc., Detailed Test Procedures (Name of Test)

Distribution:

Source: "Standard Procedures for USAF OT&E," July 1974.

generated by the contractor or government agency. They are of particularly critical interest for high-visibility systems that may be experiencing some development difficulties. Techniques and formats should be determined before the start of testing. They may be exercised during pretest trials.

5.5.6.2 Final Test Report

The final test report disseminates the test information to decision authorities, program office staff and the acquisition community. It provides a permanent record of the execution of the test and its results. The final test report should relate the test results to the critical issues and address the objectives stated in the test design and test plan. A final test report may be separated into two sections — a main section providing the essential information about test methods and results, and a second section consisting of supporting appendices to provide details and supplemental information. Generally, the following topics are included in the main body of the report:

(1) Test purpose

(2) Issues and objectives

(3) Method of accomplishment

(4) Results (keyed to the objectives and issues)

(5) Discussion, conclusions and recommendations.

Appendices of the final test report may address the following topics:

(1) Detailed test description

(2) Test environment

(3) Test organization and operation

(4) Instrumentation

(5) Data collection and management

(6) Test data

(7) Data analysis

(8) Modeling and simulation

(9) Reliability, availability and maintainability information

(10) Personnel

(11) Training

(12) Safety

(13) Security

(14) Funding

(15) Asset Disposition.

The final test report may contain an evaluation and analysis of the results, or the evaluation may be issued separately. The analysis tells what the results are, whereas an evaluation tells what the results mean. The evaluation builds on the analysis and generalizes from it, showing how the results apply outside the test arena. It shows what the implications of the test are and may provide recommendations. The evaluation may make use of independent analyses of all or part of the data; it may employ data from other sources and may use modeling and simulation to generalize the results and extrapolate to other conditions. In the case of the Army, a separate Independent Evaluation Report is prepared by technical independent evaluators and operational independent evaluators.

5.6 OTHER TEST-RELATED STATUS REPORTS

5.6.1 End of Test Phase Report

The Services are required by DODI 5000.2 to submit to OSD copies of their formal DT&E and/or OT&E reports that are prepared at the end of each phase of DT&E or OT&E. For major defense acquisition programs/reviews, draft reports must be received by the Defense Acquisition Board (DAB) executive secretary no later than 45 days before a scheduled review.

5.6.2 Low-Rate Initial Production Report

Before an ACAT I or designated program can proceed beyond (MS III) Low-Rate Initial Production (LRIP), the DOT&E must submit a report to the Secretary of Defense and the Senate and House of Representatives Committees on Armed Services and on Appropriations. This report addresses whether the OT&E performed was adequate and whether the OT&E results confirm that the items or components tested are effective and suitable for use in combat by typical military users.

5.7 SUMMARY

A wide range of documentation is available to the test manager and should be used to develop T&E programs that address all relevant issues. The program manager must work to ensure that T&E requirements are *considered at the outset when the acquisition* strategy is formulated. He must also require early, close coordination and a continuing dialogue among those responsible for the SEMP, the ILSP and the TEMP.

6

TYPES OF TEST AND EVALUATION

6.1 INTRODUCTION

This chapter provides a brief introduction to development test and evaluation (DT&E) and operational test and evaluation (OT&E) — two principal types of test and evaluation (T&E); it also discusses the role of qualification testing as a subelement of development testing. Other important types of T&E are introduced. They include: multi-Service testing; joint T&E; live fire testing; nuclear, chemical and biological testing; and nuclear hardening and survivability testing. As Figure 6-1 illustrates, DT&E and OT&E are performed throughout the acquisition process and identified by nomenclature that may change with the phase of the acquisition cycle in which they occur.

6.2 DEVELOPMENT TEST AND EVALUATION (DT&E)

Development test and evaluation is T&E conducted throughout the acquisition process to assist in engineering design and development and to verify that technical performance specifications have been met. The DT&E is planned and monitored by the developing agency and is normally conducted by the contractor. However, the development agency may perform technical compliance tests before OT&E. It includes the T&E of components, subsystems, preplanned product improvement (P^3I) changes, hardware/software integration and production qualification testing. It encompasses the use of models, simulations and test-beds, and prototypes or full-scale engineering development models of the system. Development test and evaluation may involve a wide degree of test complexity, depending upon the type of system or test article under development; e.g., tests of electronic breadboards or brassboards, components, subsystems or experimental prototypes.

Development test and evaluation supports the system design process through a test-analyze-fix-retest approach that involves both contractor and government personnel. Because contractor testing plays a pivotal role in the total test program, it is important the contractor establishes an integrated test plan early to ensure that the scope of the contractor's test program satisfies government and contractor test objectives.

The program manager remains responsible for the ultimate success of the overall program. He and the test specialists on his staff must foster an environment that provides the contractor with sufficient latitude to pursue innovative solutions to technical problems and, at the same time, provides the data needed to make rational trade-off decisions between cost, schedule and performance as the program progresses.

6.2.1 Production Qualification Tests (PQT)

Qualification testing is a form of development testing that verifies the design and

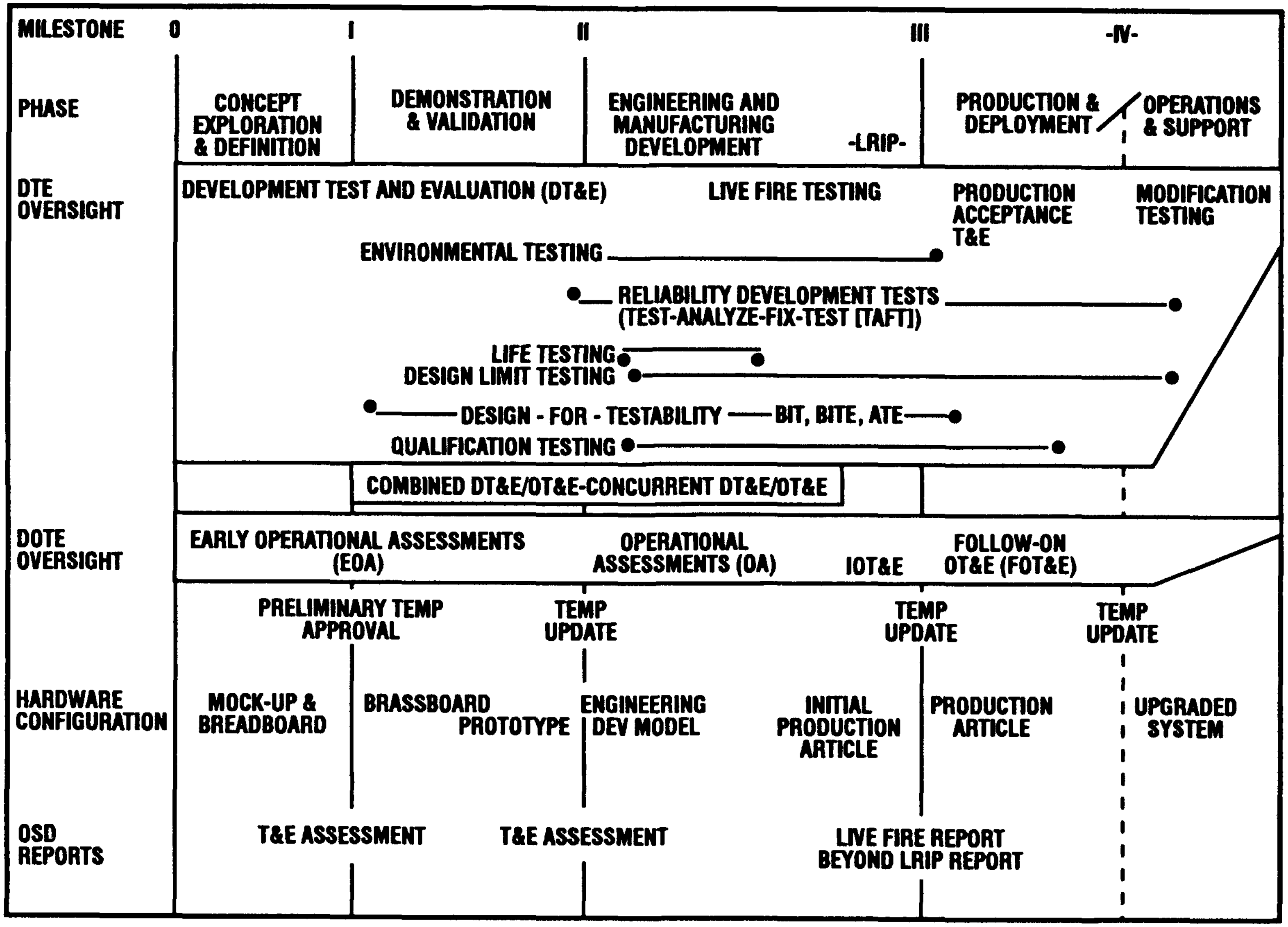


Figure 6-1. Testing During Acquisition

manufacturing process. Production qualification tests are formal contractual tests that confirm the integrity of the system design over the specified operational and environmental range. These tests usually use prototype or preproduction hardware fabricated to the proposed production design specifications and drawings. Such tests include contractual reliability and maintainability demonstration tests required before production release. Production qualification T&E must be completed before Milestone III.

Production qualification tests may be conducted on low-rate initial production items to ensure the effectiveness of the manufacturing process, equipment and procedures. These tests are conducted on each item or a sample lot taken at random from the first production lot and are repeated if the process or design is changed significantly or a second or alternative source is brought on line. These tests are also conducted against contractual design and performance requirements.

6.3 OPERATIONAL TEST AND EVALUATION (OT&E)

6.3.1 The Difference Between Development and Operational Testing

Air Force Manual 55-43, published in June 1979, once contained the following account of the first OT&E; this anecdote serves as an excellent illustration of the difference between development and operational testing:

> The test and evaluation of aircraft and air weapon systems started with the contract awarded to the Wright brothers in 1908. This contract specified a craft which would lift two men with a total weight of 350 pounds, carry enough fuel for a flight of 125 miles, and fly 40 miles per hour in still air. The contract also required that testing be conducted to assure this capability.
>
> What we now call development test and evaluation (DT&E) was satisfied when the Wright brothers (the developer) demonstrated that their airplane could meet those first contract specifications. However, no immediate military mission had been conceived for the Wright Flyer. It was shipped to Fort Sam Houston, Texas, where Captain Benjamin D. Foulois, the pilot, had orders to "teach himself to fly." He had to determine the airplane's performance, how to maintain it, and the kind of organization that would use it. Cavalry wagon masters had to be trained as airplane mechanics, and Captain Foulois was his own instructor pilot.
>
> In the process, Captain Foulois subjected the Wright Flyer to test and evaluation under operational conditions. Foulois soon discovered operational deficiencies. For example, there was no seat on the airplane. During hard landings, Foulois' 130 pound frame usually parted company from the airplane. To correct the problem, Foulois bolted an iron tractor seat to the airplane. The seat helped, but Foulois still toppled from his perch on occasion. As a further improvement, Foulois looped his Sam Browne belt through the seat and strapped himself in. Ever since then, contoured seats and safety belts — a product of this earliest "operational" test and evaluation -have been part of the military airplane.

Captain Foulois' experience may seem humorous now, but it dramatically illustrates the need for operational testing. It also

Table 6-1. Differences Between DT and IOT&E

DT	IOT&E
• Controlled by Program Manager	• Controlled by Independent Agency
• One-on-One Tests	• Many-on-Many Tests
• Controlled Environment	• Realistic/Tactical Environment with Operational Scenario
• Contractor Involvement	• No System Contractor Involvement
• Trained, Experienced Operators	• User Troops Recently Trained on Equipment
• Precise Performance Objectives and Threshold Measurement	• Performance Measurement of Operational Effectiveness and Suitability
• Test to Specification	• Test to Requirements
• Development Test Article	• Production Representative Test Article

shows that operational testing has been going on for a long time.

As shown in Table 6-1 where development testing is focused on meeting detailed technical specifications, the operational test focuses on the actual functioning of the equipment in a realistic combat environment in which the equipment must interact with humans and peripheral equipment. Where DT&E and OT&E are separate activities and are conducted by different test communities, the communities must interact frequently and are generally complementary. The DT&E provides a view of the potential to reach technical objectives, and OT&E provides an assessment of the system's potential to satisfy user requirements.

6.3.2 The Purpose of Operational Test and Evaluation

Operational Test and Evaluation is defined in Title 10, U.S. code:

Definitions of operational effectiveness and operational suitability, outlined in DoDI 5000.2, PT 15 are listed below:

Operational Effectiveness: The overall degree of mission accomplishment of a system when used by representative personnel in the environment planned or expected (e.g. natural, electronic, threat etc.) for operational employment of the system considering organization, doctrine, tactics, survivability, vulnerability, and threat (including countermeasures, initial nuclear weapons effects, nuclear, biological and chemical contamination (NBCC) threats).

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, natural environmental effects and impacts, documentation and training requirements.

In each of the Services, operational testing is conducted under the auspices of an organization that is independent of the development agency, in as operationally realistic environments as possible, with hostile forces representative of the anticipated threat and with typical users operating and maintaining the system. In other words, "OT&E is conducted to ensure that new systems meet the user's requirements, operate satisfactorily, and are supportable under actual field conditions," (Reference 2). The major questions addressed in OT&E are shown in Figure 6-2.

Early Operational Assessment, Operational Assessment (EOA, OA): The operational test normally takes place during the concept exploration/definition and demonstration/validation phases and is used to provide an early assessment of potential operational effectiveness and suitability and to project the system's potential to meet the user's requirements.

6.3.3 Initial Operational Test and Evaluation

The OT&E performed in support of the full-rate production decision (Milestone (MS) III) is known as Initial Operational Test and Evaluation (IOT&E). The IOT&E begins during the Engineering and Manufacturing Development (EMD) Phase and ends before the full-rate production decision. More than one IOT&E may be conducted on the system. The operational test is conducted on a production representative system using typical operational personnel in as realistic a scenario as possible to verify a system's operational effective-

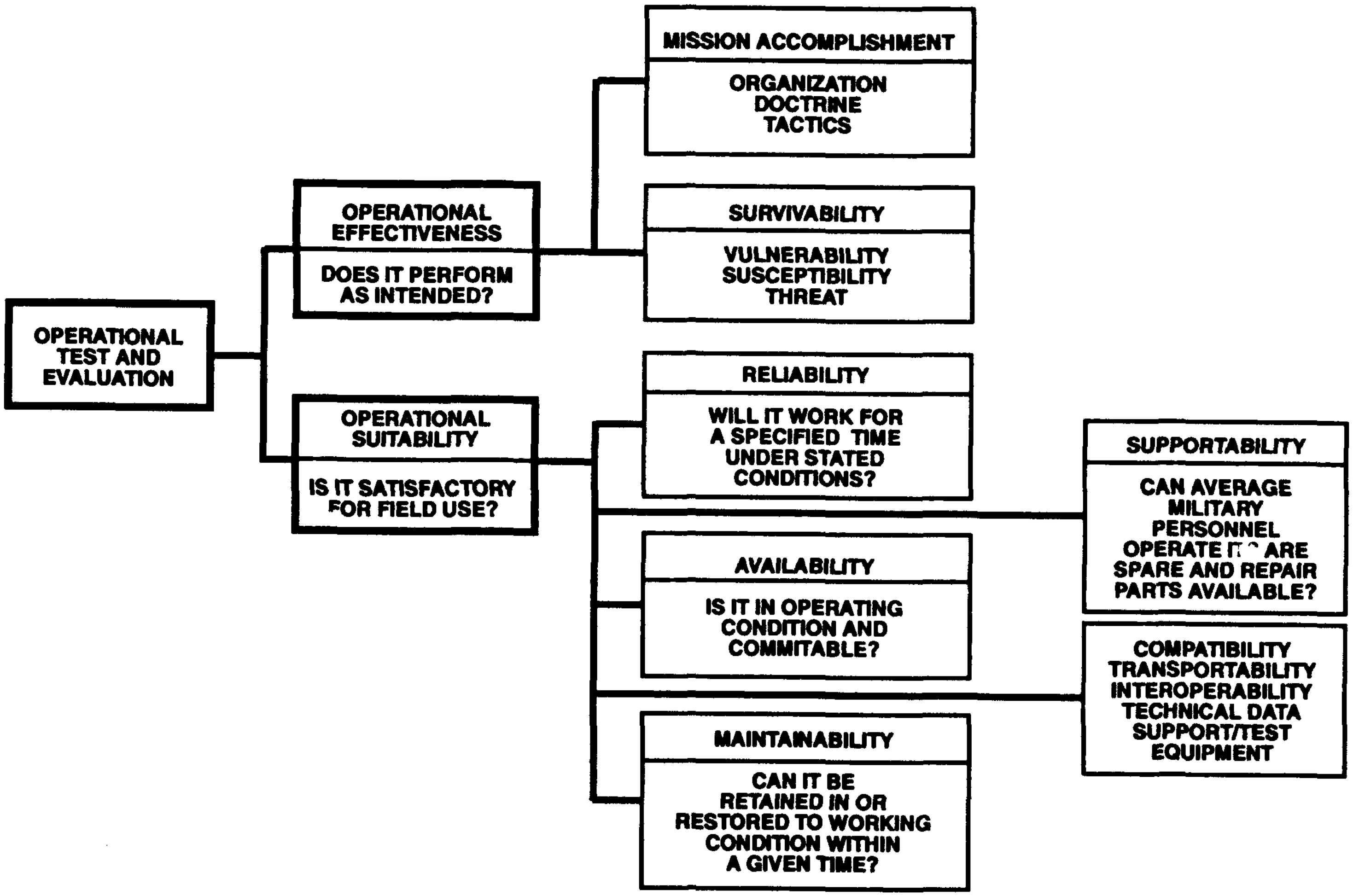


Figure 6-2. Sample Hierarchy of Questions Addressed In Operational Test and Evaluation

ness and suitability and to ensure that the system meets operational thresholds.

6.3.4 Follow-on Operational Test and Evaluation

The OT&E performed after the start of full-rate production may be called follow-on operational test and evaluation (FOT&E) and is conducted during production and deployment. It, too, is sometimes divided into two separate activities. Preliminary FOT&E is normally conducted after the initial operational capability is attained to assess full system capability. It is conducted by the OT&E organization to verify the correction of deficiencies, if required, and to assess system training and logistics status. Subsequent FOT&E is conducted on production items throughout the life of a system. The results are used to refine estimates of operational effectiveness and suitability; to update training, tactics, techniques and doctrine; and to identify operational deficiencies and evaluate modifications. This FOT&E is conducted using the operating command.

6.4 MULTI-SERVICE TEST AND EVALUATION

Multi-Service test and evaluation is T&E conducted on a system being acquired for use by more than one Service. All affected Services and their respective operational test agencies participate in planning, conducting, reporting and evaluating the multi-Service test program. One Service is designated the lead Service and is responsible for the management of the program. The lead Service is charged (by DoDI 5000.2) with the preparation and coordination of a single report that reflects the system's operational effectiveness and suitability for each Service.

The management challenge in a multi-Service test program stems from the fact that the items undergoing testing will not necessarily be used by each of the Services for identical purposes. Differences among the Services usually exist in performance criteria, tactics, doctrine, configuration of armament or electronics and the operating environment. As a result, a deficiency or discrepancy considered disqualifying by one Service is not necessarily disqualifying for all Services. It is incumbent upon the lead Service to establish a discrepancy reporting system that permits each participating Service to document all discrepancies noted. At the conclusion of a multi-Service T&E, each participating OT&E agency prepares an independent evaluation report in its own format and submits that report through its normal Service channels. The lead Service OT&E agency prepares the documentation that goes forward to the Defense Acquisition Board; this documentation is coordinated with all participating OT&E agencies.

6.5 JOINT TEST AND EVALUATION

Joint T&E is not the same as multi-Service T&E. Joint T&E is a specific program activity sponsored and funded by the Office of the Secretary of Defense (OSD). Joint T&E programs are not acquisition oriented; they are a means of examining joint-Service tactics and doctrine. Past joint-test programs have been conducted to provide information required by the Congress, the OSD, the commanders of the Unified and Specified Commands and the Services. Joint tests are usually characterized as either Joint Development T&E or Joint Operational T&E. Joint development T&Es focus on obtaining information on system requirements, system performance, system interoperability, technical concepts, technical improvements, improved testing methodologies or test resource requirements.

Joint operational tests and evaluations are conducted using actual fielded equipment, simulators or surrogate equipment in an exercise or operational environment to obtain data pertinent to operational doctrine, tactics and procedures.

The OSD reviews candidate nominations for joint test programs each year; and, if a proposal is deemed appropriate by the feasibility study, a lead Service is selected and tasked to plan and execute the program using a test force of participating Service personnel.

The commanders of the four-Service operational test agencies — the Army Operational Test and Evaluation Command (OPTEC), the Navy Operational Test and Evaluation Force (OPTEVFOR), the Air Force Operational Test and Evaluation Center (AFOTEC) and the Marine Corps Operational Test and Evaluation Activity (MCOTEA) — have signed a Memorandum of Agreement on Multi-Service OT&E and Joint T&E (Reference 37) that stipulates how both types of programs are to be managed.

6.6 LIVE FIRE TESTING

The Live Fire Test Program was mandated by the Congress in the National Defense Authorization Act for Fiscal 1987 (Public Law 99-661) passed in November 1986. Specifically, this law stipulated that a major defense acquisition program may not proceed beyond low-rate initial production until realistic survivability or (in the case of missiles and munitions) lethality testing has been completed.

In 1984, before the passage of this legislation, the OSD had chartered a joint test program designed to address similar questions relative to systems already in field use. This program, the Joint Live Fire Test, was initially divided into two distinct parts: Armor/Antiarmor and Aircraft. The program's objectives are to:

- Gather empirical data on the vulnerability of existing U.S. systems to Soviet weapons;
- Gather empirical data on the lethality of existing U.S. weapons against Soviet systems;
- Provide insights into the design changes necessary to reduce vulnerabilities and improve lethalities of existing U.S. weapon systems;
- Calibrate current vulnerability and lethality models.

The legislated Live Fire Test (LFT) Program complements the older Joint Live Fire (JLF) Program. While the JLF Program was designed to test systems that were fielded before being completely tested, the spirit and intent of the LFT legislation is to avoid the need to play "catch-up." This program not only requires the Services to test their weapons systems as early as possible against the expected combat threat but also before MS III to identify design characteristics that cause undue combat damage or measure munitions lethality. Remedies for deficiencies can entail required retrofits, production stoppages or other more time-consuming solutions. The essential feature of live fire testing is that appropriate threat munitions are fired against a major U.S. system configured for combat to test its vulnerability and/or that a major U.S. munitions or missile is fired against a threat target configured for combat to test the lethality of the munitions or missile.
Live Fire Test and Evaluation Guidelines were first issued by the Deputy Director, T&E (Live Fire Testing) in May 1987 to supplement DOD Test and Evaluation

Master Plan guidelines (DOD 5000.2-M) in areas pertaining to live fire testing (Reference 34). These guidelines encompass all major defense acquisition programs and define LFT requirements.

6.7 NUCLEAR, BIOLOGICAL AND CHEMICAL WEAPONS TESTING

The testing of nuclear, biological and chemical (NBC) weapons is highly specialized and regulated. Program managers involved in these areas are advised to consult authorities within their chain of command for the specific directives, instructions and regulations that apply to their individual situations. Nuclear weapons tests are divided into categories in which the responsibilities of the Department of Energy (DOE), the Defense Nuclear Agency (DNA) and the military Services are clearly assigned. The DOE is responsible for nuclear warhead technical tests; the DNA is responsible for nuclear weapons effects tests. The Services are responsible for the testing of Service-developed components of nuclear subsystems. All nuclear tests are conducted within the provisions of the Limited Test Ban Treaty that generally restricts nuclear detonations to the underground environment. Nuclear weapons testing requires extensive coordination between Service and DOE test personnel (Reference 55).

Since the United States signed and ratified the Geneva Protocol of 1925, U.S. policy has been never to be the first to use lethal chemical weapons; it may, however, retaliate with chemical weapons if so attacked. With the signing and ratification of the 1972 Biological and Toxin Weapon Convention, the United States formally adopted the position that it would not employ biological or toxin weapons under any circumstances. All such weapons were reported destroyed in the early '70s (Reference 38).

Regarding retaliatory capability against chemical weapons, the Service Secretaries are responsible for ensuring that their organizations establish requirements and determine the military characteristics of chemical deterrent items and chemical defense items. The Army has been designated the DOD executive agent for DOD chemical warfare, research, development and acquisition programs (Reference 39).

United States policy on chemical warfare seeks to:

- Deter the use of chemical warfare weapons by other nations;

- Provide the capability to retaliate if deterrence fails;

- Achieve the early termination of chemical warfare at the lowest possible intensity (Reference 39).

In addition to the customary development tests (conducted to determine if a weapon meets technical specifications) and operational tests (conducted to determine if a weapon will be useful in combat), chemical weapons testing involves two types of chemical tests — chemical mixing and biotoxicity. Chemical-mixing tests are conducted to obtain information on the binary chemical reaction. Biotoxicity tests are performed to assess the potency of the agent generated. Chemical weapons testing, of necessity, relies heavily on the use of nontoxic stimulants, since such substances are more economical and less hazardous and open-air testing of live agents has been restricted since 1969 (Reference 39).

6.8 NUCLEAR HARDNESS AND SURVIVABILITY TESTING

Nuclear hardness is a quantitative description of the physical attributes of a system or component that will allow it to survive in a given nuclear environment. Nuclear survivability is the capability of a system to survive in a nuclear environment and to accomplish a mission. Department of Defense policy requires the incorporation of nuclear hardness and survivability features in the design, acquisition and operation of major and nonmajor systems that must perform critical missions in nuclear conflicts. Nuclear hardness levels must be quantified and validated (Reference 12).

The T&E techniques used to assess nuclear hardness and survivability include: nuclear testing, physical testing in a simulated environment, modeling, simulation and analysis. Although nuclear tests provide a high degree of fidelity and valid results for survivability evaluation, they are not practical for most systems due to cost, long lead times and international treaty constraints. Underground testing is available only on a prioritized basis for critical equipment and components and is subject to a frequently changing test schedule. Physical testing provides an opportunity to observe personnel and equipment in a simulated nuclear environment. Modeling, simulation and analysis are particularly useful in the early stages of development to provide early projections before system hardware is available. These methods are also used to furnish assessments in an area that, because of safety or testing limitations, cannot be directly observed through nuclear or physical testing.

6.9 SUMMARY

Test and evaluation is a technique used to address critical performance questions during system development. These questions may involve several issues including: technical (development testing); effectiveness, suitability and supportability (operational testing); those affecting more than one Service (multi-Service and joint testing); vulnerability and lethality (live fire testing), nuclear survivability; or the use of other than conventional weapons (i.e., nuclear, biological or chemical).

II

MODULE

Developmental Test and Evaluation

Material acquisition is an iterative process of designing, building, testing, identifying deficiencies, fixing, retesting and repeating. Development Test and Evaluation (DT&E) is an important aspect of this process. The DT&E is performed in the factory, laboratory and on the proving ground. It is conducted by subcontractors, as they are developing the components and subassembly; the prime contractor, as he assembles the components and ensures integration of the system; and by the government, to demonstrate how well the weapon system meets its technical and operational requirements. This module describes development testing and the various types of activities it involves. The module also discusses how development testing is used to support the technical review process.

7

INTRODUCTION TO DEVELOPMENT TEST AND EVALUATION

7.1 INTRODUCTION

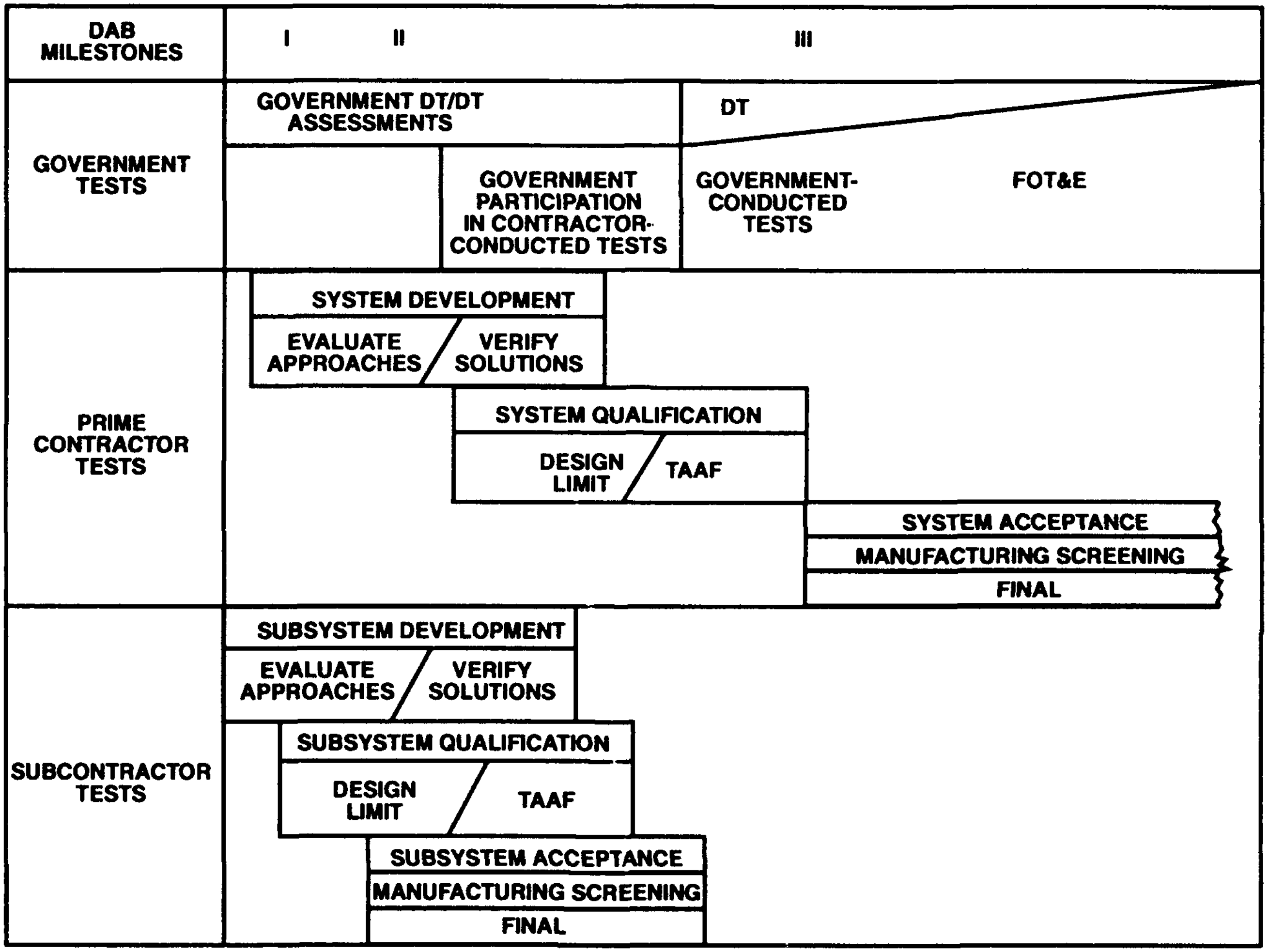
Development test and evaluation (DT&E) is the test and evaluation conducted to demonstrate that the engineering design and development process is complete. It is used by the contractor to reduce risk, validate and qualify the design, and ensure that the product is ready for government acceptance. The DT&E results are evaluated to ensure that design risks have been minimized and the system will meet specifications. The results are also used to estimate the system's military utility when it is introduced into service. Development test and evaluation serves a critical purpose in reducing the risks of development by testing selected high-risk components or subsystems. Finally, DT&E is the government developing agency tool used to confirm that the system performs as technically specified and that the system is ready for field testing. This chapter provides a general discussion of contractor and government DT&E activities, stresses the need for an integrated test program, describes some special-purpose development tests (DTs) and discusses several factors that may influence the extent and scope of the DT&E program.

7.2 DT&E RESPONSIBILITIES

As illustrated in Figure 7-1, the primary participants in testing are the prime contractor, subcontractor, Service materiel developer or developing agency and the operational test and evaluation (OT&E) agency. In some Services, there are also independent evaluation organizations that assist the testing organization in designing and evaluating development tests. As the figure shows, system development testing is performed principally by contractors during the early development stages of the acquisition cycle and by government test/evaluation organizations during the later phases.

Army testing of the Advanced Attack Helicopter illustrates the type of development testing performed by contractors and the relationship of this type of testing to government DT&E activities.

During the contractor competitive Phase I testing of the Army Advanced Attack Helicopter (AAH), prime contractor and subcontractor testing included design support tests, testing of individual components, establishing limited fatigue lives, and bench testing of dynamic components to demonstrate sufficient structural integrity to conduct the Army competitive flight test program. Complete dynamic system testing was conducted utilizing ground test vehicles. Besides supporting the contractor's development effort, these tests provided information for the Army technical review process as the systems, preliminary and critical design reviews were conducted. Following successful completion of the



Source: Adapted and Modified from "Solving the Risk Equation in Transiting from Development to Production," January 19, 1984.

Figure 7-1. Contractor's Integrated Test Requirements ground test vehicle qualification testing, first flights were conducted on the two types of competing helicopters. Each aircraft was being flown 300 hours before delivery of two of each competing aircraft to the Army. The contractor flight testing was oriented toward flight-envelope development, demonstration of structural integrity, and evaluation and verification of aircraft flight handling qualities. Some weapons system testing was conducted during this phase. Government testers used much of the contractor's testing data to develop the test data matrices as part of the government's DT and operational test (OT) planning efforts. The use of contractor test data reduced the testing required by the government and added validity to the systems already tested and data received from other sources.

7.2.1 Contractor Testing

Materiel development, testing and evaluation are an iterative process in which a contractor designs hardware and software, evaluates performance, makes changes as necessary, and retests for performance and technical compliance. Contractor testing plays a primary role in the total test program, and the results of contractor tests are useful to the government evaluator in supporting government test objectives. It is important that government evaluators, as appropriate, oversee contractor system tests and use test data from them to address government testing issues. It is not uncommon for contractor testing to be conducted at government test facilities, since contractors often do not have the required specialized facilities (e.g., for testing hazardous components or for missile flight tests). This enables government evaluators to monitor the tests more readily and increases government confidence in the test results.

Normally, a Request For Proposal (RFP) requires that the winning contractor submit an Engineering Design Test Plan within 60 to 90 days after contract initiation for coordination with government test agencies and approval. This test plan should include testing required by the Statement of Work (SOW), specifications, various MIL-SPECs and MIL-STDs, and testing expected as part of the engineering development and integration process. When approved, the contractor's test program automatically becomes part of the development agency's Integrated Test Plan.

If the contractor has misinterpreted the RFP requirements and the Engineering Design Test Plan does not satisfy government test objectives, the iterative process of amending the contractor's test program begins. This iterative process must be accomplished within limited bounds so the contractor can meet the test objectives without significant effects on contract cost, schedule or scope.

7.2.2 Government Testing

Government testing is performed to: demonstrate how well the materiel system meets its technical compliance requirements, provide data to assess developmental risk for decision-making, verify that the technical and support problems identified in previous testing have been corrected, and ensure that all critical issues to be resolved by testing have been adequately considered. All previous testing, from the contractor's bench testing through development agency testing of representative prototypes, is considered during government evaluation.

Government materiel development organizations include major materiel acquisition commands and, in some cases, operational commands. The materiel acquisition commands have test and evaluation (T&E) organizations that conduct government development testing. In addition to monitoring and participating in contractor testing, these organizations conduct development testing on selected high-concern areas to evaluate the adequacy of systems engineering, design, development and performance to specifications. The program management office (PMO) must be involved in all stages of testing that these organizations perform.

In turn, the materiel development/test and evaluation agencies conduct T&E of the systems in the development stage to ensure they meet technical and operational requirements. These organizations operate government proving grounds, test facilities and labs; and they must be responsive to the needs of the program manager (PM) by providing test facilities, personnel and data acquisition services, as required.

7.2.3 Program Manager's Role

The PM is responsible for coordinating the total T&E program. He performs this task with the assistance of the T&E working group whose members are assembled from development agency, user, technical and operational T&E, logistics, and training organizations. The PM must remain active in all aspects of testing including planning, funding, resourcing, execution and reporting. He plays an important role as the interface between the contractor and the government testing community. Recent emphasis on early T&E highlights a need for early government tester involvement in contractor testing. For example, during development of the AAH test, it was found that having program management personnel on the test sites improved test continuity, facilitated the flow of spare and repair parts, provided a method of monitoring contractor performance, and kept the Service headquarters informed with timely status reports.

7.3 TEST PROGRAM INTEGRATION

During the development of a weapon system, there are a number of tests conducted by subcontractors, the prime contractor and the government. To ensure these tests are properly time-phased, that adequate resources are available, and to minimize unnecessary testing, a coordinated test program must be developed and followed. The Test and Evaluation Master Plan (TEMP) normally does not provide a sufficient level of detail concerning contractor or subcontractor tests. A contractor or PMO Integrated Test Plan must also be developed to describe these tests.

7.3.1 Integrated Test Plan

The Integrated Test Plan (ITP) is used to record the individual test plans for the subcontractor, prime contractor and government. The prime contractor should be contractually responsible for the preparation and updating of the ITP, and the contractor and Service-developing agency should ensure that it remains current. The ITP includes all developmental tests that will be performed by the prime contractor and the subcontractors at both the system and subsystem levels. It is a detailed, working-level document that assists in identifying risk as well as duplicative or missing test activities. A well-maintained ITP facilitates the most efficient use of test resources.

7.3.2 Single Integrated Test Policy

Most Services have adopted a single integrated contractor/government test policy, thereby reducing much of the government testing requirements. This policy stresses independent government evaluation and permits an evaluator to monitor contractor and government test programs and evaluate the system from an independent perspective. The policy stresses the use of all available test data for system evaluation.

7.4 AREAS OF DT&E FOCUS

7.4.1 Life Testing

Life testing is performed to assess the effects of long-term exposure to various portions of the anticipated environment. These tests are used to ensure the system will not fail prematurely due to metal fatigue, component aging or other problems caused by long-term exposure to environmental effects. It is important that the requirements for life testing are identified early and integrated into the system test plan. Life tests are time-consuming and costly; therefore, life testing requirements and life characteristics must be carefully analyzed concurrent with the initial test design. Aging failure data must be collected early and analyzed throughout the testing cycle. If life characteristics are ignored until results of the test are available, extensive redesign and project delays may be required. Accelerated life testing techniques are available and may be used whenever applicable.

7.4.2 Design Evaluation/ Verification Testing

Design evaluation and verification testing is conducted by the contractor and/or the development agency with the primary objective of influencing system design. Design evaluation is fully integrated into the development test cycle; and its purposes are to:

(1) Determine if critical system technical characteristics are achievable;

(2) Provide data for refining and making the hardware more rugged to comply with technical specification requirements;

(3) Eliminate as many technical and design risks as possible or to determine the extent to which they are manageable;

(4) Provide for evolution of design and verification of the adequacy of design changes;

(5) Provide information in support of development efforts;

(6) Ensure components, subsystems and systems are adequately developed before beginning operational tests.

7.4.3 Design Limit Testing

Design limit tests are integrated into the test program to ensure the system will provide adequate performance when operated at outer performance limits and when exposed to environmental conditions expected at the extreme of the operating envelope. The tests are based on mission profile data. Care must be taken to ensure all systems and subsystems are exposed to the worst-case environments, with adjustments made because of stress amplification factors and cooling problems. Care must also be taken to ensure that the system is not operated beyond the specified design limits; for example, an aircraft component may have to be tested at temperature extremes from an Arctic environment to a desert environment.

7.4.4 Reliability Development Testing (RDT)

Reliability development testing (RDT) or reliability growth testing (RGT) is a planned test, analyze, fix and test (TAFT) process in which development items are tested under actual or simulated mission-profile environments to disclose design deficiencies and to provide engineering information on failure modes and mechanisms. The purpose of RDT is to provide a basis for early incorporation of corrective actions and verification of their effectiveness in improving the reliability of equipment. Reliability development testing is conducted under controlled conditions with simulated operational mission and environmental profiles to determine design and manufacturing process weaknesses. The RDT process emphasizes reliability growth rather than a numerical measurement. Reliability growth during RDT is the result of an iterative design process because, as the failures occur, the problems are identified, solutions proposed, the redesign is accomplished, and the RDT continues. A substantial reliability growth TAFT testing effort was conducted on the F-18 DT&E for selected avionics and mechanical systems. Although the TAFT effort added $100 million to the Research, Development, Test and Evaluation (RDT&E) Program, it is estimated that many times that amount will be saved through lower operational and maintenance costs throughout the system's life.

7.4.5 Reliability, Availability and Maintainability (RAM)

Reliability, availability and maintainability (RAM) requirements are assessed during all contractor and government testing. The data are collected from each test event and placed in a RAM data base, which is managed by the development agency. Contractor and government development tests provide a measure of the system's common RAM performance against stated specifications in a controlled environment. The primary emphasis of RAM data collection during the DT is to provide an assessment of the system RAM parameters growth and a basis for assessing the consequences of any differences anticipated during field operations.

7.5 SYSTEM DESIGN FOR TESTING

Built-in test (BIT), built-in-test equipment (BITE) and automated test equipment (ATE) are major areas that must be considered from the start of the design effort. Design for testing (Figure 7-2) addresses the need to: (1) collect data during the development process concerning particular performance characteristics; (2) enable efficient and economical production by providing ready access to, and measurement of, appropriate acceptance parameters; and (3) enable rapid and accurate assessment of the status of the product to the lowest repairable element when deployed. Many hardware systems have testing circuits designed and built-in. This early planning by design engineers allows easy testing for fault isolation of circuits, both in system development phases and during operational testing and deployment. There are computer chips in which more than half of the circuits are for test/circuit check functions. This type of circuit design requires early planning by the PM to ensure the RFP requirements include the requirement for designed/built-in test capability. Evaluation of these BIT/BITE/ATE systems must be included in the test program.

7.6 IMPACT OF WARRANTIES ON T&E

A warranty or guarantee is a commitment provided by a supplier to deliver a product that meets specified standards for a specified time. With a properly structured warranty, the contractor must meet technical and operational requirements. If the product should fail during that warranty period, the contractor must replace or make repairs at no additional cost to the government. The Defense Appropriations Act of 1984 requires warranties or guarantees on all weapon systems procurement. This act makes warranties a standard item on most

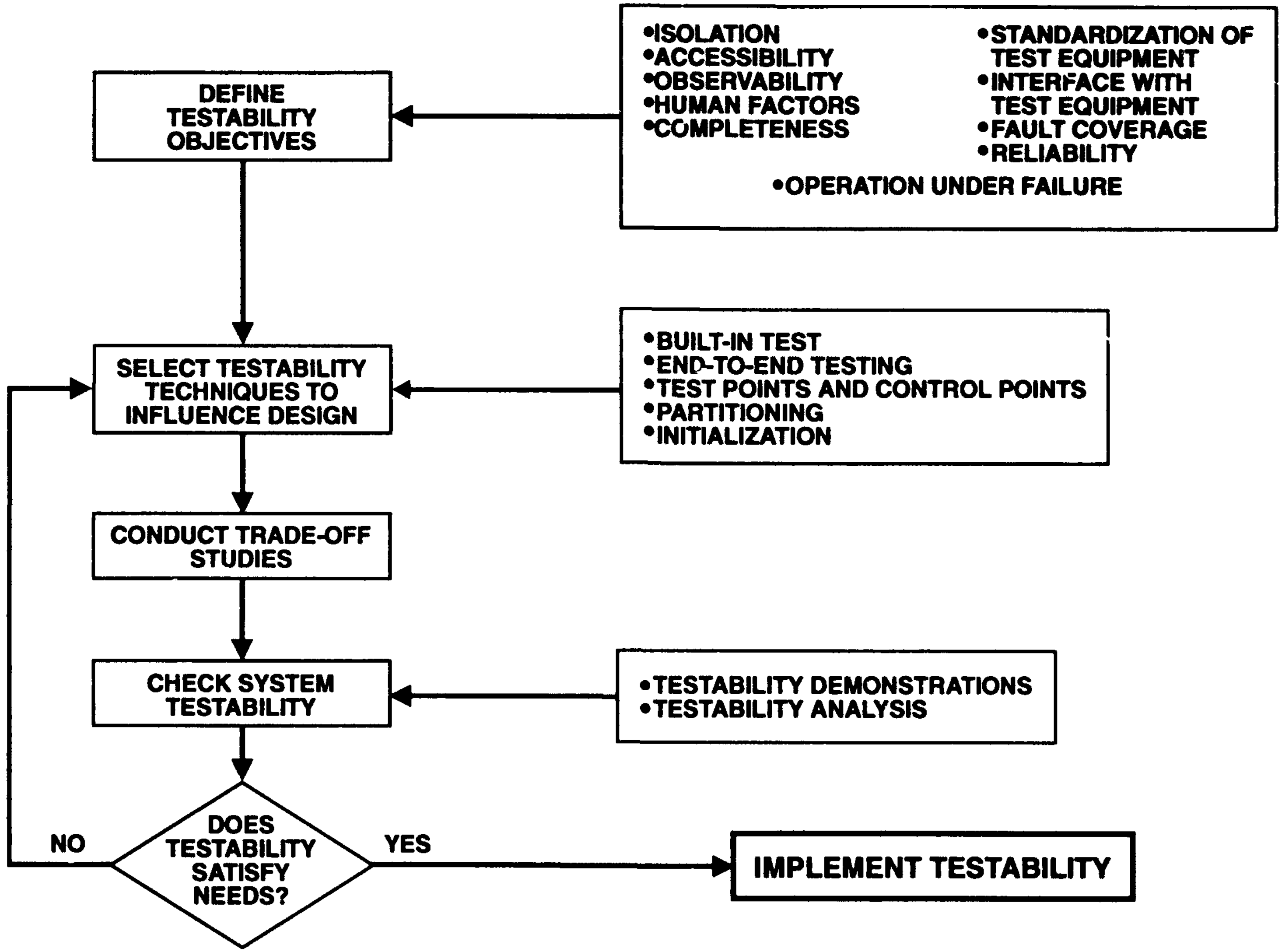


Figure 7-2. Design For Testing: Procedures

fixed-price production contracts. Incentives are the main thrust of warranties; and, as prescribed in MIL-STD-781, the government will perform a reliability demonstration test on the system to determine these incentives. Although warranties have favorable advantages to the government during the early years of the contract, warranties do not affect the types of testing performed to ensure the system meets technical specifications and satisfies operational effectiveness and suitability requirements. Warranties do, however, affect the amount of testing required to establish reliability. Because the standard item is warranted, less emphasis on that portion of the item can allow for additional emphasis on other aspects of the item not covered under the warranty. Further, the government may tend to have more confidence in contractor test results and may be able, therefore, to avoid some duplication of test effort. The warranty essentially shifts the burden of risk from the government to the contractor. Warranties can significantly increase the price of the contract, especially if high-cost components are involved.

7.7 DT&E OF LIMITED PROCUREMENT QUANTITY PROGRAMS

Programs that involve the procurement of relatively few items, typically over an extended period, are normally subjected to standard DT&E. Occasionally, a unique test approach that deviates from the standard timing and reporting schedule will be used. The DT&E principle of components, subsystems, prototypes and first-production models of the system is normally applied to limited procurement. It is important that DT&E and OT&E organizations work together to ensure that T&E plans are integrated into the overall acquisition strategy.

7.8 SUMMARY

Development test and evaluation is an iterative process of designing, building, testing, identifying deficiencies, fixing, retesting and repeating. It is performed in the factory, laboratory and on the proving ground by the contractors and the government. Contractor and government testing is combined into one integrated test program and conducted to determine if the technical development of the acquisition process have been met and to provide data to the decision authority.

8

DT&E SUPPORT OF TECHNICAL REVIEWS AND MILESTONE DECISIONS

8.1 INTRODUCTION

Throughout the acquisition process, development test and evaluation (DT&E) is oriented toward the demonstration of specifications showing the completeness and adequacy of systems engineering, design, development and performance. A critical purpose of DT&E is to identify the risks of development by testing and evaluating selected high-risk components or subsystems. Development test and evaluation is the developer's tool to show that the system performs as specified or that deficiencies have been corrected and the system is ready for operational testing and fielding. The DT&E results are used throughout the system engineering process to provide valuable data in support of formal design reviews. This chapter describes the types of development testing used throughout the system acquisition cycle to support the materiel acquisition process. It also describes the objectives of the various tests conducted during the DT&E process and discusses their relationship to the formal design reviews essential to the systems engineering process.

8.2 DT&E AND THE SYSTEM ACQUISITION CYCLE

As illustrated in Figure 8-1, DT&E is conducted throughout the system life cycle. Development test and evaluation may begin before program initiation (Milestone 0) with the evaluation of evolving technology, and it continues after the system is in the field.

8.2.1 DT&E Prior to Program Initiation

Prior to program initiation, technology feasibility testing is conducted to confirm that the technology considered for the proposed weapon development is the most advanced available and that it is technically feasible.

8.2.2 DT&E During the Concept Exploration and Definition Phase

Development testing that takes place during the Concept Exploration and Definition Phase is conducted by a contractor or the government to assist in selecting preferred alternative system concepts, technologies and designs. The testing conducted depends on the state of development of the test article's design. Government test evaluators participate in this testing because information obtained can be used to support the Systems Requirements Review. The information obtained from these tests may also be used to support a program start decision by the Services or the Office of the Secretary of Defense (OSD).

8.2.3 DT&E During the Demonstration and Validation Phase

Development testing conducted during the Demonstration and Validation Phase is

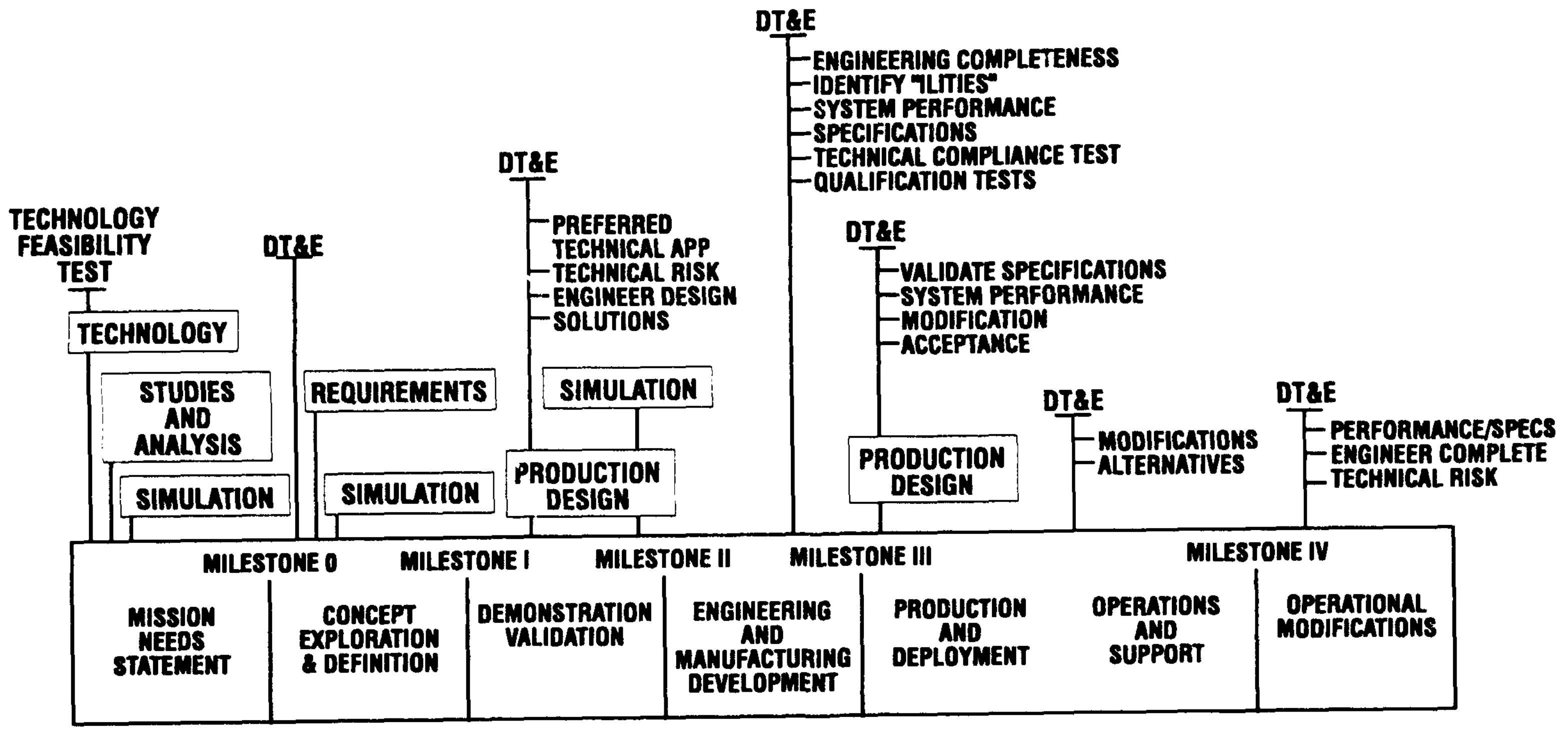


Figure 8-1. Relationship of DT&E to the Acquisition Process

used to demonstrate that: technical risk areas have been identified and can be reduced to acceptable levels; the best technical approach can be identified; and, from this point on, engineering efforts will be required rather than experimental efforts. It supports the Milestone II decision that considers entry into Engineering and Manufacturing Development (EMD) and, as appropriate, low-rate initial production. This DT&E includes contractor/government integrated testing, engineering design testing and advanced development verification testing.

Development testing during this period is most often conducted at the contractor's facility. It is conducted on components, subsystems, brassboard configurations or advanced development prototypes to evaluate the potential application of technology and related design approaches before EMD. Component interface problems and equipment performance capabilities are evaluated. The use of properly validated analysis, modeling and simulation is encouraged, especially during the early phases to assess those areas that, for safety or testing capability limitations, cannot be observed directly through testing. Models can provide early projections of systems performance, effectiveness and suitability and can reduce testing costs. This test and evaluation also may include initial environmental assessments.

Army testing of the Advanced Attack Helicopter (AAH) provides an example of the type of activities that occur during development tests (DTs). The early DT&E of the AAH was conducted by the Army Engineering Flight Activity. The test was conducted in conjunction with an Early Operational Test, and candidate designs were flown more than 90 hours to evaluate flight handling qualities and aircraft performance. This test also included the firing of the 30 millimeter cannon and the 2.75-inch rockets. Reliability, availability and maintainability (RAM) data were obtained throughout the test program; and these data, along with RAM data provided from early contractor testing, became a part of the system's RAM data base. After evaluating the results, the Army selected a contractor to proceed with the next development phase of the AAH.

8.2.4 DT&E During the Engineering and Manufacturing Development Phase

Development test and evaluation conducted during the EMD Phase provides the final technical data for determining a system's readiness to transition into either low-rate initial production (LRIP) or full-rate production. It is conducted using advanced engineering development models and is characterized by engineering and scientific approaches under controlled conditions. The qualification testing provides quantitative and qualitative data for use in the system's evaluation. The evaluation results are used by the development community and are also provided to Service and OSD decision authorities. These tests measure technical performance including: effectiveness, reliability, availability, maintainability, compatibility, interoperability, safety and supportability. They include tests of human engineering and technical aspects of the system. Demonstrations of whether engineering is reasonably complete and if solutions to all significant design problems are in hand are also included. Development test and evaluation may be conducted on LRIP articles as a prelude certifying the system ready for initial operational test and evaluation (IOT&E). The Navy calls this DT&E for certification TECHEVAL (technical evaluation).

As an example of testing done during the EMD Phase, the Army AAH was flown in a series of engineering design tests (EDTs). The EDT-1, -2 and -4 were flown at the contractor's facility. (The EDT-3 requirement was deleted during program restructuring.) The objectives of these flight tests were to evaluate the handling characteristics of the aircraft, check significant performance parameters and confirm the correction of deficiencies noted during earlier testing. The EDT-5 was conducted at an Army test facility, Yuma Proving Ground. The objectives of this test were the same as earlier EDTs; however, the testers were required to ensure that all discrepancies were resolved before the aircraft entered operational testing. During the EDTs, operational test personnel were completing operational test design, bringing together test resources and observing the DT&E tests. Additionally, operational test (OT) personnel were compiling test data, such as the system contractor's test results, from other sources. The evolving DT results and contractor data were made available to the Critical Design Review members to ensure that each configuration item design was essentially completed. The Army conducted a physical Configuration Audit to provide a technical examination to verify that each item "as built" conformed to the technical documentation defining that item.

8.2.5 DT&E During the Production and Deployment Phase

Development testing may be conducted after the LRIP Phase and after the full-rate production decision is made at Milestone III. This testing is normally tailored to verify correction of identified design problems and demonstrate the system modification's readiness for production. This testing is conducted under controlled conditions and provides quantitative and qualitative data. This testing is conducted on production items delivered from either the pilot or initial production runs. To ensure that the items are produced according to contract specification, quantity production processes are used. This testing determines whether the system has successfully transitioned from engineering development prototype to production and whether it meets design specifications.

8.2.6 DT&E During Operations and Support

The development testing, which occurs soon after the initial operating capability or initial deployment, assesses the deployed system's operational readiness and supportability. It ensures that all deficiencies noted during previous testing have been corrected, evaluates proposed product improvements and block upgrades, and ensures that integrated logistics support is complete. It also evaluates the resources on hand and determines if the plans to ensure operational phase readiness and support objectives are sufficient to maintain the system for the remainder of its acquisition life cycle. Near the end of the system's life, DT&E is performed to assist in modifying the system to help meet new threats, add new technologies, or aid in disposal.

Once a system approaches the end of its usefulness, the development testing conducted is concerned with the monitoring of a system's current state of operational effectiveness, suitability and readiness to determine whether major upgrades are necessary or deficiencies warrant consideration of a new system replacement. Tests are normally conducted by the operational testing community; however, the DT&E community may be required to assess the technical aspects of the system.

8.3 DT&E AND THE REVIEW PROCESS

8.3.1 The Technical Review Process

Technical reviews and audits are conducted by the government and the contractor as part of the system engineering process to ensure the design meets the system, subsystem and software specifications. Each review is unique in its timing and orientation. Some reviews build on previous reviews and take the design and testing effort one step closer to the final system design to satisfy the operational concept/purpose for the weapon system. Table 8-1 illustrates the sequencing of the technical reviews in relation to the test and evaluation phases.

The review process was established to ensure that the system under development would meet government requirements. The reviews evaluate data from contractor and government testing, engineering analysis, and models to determine if the system or its components will eventually meet all functional and physical specifications and to determine the final system design. The system specification is very important in this process. It is the document used as a bench mark to compare contractor progress in designing and developing the desired product. The requirements and direction for these formal technical reviews and audits are set forth in MIL-STD-1521B.

8.3.2 Testing in Support of Technical Reviews

The testing community must be continually involved in supporting the technical reviews of their systems. Decisions made at these reviews have major impacts on the system test design, resources required to test, and the development of the Test and Evaluation Master Plan (TEMP) and other documentation. A more detailed discussion of testing to support the technical reviews is provided in the Systems Engineering Management Guide (Reference 45). The reviews focus primarily on government technical specifications for the system. Figure 8-2 illustrates the program specifications and how they are developed in the system life cycle.

8.3.3 Formal Reviews

8.3.3.1 Systems Requirements Review (SRR)

The SRR is normally conducted during the System Concept Exploration and Definition Phase or early Demonstration and Validation Phase. It consists of a review of the system/system segment specifications, also known as the "A" specifications (System Functional Block Diagram, Reference 45, Chapter 12), and is conducted after the accomplishment of functional analysis and preliminary requirements allocation. During this review, the systems engineering management activity and its output are reviewed for responsiveness to the Statement of Work requirements. The primary function of the SRR is to ensure that systems requirements have been completed and properly identified and that there is a mutual understanding between the contractor and the government. During the review, the contractor describes his progress and any problems in risk identification and ranking, risk avoidance and reduction, trade-off analysis, producibility and manufacturing considerations, and hazards considerations. The results of integrated test planning are reviewed to ensure the adequacy of planning to assess the design and to identify risks. Issues of testability of requirements should be discussed.

8.3.3.2 Systems Design Review (SDR)

The SDR is conducted as a final review before submittal of the Demonstration and Validation Phase products or as the initial EMD Phase review. The "A" specification is validated to ensure that the most current specification is included in the System Functional Baseline and that they are adequate

Table 8-1. Technical Reviews and Audits

		WHEN	PURPOSE	DOCUMENTATION/ DATA
System Requirements Review	SRR	Late C/E	- Evaluate System Functional Requirements	- Prelim "A" Spec - Prelim Planning Documentation - FFBD, RAS, MBN Analysis
System Design Review	SDR	Late Dem Val	- Evaluate System Design - Validate "A" Spec - Establish System Level Functional Baseline	- "A" Spec - Prelim "B" Spec - Design Documents - RAS, SSD, TLS
Software Specification Review	SSR	Early EMD Prior SW PDR	- Evaluate SW Performance Requirements - Validate "B-6" Specs - Establish SW Specs Baseline	- "B-5" Spec - (SRS & IRS) - Ops Concept Doc
Preliminary Design Review	PDR	Early EMD	- Validate "B" Specs - Establish HW Allocated Baseline - Evaluate Preliminary Design HW & SW	- "B" Spec - Des Doc Test Plan - ICD, Engr Drawings - Preliminary SDD - IDD
Critical Design Review	CDR	Early/Mid EMD	- Evaluate CI Design - Determine Readiness For Fabrication	- Prelim C, D, E Specs - Detail Design Documents Include SDD - IDD
Test Readiness Review	TRR	Mid/Late EMD	- Approve SW Test Procedures - Determine Readiness For Formal Test	- SW Test Plan/ Procedures - Informal SW Test Results
Functional Configuration Audit	FCA	Late EMD	- Verify CI Actual Performance Complies With Hardware Development Or SRS & IRS	- Test Plans & Descriptions - Software Test Reports
Formal Qualification Review	FQR	Late EMD	- Verify CI's Perform in System Environment	- Test Reports - A, B, C Specs - O & S Docs
Production Readiness Review	PRR	Incrementally EMD	- Assess Risk For Production Go-Ahead	- Prod Planning Documents
Physical Configuration Audit	PCA	Late EMD Early PROD	- Format Examination of The As-Built	- Final Spec "C" - Listings - Lvl II&III Drawing

SPECIFICATION	WHEN PREPARED	PREPARING AGENT	APPROVING AGENT	CONTENT	BASELINE
SYSTEM/ SEGMENT (TYPE A)	CE PHASE	DEV/PROG MGR INDUSTRY	DEV/PROG MGR USER	DEFINES MISSION AND TECH REQMTS; ALLOCATES REQUIREMENTS TO FUNCTIONAL AREAS; DOCUMENTS DESIGN CONSTRAINTS; DEFINES INTERFACES	FUNCTIONAL
DEVELOPMENT (TYPE B, PART 1, DESIGN-TO)	LATE D/V PHASE	INDUSTRY	DEV/PROG MGR	DETAILS DESIGN REQUIREMENTS; STATES, DESCRIBES PERFORMANCE CHARACTERISTICS OF EACH CI; DIFFERENTIATES REQUIREMENTS ACCORDING TO COMPLEXITY AND DISCIPLINE SETS	ALLOCATED
PRODUCT (TYPE C, PART II, BUILD-TO)	LATE EMD PHASE	INDUSTRY	DEV/PROG MGR	DEFINES FORM, FIT, FUNCTION, PERFORMANCE AND TEST REQUIREMENTS FOR ACCEPTANCE	PRODUCT
PROCESS (TYPE D)	EMD/PROD PHASES	INDUSTRY	DEV/PROG MGR	DEFINES PROCESS PERFORMED DURING FABRICATION	
MATERIAL (TYPE E)	EMD/PROD PHASES	INDUSTRY	DEV/PROG MGR	DEFINES PRODUCTION OF RAW OR SEMI FABRICATED MATERIAL USED IN FABRICATION	

Source: Systems Engineering Management Guide

Figure 8-2. Specifications Summary

and cost-effective to satisfy validated mission requirements. The SDR encompasses the total system requirement of operations, maintenance, test, training, computers, facilities, personnel and logistics considerations. A technical understanding should be reached on the validity and the degree of completeness of specifications, design, operational concept documentation, software requirements specifications and interface requirements specifications during this review.

8.3.3.3 Software Specification Review (SSR)

The SSR is a formal review of the computer system configuration item (CSCI) requirements, normally held after a SDR but before the start of a CSCI preliminary design. Its purpose is to validate the allocated baseline for preliminary CSCI design by demonstrating to the government the adequacy of the software requirements specifications, interface requirements specifications and operational concept documentation.

8.3.3.4 Preliminary Design Review (PDR)

The PDR is a formal technical review of the basic approach for a configuration item. It is conducted at the configuration item and system level early in the EMD Phase to confirm that the preliminary design logically follows SDR findings and meets the system requirements. The review results in an approval to begin the detailed design. The draft Type B specifications are reviewed during the PDR. The purpose of the PDR is to: evaluate the progress, technical adequacy and risk resolution (on technical, cost and schedule basis) of the configuration item design approach; conduct DT and OT activities to measure the performance of each configuration item (CI); and establish the existence and compatibility of the physical and functional interface among the CI and other equipment.

8.3.3.5 Critical Design Review (CDR)

The CDR may be conducted on each configuration item and/or at the system level. It is conducted during the EMD Phase when the detailed design is essentially complete and prior to the Functional Configuration Audit. During the CDR, the overall technical program risks associated with each configuration item are also reviewed on a technical, cost and schedule basis. It includes a review of the "C" specifications and the status of both the system's hardware and software. Input from qualification testing should assist in determination of readiness for design freeze and LRIP.

8.3.4.6 Test Readiness Review (TRR)

The TRR is a formal review of the contractor's readiness to begin CSCI testing. A government witness will observe the system demonstration to verify that the system is ready to proceed with CSCI testing. It is conducted after the software test procedures are available and computer software components testing is complete. The purpose of the TRR is for the program management office (PMO) to determine whether the contractor is ready to begin CSCI testing.

8.3.4.7 Functional Configuration Audit (FCA)

The FCA is a formal review to verify that the configuration item's performance complied with its development specification. The "B" specifications are derived from the system requirements and baseline documentation. During the FCA, all relevant test data is reviewed to verify that the item has performed as required by its functional and/or allocated configuration identification. The audit is conducted on the item representative (prototype or production) of the configuration to be released for production. The audit consists of a review of the contractor's test procedures and results. Information provided will be used by the FCA to determine the status of planned tests.

8.3.3.9 Physical Configuration Audit (PCA)

The PCA is a formal review which establishes the product baseline as reflected in an early production configuration item. It is the examination of the as-built version of hardware and software configuration items against its technical documentation. The PCA also determines that the acceptance testing requirements prescribed by the documentation are adequate for acceptance of production units of a CI by quality assurance activities. It includes a detailed audit of engineering drawings, final Part II product specifications, technical data and plans for testing that will be utilized during production. The PCA is performed on all first articles and on the first CIs delivered by a new contractor.

8.3.3.9 Formal Qualification Review (FQR)

The FQR is a systems-level configuration audit that may be conducted after system testing is completed. The objective is to verify that the actual performance of the CI, as determined through testing, complies with its Type B specifications and to document the results of the qualification tests. The FQR and FCA are often performed at the same time; however, if sufficient test results are not available at the FCA to ensure the CI will perform in its operational environment, the FQR can be scheduled for a later time.

8.3.3.10 Production Readiness Review (PRR)

The PRR is an assessment of the contractor's ability to produce the items on the contract. It is usually a series of reviews conducted before an LRIP or full-scale production decision. For more information, see Chapter 10, paragraph 10.3 (Production Readiness Reviews).

8.3.3.11 Configuration Change Control

The Configuration Change Control review is an assessment of the impact of engineering or design changes. It is conducted by the engineering, T&E and PM portions of the PMO. Most approved Class I engineering change proposals will require additional testing, and the test manager must accommodate the new schedules and resource requirements. Adequate testing must be accomplished to ensure integration and compatibility of these changes. For example, an engineering change review was conducted to replace the black and white monitors and integrate color monitors into the Airborne Warning and Control System (AWACS). Further, the AWACS operating software had to be upgraded to handle color enhancement. The review was conducted by the government PMO; and sections of the PMO were tasked to contract, test, engineer, logistically support, control, cost and finance the change to completion. Configuration control and engineering changes are discussed in MIL-STD-480.

8.4 SUMMARY

Design reviews are an integral and essential part of the system engineering process. The meetings range from very formal reviews by government and contractor PMs to informal technical reviews concerned with product or task elements of the work breakdown structure. Reviews may be conducted in increments over time. All reviews share the common objective of determining the technical adequacy of the existing design to meet technical requirements. The DT/OT assessments and test results are made available to the reviews, and it is important that the test community be involved.

9

COMBINED AND CONCURRENT TESTING

9.1 INTRODUCTION

The terms "concurrency," "concurrent testing," and "combined testing" are sometimes subject to misinterpretation. Concurrency is defined as an approach to system development and acquisition in which phases of the acquisition process, which normally occur sequentially, overlap to some extent. For example, a weapon system enters the production phase while development efforts are still underway.

Concurrent testing refers to circumstances when development testing and operational testing take place at the same time as two parallel but separate and distinct activities. In contrast, combined testing refers to a single test program conducted to support development test (DT) and operational test (OT) objectives. This chapter discusses the use of combined testing and concurrent testing, and highlights some of the advantages and disadvantages associated with these approaches.

9.2 COMBINING DEVELOPMENT TEST AND OPERATIONAL TEST

Certain test events can be organized to provide information useful to development testers and operational testers. For example, a prototype free-fall munitions could be released from a fighter aircraft at operational employment conditions instead of from a static stand to satisfy DT and OT objectives. Such instances need to be identified to prevent unnecessary duplication of effort and to control costs. A combined testing approach is also appropriate for certain specialized types of testing. For example, in the case of nuclear survivability and hardness testing, systems cannot be tested in a totally realistic operational environment; therefore, a single test program is often used to meet both development and operational test objectives.

The DODI 5000.2 encourages combined testing and states that "a combined DT&E [development test and evaluation] and OT&E [operational test and evaluation] approach should be considered when there are time and cost savings. The combined approach must not compromise either DT or OT objectives." If this approach is elected, planning efforts must be carefully coordinated early in the program to ensure data is obtained to satisfy the needs of both the developing agency and the independent operational tester. Care must also be exercised to ensure a combined test program contains dedicated operational test events to satisfy the requirement for an independent evaluation. A final independent phase of OT&E testing shall be required for beyond low-rate initial production (BLRIP) decisions. In all combined test programs, provisions for separate independent development and operational evaluations of test results should be provided.

Service regulations describe the sequence of activities in a combined testing program as follows:

Although OT&E is separate and distinct from DT&E, most of the generated data are mutually beneficial and freely shared. Similarly, the resources needed to conduct and support both test efforts are often the same or very similar. Thus, when sequential DT&E and OT&E efforts would cause delay or increase the acquisition cost of the system, DT&E and OT&E are combined. When combined testing is planned, the necessary test conditions and data required by both DT&E and OT&E organizations must be integrated. Combined testing can normally be divided into three segments. In the first segment, DT&E event[s] usually assume priority because critical technical and engineering tests must be accomplished to continue the engineering and development process. During this early period, OT&E personnel participate to gain familiarity with the system and to gain access to any test data that can support OT&E. Next, the combined portion of the testing frequently includes shared objectives or joint data requirements. The last segment normally contains the dedicated OT&E or separate OT&E events to be conducted by the OT&E agency. The OT&E agency and implementing command must ensure the combined test is planned and executed to provide the necessary operational test information. The OT&E agency provides an independent evaluation of the OT&E portion and is ultimately responsible for achieving OT&E objectives.

The testing of the Navy's F-14 aircraft has been cited as an example of a successful combined test and evaluation (T&E) program (Reference 112). A key factor in the success of the F-14 approach was the selection of a T&E coordinator responsible for supervising the generation of test plans that integrated the technical requirements of the developers with the operational requirements of the users. The T&E coordinator was also responsible for the allocation of test resources and the overall management of the test. In a paper for the Defense Systems Management College, Mr. Thomas Hoivik describes the successful F-14 test program as follows:

> The majority of the Navy developmental and operational testing took place during the same period and even on the same flights. Maximum use was made of contractor demonstrations witnessed by the Navy testing activities to obviate the retesting of a technical point already demonstrated by the contractor. Witnessing by testing activities was crucially important and allowed the contractor's data to be readily accepted by the testing activities. This approach also helped to eliminate redundancy in testing, i.e. the testing of the same performance parameter by several different activities which has been a consistent and wasteful feature of Navy testing in the past.

Obviously, this approach placed a great deal of responsibility directly on the shoulders of the T&E Coordinator, and required his staff to deal knowledgeably with a wide-ranging and complex test plan.

9.3 CONCURRENT TESTING

In 1983, a senior DOD test and evaluation official testified that a concurrent testing approach is usually not an effective strategy (Reference 106). He acknowledged, however, that certain test events may provide information useful to development and operational testers, and test planners must be alert to identify those events. His testimony included the following examples of situations where a concurrent testing approach was unsuccessful:

(1) During AAH (Advanced Attack Helicopter) testing in 1981, the Target Acquisition Designation System (TADS) was undergoing developmental and operational testing at the same time. The schedule did not allow enough time for qualification testing (a development test activity) of the TADS prototype prior to a full field test of the total aircraft system, nor was there time to introduce changes to TADS problems discovered in tests. As a result, the TADS performed poorly and was unreliable during the operational test. The resulting DSARC [Defense Systems Acquisition Review Council] action required the Army to fix and retest the TADS prior to release of second year and subsequent production funds.

(2) When the AIM-7 Sparrow air-to-air missile was tested an attempt was made to move into operational testing while developmental reliability testing was still underway. The operational test was suspended after less than two weeks because of poor reliability of the test missiles. The program concentrated on an intensive reliability improvement effort. A year after the initial false start, a full operational test was conducted and completed successfully.

(3) The Maverick missile had a similar experience of being tested in an operational environment before component reliability testing was completed. As a result, reliability failures had a major impact on the operational testers and resulted in the program being extended.

9.4 ADVANTAGES AND LIMITATIONS

Before adopting a combined or concurrent testing approach, program and test managers are advised to consider the advantages and disadvantages summarized in Table 9-1.

9.5 SUMMARY

A combined or concurrent testing approach may offer an effective means of shortening the time required for testing and achieving cost savings. If such an approach is used, extensive coordination is required to ensure the development and operational requirements are addressed.

It is possible to have combined test teams, consisting of DT&E and OT&E personnel, involved throughout the testing process. The teams can provide mutual support and share mutually beneficial data as long as the test program is carefully planned and executed and reporting activities are conducted separately.

Table 9-1. Combined vs. Concurrent Testing: Advantages and Limitations

ADVANTAGES	LIMITATIONS
COMBINED TESTING	
• Shortens time required for testing and, thus, the acquisition cycle • Achieves cost savings by eliminating redundant activities. • Early involvement of OT&E personnel during system development increases their familiarity with system. • Early involvement of OT&E personnel permits communication of operational concerns to developer in time to allow changes in system design.	• Requires extensive early coordination. • Test objectives may be compromised. • Requires development of DT/OT common test data base. • Combined testing programs are often conducted in a development environment. • Test will be difficult to design to meet DT and OT requirements. • The system contractor is prohibited by law from participating in IOT&E. • Time constraints may result in less coverage than planned for OT&E objectives.
CONCURRENT TESTING	
ADVANTAGES	LIMITATIONS
• Shortens time required for testing and, thus, the acquisition cycle. • Achieves cost savings by overlapping redundant activities. • Provides earlier feedback to the development process.	• Requires extensive coordination of test assets. • If system design is unstable and far-reaching modifications are made, OT&E must be repeated. • Concurrent testing programs often do not have development test data available for OT&E planning and evaluation. • Contractor personnel frequently perform maintenance functions in a DT&E. Logistic support by user must be available earlier for IOT&E. • Limited test assets may result in less coverage than planned for OT&E objectives.

10

PRODUCTION RELATED TESTING ACTIVITIES

10.1 INTRODUCTION

Most of the test and evaluation (T&E) discussed in this guidebook concerns the testing of the actual weapon or system being developed, but the program manager (PM) must also evaluate production-related test activities and the production process. This chapter describes production management and the production process testing required to ensure the effectiveness of the manufacturing process and the producibility of the system's design.

Normally, the development test (DT) and operational test (OT) organizations are not involved directly in this process. Usually, the manufacturing and quality assurance sections of the program office and a representative of the government Defense Plant Representatives Office (DPRO) oversee/perform many of these functions.

10.2 PRODUCTION MANAGEMENT

Production (manufacturing) management is defined as "the effective use of resources to produce, on schedule, the required number of end items that meet specified quality, performance, and cost. Production management includes, but is not limited to, industrial resource analysis, producibility assessment, producibility engineering and planning, production engineering, industrial preparedness planning, post-production planning, and productivity enhancement," (DODD 5000.34). Production management begins early in the acquisition process, as early as the Concept Exploration and Definition (CED) Phase, and is specifically addressed at each program milestone (MS) decision point. For instance, during the CED, production feasibility, costs and risks should be addressed. Before MS I, the PM must conduct an industrial resource analysis (IRA) to determine the availability of production resources (e.g., capital, material, manpower) required to support the production of the weapon system. On the basis of the results of the IRA, critical materials, deficiencies in the U.S. industrial base and requirements for new or updated manufacturing technology can be identified. Analysis of the industrial-base capacity is one of the considerations in preparing the Integrated Program Summary for the MS I decision. As development proceeds, the manufacturing strategy is developed; and detailed plans are made for the Production Phase. Independent producibility assessments, conducted in preparation for the transition from development to production, are reviewed at the MS II decision point. At MS II, the Engineering and Manufacturing Development (EMD) decision, the producibility of the system design concept is evaluated to verify that the system can be manufactured in compliance with the production-cost and the industrial-base goals and thresholds.

The MS III Production and Deployment decision is supported by an assessment of the readiness of the system to enter production. The system cannot enter the Full-Rate Production and Deployment Phase until it is determined the principal contractors have the necessary resources (i.e., physical, financial, and managerial capacity) to achieve the cost and schedule commitments and to meet peacetime and mobilization requirements for production of the system. The method of assessing production readiness in preparation for MS III is the Production Readiness Review (PRR), which is conducted by the PM and staff. An independent assessment of production readiness is conducted at the same time by the DOD Product Engineering Services Office (DPESO) of Deputy Director, Defense Research and Engineering (DDDR&E). The DPESO reports its findings directly to the defense acquisition executive.

10.3 PRODUCTION READINESS REVIEW (PRR)

The policy, procedures and guidelines for PRRs are delineated in MIL-STD-1521 as follows:

> This review is intended to determine the status of completion of the specific actions which must be satisfactorily accomplished prior to executing a production go-ahead decision. The review is accomplished in an incremental fashion during the Engineering and Manufacturing Phase, usually two initial reviews and one final review to assess the risk in exercising the production go-ahead decision. In its earlier stages the PRR concerns itself with gross level manufacturing concerns such as the need for identifying high risk/low yield manufacturing processes or materials or the requirement for manufacturing development effort to satisfy design requirements. Timing of the incremental PRR's is a function of program posture and is not specifically locked into other reviews.

The conduct of a PRR is the responsibility of the PM, who usually appoints a director. The director assembles a team comprised of individuals in the disciplines of design, industry, manufacturing, procurement, inventory control, contracts, engineering and quality training. The PRR director organizes and manages the team effort and supervises preparation of the findings. The PRR is conducted as a time-phased effort during the EMD Phase following the guidelines presented in DODI 5000.2.

10.4 QUALIFICATION TESTING

Qualification testing is performed to verify the design and manufacturing process, and it provides a baseline for subsequent acceptance tests. The production qualification testing is conducted at the unit, subsystem and system level on production items and is completed before the production decision. The results of these tests are a critical factor in assessing the system's readiness for production. Downline production qualification tests are performed to verify process control and may be performed on selected parameters rather than at the levels originally selected for qualification.

10.4.1 Production Qualification Tests (PQT)

Production qualification tests are a series of formal contractual tests conducted to ensure design integrity over the specified operational and environmental range. The tests are conducted on prototype or preproduction items fabricated to the proposed production design drawings and specifications. The PQTs include all contractual reliability and maintainability demonstration tests required prior to production release. For volume acquisitions, these tests are a constraint to production release.

10.4.2 First Article Tests (FAT)

First article tests consist of a series of formal contractual tests conducted to ensure the effectiveness of the manufacturing process, equipment and procedures. These tests are conducted on a random sample from the first production lot. These series of tests are repeated if the manufacturing process, equipment or procedure is changed significantly and when a second or alternative source of manufacturing is brought on line.

10.5 TRANSITION TO PRODUCTION

In an acquisition process, often the first indication that a system will experience problems is during the transition from EMD to low-rate initial production. This transition continues over an extended period, often months or years; and during this period, the system is undergoing stringent contractor and government testing. There may be unexpected failures requiring significant design changes, which impact on quality, producibility, supportability and may require program schedule slippage.

Long periods of transition usually indicate that insufficient attention to design or producibility was given early in the Combat Exploration/Definition (CED) or Demonstration and Validation phases.

10.5.1 The Transition Plan

Producibility engineering and planning (PEP) is the common thread that guides a system from CED to production. The plan is a management tool used to ensure that adequate risk-handling measures have been taken to transition from development to production. It contains a checklist to be used during the readiness reviews. The plan should tie together the applications of designing, testing and manufacturing activities to reduce data requirements, duplication of effort, costs and scheduling and to ensure early success of the EMD first production article.

10.5.2 Testing During the Transition

Testing accomplished during the transition from development to production will include acceptance testing, manufacturing screening and final testing. These technical tests are performed by the contractor to ensure the system will transition smoothly and that test design and manufacturing issues affecting design are addressed. During this same period, the government will be using the latest available configuration item to conduct the initial operational test and evaluation (IOT&E). The impact of these tests may overwhelm the configuration management of the system unless careful planning is accomplished to handle these changes.

10.6 LOW-RATE INITIAL PRODUCTION (LRIP)

Low-rate initial production is the production of a system in limited quantity to provide articles for operational test and evaluation and establish an initial production base. Also, it permits an orderly increase in the production rate sufficient to lead to full-rate production upon successful completion of operational testing. The decision to have an LRIP is made at the MS II approval of the program acquisition strategy. At that time, the PM must identify: (1) the quantity to be produced during LRIP and (2) the quantity of LRIP articles to be used for IOT&E (to be approved by the Director, Operational Test and Evaluation (DOT&E)). When the decision authority thinks the systems will not perform to expectation, he will direct that it not proceed beyond LRIP. The DOT&E submits a report, on all major systems, to congressional committees before the MS III decision to proceed beyond LRIP is made.

10.7 PRODUCTION ACCEPTANCE TEST AND EVALUATION (PAT&E)

Production acceptance test and evaluation ensures that production items demonstrate the fulfillment of the requirements and specifications of the procuring contract or agreements. The testing also ensures the system being produced demonstrates the same performance as the preproduction models. The procured items or system must operate in accordance with Type A (system) and Type C (production) specifications. The PAT&E is usually conducted by the program office quality assurance section at the contractor's plant and may involve operational users.

For example, for the Rockwell B-1B Bomber production acceptance, Rockwell and Air Force quality assurance inspectors reviewed all manufacturing and ground testing results for each aircraft. In addition, a flight test team, composed of contractor and Air Force test pilots, flew each aircraft a minimum of 10 hours, demonstrating all onboard aircraft systems while in flight. Any discrepancies in flight were noted, corrected and tested on the ground; they were then retested on subsequent checkouts and acceptance flights. Once each aircraft had passed all tests and all systems were fully operational, Air Force authorities accepted the aircraft. The test documentation also became part of the delivered package. During this test period, the program office monitored each aircraft's daily progress.

10.8 SUMMARY

A primary purpose of production-related testing is to lower the production risk in a major defense acquisition program. The PM must ensure the contractor's manufacturing strategy and capabilities will result in the desired product within acceptable cost. The LRIP and PAT&E also play major roles in ensuring the production unit is identical to the design drawings and conforms to the specifications of the contract.

Table 10-1. PRR Guidelines Checklist

PRODUCT DESIGN
- Producible at low risk
- Stabilized at low rate of change
- Validated
- Reliability, maintainability and performance demonstrated
- Components engineering has approved all parts selections

INDUSTRIAL RESOURCES
- Adequate plant capacity (peacetime and and wartime demands)
- Facilities, special production and test equipment, and tooling identified
- Needed plant modernization (CAD/CAM, other automation) accomplished, which produces an invested captive payback in two to five years
- Associated computer software developed
- Skilled personnel and training programs available

PRODUCTION ENGINEERING AND PLANNING
- Production plan developed (reference MIL-STD-1528)
- Production schedules compatible with delivery requirements
- Manufacturing methods and processes integrated with facilities, equipment, tooling and plant layout
- Value engineering applied
- Alternate production approaches available
- Drawings, standards and shop instructions are explicit
- Configuration management adequate
- Production policies and procedures documented
- Management information system adequate
- Contractor's management structure is acceptable to the PMO
- The PEP checklist has been reviewed

MATERIALS
- All selected materials approved by contractor's materiel engineers
- Bill of materials prepared
- "Make-or-Buy" decisions complete
- Procurement of long lead-time items identified
- Sole-source and government-furnished items identified
- Contractor's inventory-control system adequate
- Contractor's material cost procurement plan complete

QUALITY ASSURANCE (QA)
- Quality plan in accordance with contract requirements
- Quality control procedures and acceptance criteria established
- QA organization participates in production planning effort

LOGISTICS
- Operational support, test and diagnostic equipment available at system deployment
- Training aids, simulators and other devices ready at system deployment
- Spares integrated into production lot flow

III

MODULE

Operational
Test and Evaluation

Operational Test and Evaluation (OT&E) is conducted to ensure a weapon system meets the validated requirements of the user in a realistic scenario. Operational tests are focused on operational requirements, effectiveness and suitability, and not on the proof of engineering specifications, as is the case with development testing. This module provides an overview of OT&E and discusses how OT&E results provide essential information for milestone decisions.

11

INTRODUCTION TO OPERATIONAL TEST AND EVALUATION

11.1 INTRODUCTION

This chapter provides an introduction to the concept of operational test and evaluation (OT&E). It outlines the purpose of OT&E, discusses the primary participants in the OT&E process, describes several types of OT&E, and includes some general guidelines for the successful planning, execution and reporting of OT&E programs.

11.2 PURPOSE OF OT&E

Operational test and evaluation is conducted for major programs by an organization that is independent of the developing, procuring and using commands. It is normally conducted in phases, each of which are keyed to a decision review in the materiel acquisition process. It is conducted with typical user operators, crews or units in realistic and operational environments. The OT&E provides the decision authority with an estimate of:

(1) The degree of satisfaction of the user's requirements expressed as operational effectiveness and suitability of the new system;

(2) The system's desirability, considering equipment already available, and operational benefits or burdens associated with the new system;

(3) The need for further development of the new system;

(4) The adequacy of doctrine, organizations, operating techniques, tactics and training for employment of the system; of maintenance support for the system; and of the system's performance in the countermeasures environment.

11.3 TEST PARTICIPANTS

The OT&E of developing systems is managed by an independent testing agency, which each Service is required to maintain. It is accomplished under conditions of operational realism whenever possible. Personnel who operate, maintain and support the system during OT&E are trained to a level commensurate with that of personnel who will perform these functions under peacetime and wartime conditions. Also, program management office (PMO) personnel and test coordinating groups play important parts in the overall OT&E planning and execution process.

11.3.1 Program Management Office

Even though operational testing is performed by an independent organization, the program manager (PM) plays an important role in its planning, reporting and funding. He must coordinate program activities with the test community, especially the operational test agencies. He ensures that testing can address the critical issues, and provides feedback from OT&E testing activities to contractors.

At each milestone review, the PM is required to brief the decision authority on the testing planned and completed on the program. It is, therefore, important that PMO personnel have a good understanding of the test program and that they work with the operational test community. This will ensure OT&E is well-planned and adequate resources are available. The PMO should involve the test community by organizing test coordinating groups at program initiation and by establishing channels of communication between the PMO and the key test organizations. The PMO can often avoid misunderstandings by aggressively monitoring the system testing and providing up-to-date information to key personnel in OSD and the Services. The PMO staff should keep appropriate members of the test community well-informed concerning system problems and the actions taken by the PMO to correct them.

11.3.2 Test Coordinating Groups

The test and evaluation (T&E) working groups, such as the Army's Test Integration Working Groups (TIWG) and Air Force's Test Planning Working Groups (TPWG), are chartered by their respective PMO to coordinate and integrate the planning and execution of the T&E program. The Army and Air Force groups are chaired by a representative of the PMO, often the deputy for test and evaluation or systems engineer. Members of these groups represent various communities including the user, development and operational testing, independent evaluation, logistics, training and contractor, as appropriate. The functions of the groups are to: facilitate the use of testing expertise, instrumentation, facilities, simulations and models; integrate test requirements; accelerate the TEMP coordination process; resolve test cost and scheduling problems; and provide a forum to ensure T&E of the system is coordinated. The existence of a test coordinating group does not alter the responsibilities of any command or headquarters; and, in the event of disagreement within a group, the issue is resolved through the normal command/staff channels. Within the Air Force, the TPWG may help to prepare the test portions of the request for proposal and related contractual documentation and to evaluate the contractors' proposals. In all Services, the groups help develop the TEMP.

11.3.3 Service Operational Test Agencies

The operational test and evaluation agencies (OTA) should become involved early in the system's life cycle, usually before program starts at Milestone (MS) I. At this time, they can begin to develop strategies for conducting of operational tests. As test planning continues, a more-detailed Test and Evaluation Master Plan (TEMP) is developed and the test resources are identified and scheduled. During the early stages, the OTAs structure an OT&E program consistent with the approved acquisition strategy for the system, identify critical operational test issues, and assess the adequacy of candidate systems. As the program moves into advanced planning, OT&E efforts are directed toward becoming familiar with the system, encouraging interface between the user and developer and further refining the critical operational issues. Each Service has an independent organization dedicated to planning, executing and reporting the results of that Service's OT&E activities. These organizations are the: Army Operational Test and Evaluation Command (OPTEC), Navy Operational Test and Evaluation Force (OPTEVFOR), Air Force Operational Test and Evaluation Center (AFOTEC), and Marine Corps Operational Test and Evaluation Activity (MCOTEA).

11.3.4 Test Personnel

Operational testing is conducted on materiel systems with "typical" user players in a realistic operational environment. It uses personnel with the same military occupational specialties as those who will operate, maintain and support the system when fielded. Participants are trained in the system's operation based on the Service's operational mission profiles. Because some operational tests consist of force-on-force tests, the forces opposing the tested system must also be trained in threat tactics and doctrine. For operational testing conducted before initial operational test and evaluation (IOT&E), most system training is conducted by the system's contractor. For IOT&E, the contractor trains the school cadre who train other troops. As the system enters full-rate production, the Services assume training responsibilities.

11.4 TYPES OF OT&E

Operational Test and Evaluation can be subdivided into two phases: operational testing performed before MS III (full-rate production/deployment decision) and the operational testing performed after MS III. The Pre-MS III OT&E includes operational assessments and IOT&E. Operational assessments begin early in the program, frequently before program start (MS I) and continue until the system is certified as ready for IOT&E. The initial operational test and evaluation is conducted just before the full-rate production/deployment decision. After full-rate production/deployment, all subsequent operational testing is referred to as follow-on operational test and evaluation (FOT&E). In the Air Force, if no research and development funding is committed to a system, qualification OT&E may be performed in lieu of IOT&E. The Navy uses the term "OPEVAL" to define IOT&E.

11.4.1 Early Operational Assessments

Early operational assessments are conducted primarily to forecast and evaluate the potential operational effectiveness and suitability of the weapon system during development. Early operational assessments start in the Concept/Definition Phase and are conducted on the developing system until MS II.

11.4.1.1 Operational Assessments

Operational assessments begin after MS II, when the OTAs start their evaluations of system-level performance. The OTA uses any testing results and data from other sources during an assessment. These data are evaluated by the OTA from an operational point of view. As the program matures, these operational assessment requirements are conducted on preproduction articles until the system is fully developed and certified ready for its IOT&E or OPEVAL in the Navy.

11.4.1.2 Initial Operational Test and Evaluation (Navy OPEVAL)

Initial operational test and evaluation is the final dedicated phase of OT&E preceding a full-rate production decision. It is the final evaluation that entails dedicated operational testing of production-representative test articles and uses typical operational personnel in a scenario that is as realistic as possible. The IOT&E is conducted by an OT&E agency independent of the contractor, PMO or developing agency. The test is defined in DODI 5000.2 as:

> All operational test and evaluation conducted on production or production representative articles, to support the decision to proceed beyond low-rate initial production. It is conducted to provide a valid estimate of expected system operational effectiveness and operational suitability. The definition of "OT&E" as spelled out in congressional legislation (see glossary) is gen-

erally considered applicable only to initial operational test and evaluation (IOT&E).

Further, IOT&E must be conducted without system contractor personnel participation in any capacity other than stipulated in service wartime tactics and doctrine as set forth in Public Law 99-661 by Congress. The results from this test are evaluated and presented to the milestone decision authority (i.e. MS III, the decision to enter full-rate production) to support the beyond-LRIP decision. This phase of OT&E addresses the critical issues identified in the Operational Requirements Document (ORD) and the TEMP.

11.4.2 Follow-On Operational Test and Evaluation

Follow-on operational test and evaluation is conducted after the MS III decision. The tests are conducted in a realistic tactical environment similar to that used in IOT&E, but many test items may be used. Normally FOT&E is conducted using fielded production systems. Specific objectives of FOT&E include testing modifications that are to be incorporated into production systems, completing any deferred or incomplete IOT&E, and assessing reliability including spares support. The tests are also used to evaluate the system in a different platform application for new tactical applications or against new threats.

11.4.3 Qualification Operational Test and Evaluation

Air Force qualification operational test and evaluation may be performed by the major command, user or AFOTEC and is conducted on minor modifications or new applications of existing equipment when no research and development funding is required. An example of a program in which QOT&E was performed by the Air Force is the A-10 Air-to-Air Self Defense Program. In this program the mission of the A-10 was expanded from strictly ground support to include an air-to-air defense role. To accomplish this the A-10 aircraft was modified with off-the-shelf AIM-9 and air-to-air missiles; QOT&E was performed on the system to evaluate its operational effectiveness and suitability.

11.5 TEST PLANNING

Operational test planning is one of the most important parts of the OT&E process. Proper planning facilitates the acquisition of data to support the determination of the weapon system's operational effectiveness and suitability. Planning must be pursued in a deliberate, comprehensive and structured manner. Careful and complete planning may not guarantee a successful test program; but inadequate planning can result in significant test problems, poor system performance and cost overruns. Operational test planning is conducted by the OTA before program start, and more-detailed planning usually starts about two years before each operational test event.

Operational planning can be divided into three phases: early planning, advanced planning and detailed planning. Early planning entails developing critical operational issues, formulating a plan for evaluations, determining the concept of operation, envisioning the operational environment and developing mission scenarios and resource requirements. Advanced planning encompasses the determination of the purpose and scope of testing and identification of measures of effectiveness (MOEs) for critical issues. It includes developing test objectives, establishing a test approach, and estimating test resource requirements. Detailed planning involves developing step-by-step procedures to be followed as well as the final coordination of resource requirements necessary to carry out OT&E.

11.5.1 Testing Critical Operational Issues (COI)

Part 15 of DODI 5000.2 defines a critical operational issue as:

> A key operational effectiveness or operational suitability issue that must be examined in operational test and evaluation to determine the system's capability to perform its mission. A critical operational issue is normally phrased as a question to be answered in evaluating a system's operational effectiveness and/or operational suitability.

One of the purposes of OT&E is to resolve COIs about the system. The first step in an OT&E program is to identify these critical issues, some of which are explicit in the ORD. Examples can be found in questions such as: "How well does the system perform a particular aspect of its mission?" "Can the system be supported logistically in the field?" Other issues arise from questions asked about system performance or how it will affect other systems with which it must operate. Critical issues provide focus and direction for the operational test. Identifying the issues is analogous to the first step in the system engineering process, that is, defining the problem. When critical operational issues are properly addressed, deficiencies in the system can be uncovered. They form the basis for a structured technique of analysis by which detailed subobjectives or MOEs can be established. During the operational test, each subobjective is addressed by an actual test measurement. After these issues are identified, the evaluation plans and test design are developed for test execution. (For more information, see the chapter on evaluation.)

11.5.2 Test Realism

Test realism for OT&E will vary directly with the degree of system maturity. Efforts early in the acquisition program should focus on active involvement of users and operationally oriented environments. Fidelity of the "combat environment" should peak during the IOT&E when force-on-force testing of the production representative system is conducted. The degree of success in replicating a realistic operational environment has a direct impact on the credibility of the IOT&E test report. Areas of primary concern for the test planner can be derived from the legislated definition of OT&E:

> (1) A <u>field test</u> includes all of the elements normally expected to be encountered in the operational arena, such as appropriate size and type of maneuver terrain, environmental factors, day/night operations, austere living conditions, etc.
>
> (2) <u>Realistic combat</u> should be replicated using appropriate tactics and doctrine, representative threat forces properly trained in the employment of threat equipment, free play responses to test stimulus, stress, "dirty" battle area (fire, smoke, NBC, ECM, etc.), wartime tempo to operations, real time casualty assessment, and forces requiring interoperability.
>
> (3) Any item means the production representative configuration of the system at that point in time, including appropriate logistics tail.
>
> (4) Typical military users are obtained by taking a cross section of adequately trained skill levels and ranks of the intended operational force. Selection of "golden crews" or the best of the

best does not provide test data reflecting the successes nor problems of the "murphy and gang" of typical units.

In his book, *Operational Test and Evaluation,* Roger Stevens states, "In order to achieve realism effectively in an OT&E program, a concern for realism must pervade the entire test program from the very beginning of test planning to the time when the very last test iteration is run." Realism is a significant issue during planning and execution of OT&E.

11.5.3 Selection of a Test Concept

An important step in the development of an OT&E program is to develop an overall test program concept. Determinations must be made regarding when OT&E will be conducted during systems development, what testing is to be done on production equipment, how the testing will be evolutionary, and what testing will have to wait until all system capabilities are developed. This concept can best be developed by considering a number of aspects such as test information requirements, system availability for test periods, and the demonstration of system capabilities. The test concept is driven by the acquisition strategy and is a road map used for planning test and evaluation events.

11.6 TEST EXECUTION

An operational test plan is only as good as the execution of that plan. The execution is the essential bridge between test planning and test reporting. The test is executed through the OTA test director's efforts and the actions of the test team. For successful execution of the OT&E plan, the test director must direct and control the test resources and collect the data required for presentation to the decision authority. He must prepare for testing, activate and train the test team, develop test procedures and operating instructions, control data management, create OT&E plan revisions, and manage each of the test trials. His data management duties will encompass collecting raw data, creating a data status matrix, ensuring data quality, processing and reducing, verifying, filing, storing, retrieving and analyzing. Once all tests have been completed and the data is reduced and analyzed, the results must be reported. A sample test organization used for the Army OT&E of the improved 81mm mortar is illustrated in Figure 11-1. (In the Army, the Deputy Test Director comes from the OTA and controls the daily operational test activity.)

11.7 TEST REPORTING

The IOT&E test report is a very important document. It must communicate the results of completed tests to decision authorities in a timely, factual, concise, comprehensive and accurate manner. The report must present a balanced view of the weapon system's successes and problems during testing, illuminating both the positive aspects and system deficiencies discovered. Analysis of test data and their evaluation may be in one report (USAF, USN) or in separate documents (USA, USMC).

There are four types of reports most frequently used in reporting OT&E results. These include status, interim, quick-look and final reports. The status report gives periodic updates (e.g., monthly, quarterly) and reports recent test findings (discreet events such as missile firings). The interim report provides a summary of the cumulative test results to date when there is an extended period of testing. The quick-look reports provide preliminary test results, are usually prepared immediately after a test event (less than 7 days) and have been used to support program decision mile-

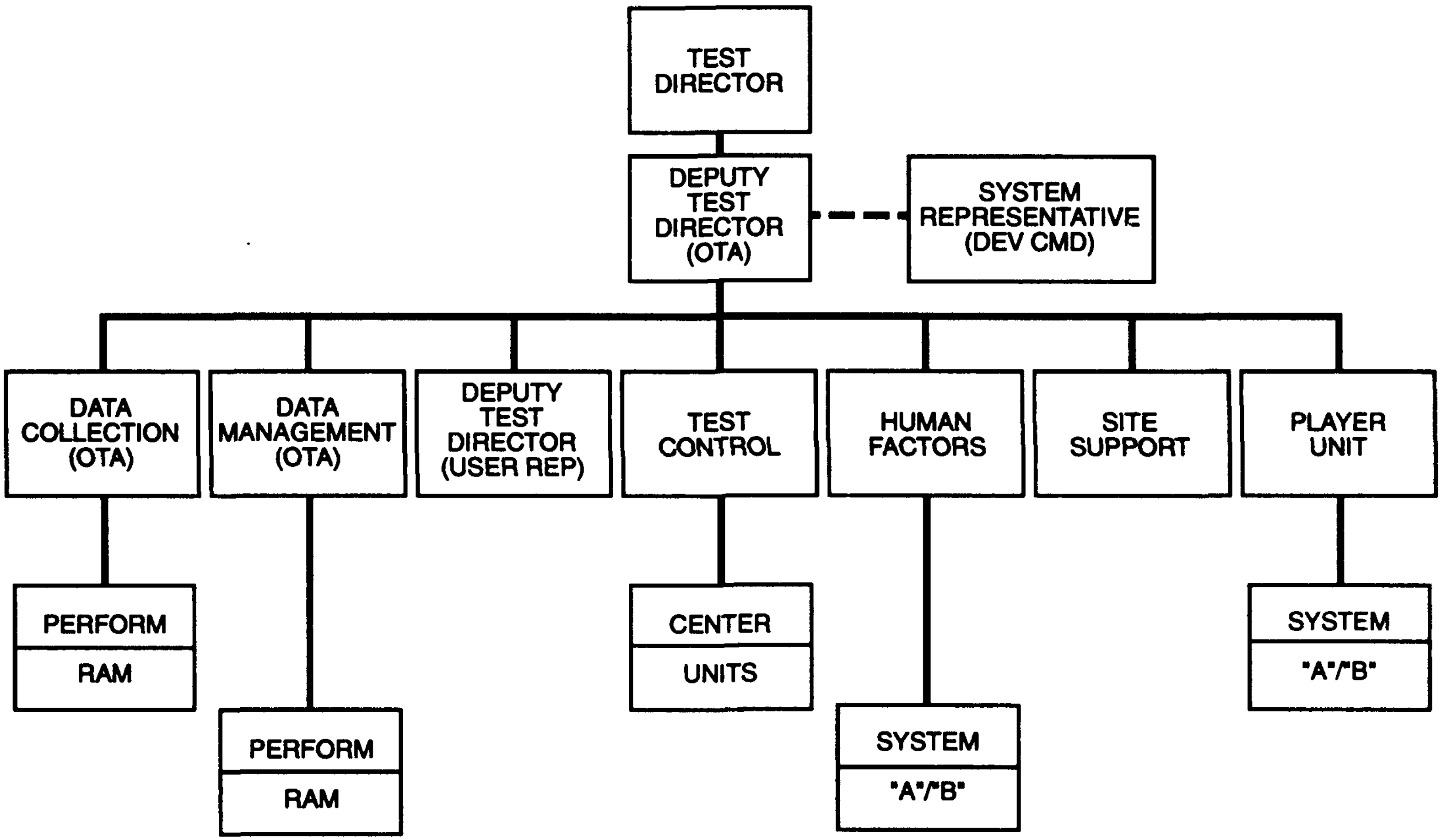


Figure 11-1. Organizational Breakdown of the I-81mm Mortar Operational Test Directorate

stones. The final T&E report (Air Force, Navy) or independent evaluation report (Army, Marine) presents the conclusions and recommendations including all supporting data and covering the entire IOT&E program.

11.8 SUMMARY

The purpose of OT&E is to assess operational effectiveness and suitability at each stage in the acquisition process. Operational effectiveness is a measure of the contribution of the system to mission accomplishment under actual conditions of employment. Operational suitability is a measure of the maintainability and reliability of the system; the effort and level of training required to maintain, support and operate it; and any unique logistic or training requirements it may have. The OT&E may provide information on tactics, doctrine, organization and personnel requirements and may be used to assist in the preparation of operating and maintenance instructions and other publications. One of the most important aspects is that OT&E provides an independent evaluation of the degree of progress made toward satisfying the user's requirements during the system development process.

12

OT&E TO SUPPORT MILESTONE DECISIONS

12.1 INTRODUCTION

Mindful of principles of objectivity and impartial evaluation, operational test and evaluation (OT&E) is conducted before each major milestone ‘MS) review to provide the decision authori, with the latest results from testing of critical operational issues. The philosophy of OT&E has been related to three terms — adequacy, quality and credibility:

Adequacy — The amount of data and realism of test conditions must be sufficient to support the evaluation of the critical operational issues.

Quality — The test planning, control of test events, and treatment of data must provide clear and accurate test reports.

Credibility — The conduct of the test and data handling must be separated from external influence and personal biases.

Operational testing is conducted to provide information to support DOD executive-level management decisions on major acquisition programs. Operational test and evaluation is accomplished using a test cycle of successive actions and documents. During the early stages of the program, the process is informal and modified as necessary. As programs mature, documentation for major systems and those designated by the Director, Operational Test and Evaluation (DOT&E) for oversight must be sent to the Office of the Secretary of Defense (OSD) for approval before the testing can be conducted or the systems can be cleared to proceed into low-rate initial production (LRIP). Figure 12-1 illustrates how OT&E relates to the acquisition process.

12.2 OT&E DURING THE CONCEPT EXPLORATION/DEFINITION PHASE (MS 0 to MS I)

The OT&E conducted during the Concept Exploration and Definition (CED) Phase is an early operational assessment (OA) focused on investigating the deficiencies identified during the mission area analysis. Operational testers participate in these evaluations to validate the OT&E requirements for future testing and to identify issues and criteria that can only be resolved through OT&E to initiate early test resource planning.

Before MS I, the OT&E objectives are to assist in evaluating alternative concepts to solve the mission area deficiencies and to assess the operational impact of the system. This early assessment also provides data to support a decision on whether to enter the Demonstration and Validation Phase. The OT&E conducted during the CED Phase supports developing estimates of:

(1) The military need for the proposed system;

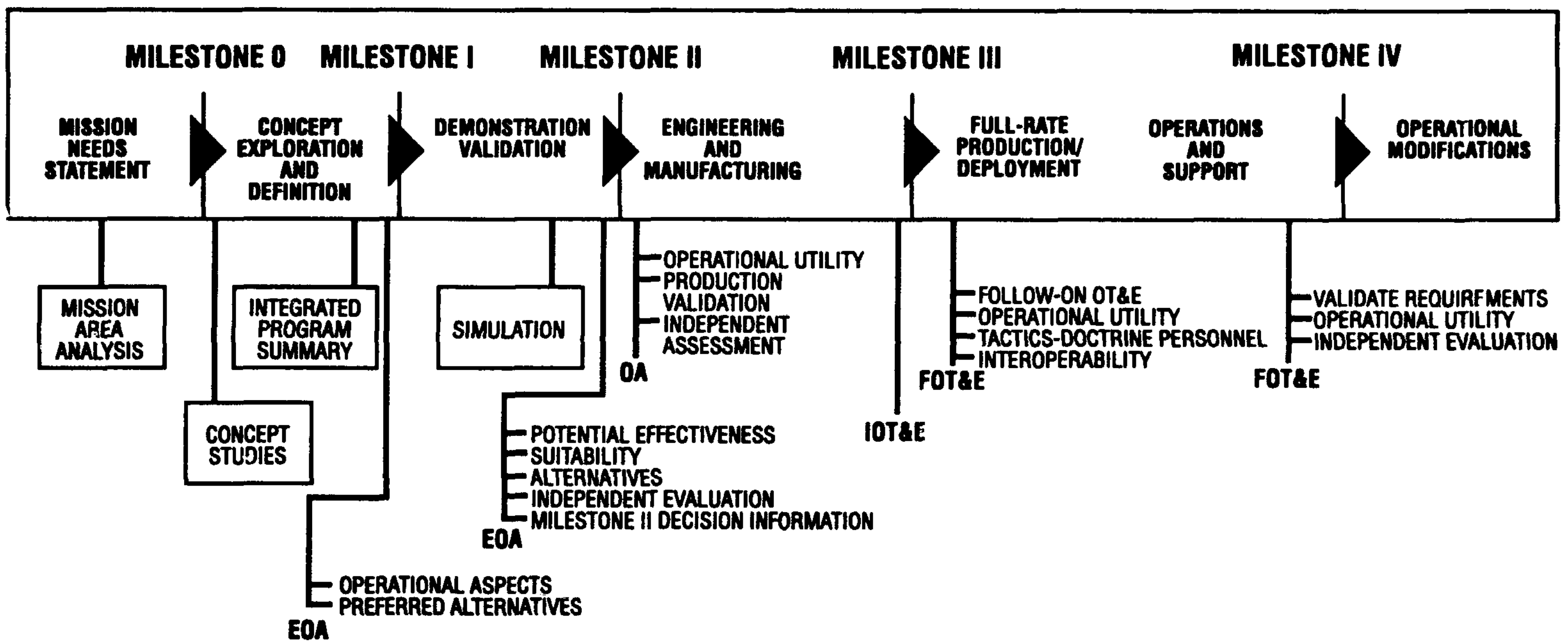


Figure 12-1. OT&E Related to the Milestone Process

(2) A demonstration that there is a sound physical basis for a new system;

(3) An analysis of concepts, based on demonstrated physical phenomena, for satisfying the military need;

(4) The system's affordability and life-cycle cost;

(5) The ability of a modification to an existing U.S. or allied system to provide needed capability;

(6) An operational utility assessment;

(7) An impact of the system on the force structure.

At MS I, there is normally no hardware available for the operational tester. Therefore, the early operational assessment is conducted from surrogate test and experiment data, breadboard models, factory user trials, mock-up/simulators, and user demonstrations (Figure 12-2). This makes early assessments difficult, and some areas cannot be covered in-depth. However, these assessments provide vital introductory information on the system's potential operational utility.

The OT&E products from this phase of testing include the information provided to the decision authority, data collected for further evaluation, input to the Test and Evaluation Master Plan (TEMP) and early test and evaluation (T&E) planning. Special logistics problems, program objectives, program plans, performance parameters and acquisition strategy are areas of primary influence to the operational tester during this phase and must be carefully evaluated to project the system's operational effectiveness and suitability.

12.3 OT&E DURING THE DEMONSTRATION AND VALIDATION PHASE (MS I to MS II)

Combined development test (DT)/OT&E or an early operational assessment during the Demonstration and Validation Phase, is conducted to support the MS II decision regarding a system's readiness to move into the Engineering and Manufacturing (EMD) Phase. In all cases, appropriate T&E must be conducted before the MS II decision, thereby providing data for identification of risk before more resources are committed. As appropriate, LRIP may be approved at MS II to verify production capability and to provide test resources needed to conduct interoperability, live fire, or operational testing.

12.3.1 Objectives of Early Operational Assessments

Early operational assessments are conducted to facilitate identification of the best design, indicate the risk level of performance for this phase of the development, examine operational aspects of the system's development, and estimate potential operational effectiveness and suitability. Additionally, an analysis of the planning for transition from development to production is initiated. Early operational assessments supporting the MS II decision are intended to:

(1) Assess the potential of the new system in relation to existing capabilities;

(2) Assess system effectiveness and suitability so that affordability can be evaluated for program cost vs. military utility;

(3) Assess the adequacy of the concept for employment, supportability and orga-

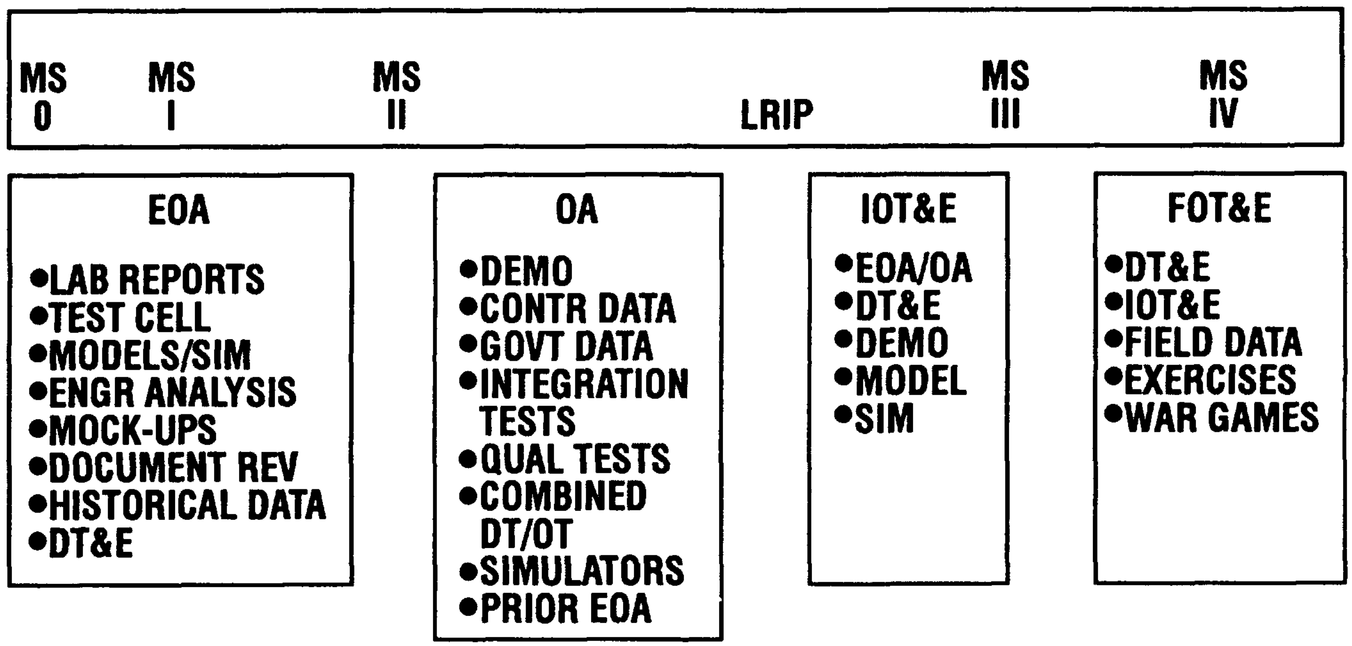


Figure 12-2. Sources of Data

nization; doctrinal, tactical and training requirements; and related critical issues;

(4) Estimate the need for the selected systems in consideration of the threat and system alternatives based on military utility;

(5) Assess the validity of the operational concept;

(6) List the key risk areas and critical operational issues that need to be resolved before EMD is initiated;

(7) Assess the need for LRIP of hardware to support initial operational test and evaluation (IOT&E) prior to the full-rate production decision;

(8) Provide data to support test planning for the EMD Phase.

During this phase, OT&E may be conducted on brassboard configurations, experimental prototypes or advanced development prototypes. Dedicated test time may be made available for the operational tester. However, the OT&E assessments may also make use of many other additional data sources. Examples of additional sources often used by the Army during this phase include: concept evaluation program tests, innovative testing, force development tests and experimentation (FDT&E), source selection tests, user participation in development test and evaluation (DT&E) and operational feasibility tests. The results from this testing, analysis and evaluation are documented in the Integrated Program Summary (IPS). These data, along with the mission needs and requirements documentation and TEMP, assist in the review of performance for the MS II decision.

12.4 OT&E DURING THE ENGINEERING AND MANUFACTURING DEVELOPMENT PHASE (MS II to MS III)

The IOT&E and operational assessments during the EMD Phase are conducted on engineering development models or production representative systems. These operational evaluations estimate the operational effectiveness and suitability and provide data on whether the system meets minimum operational thresholds. Just before the full-rate production MS III decision, the dedicated T&E is conducted on equipment that has been formally certified by the program manager as being ready for the "final OT&E." This dedicated IOT&E is conducted in a test environment as operationally realistic as possible.

12.4.1 OT&E Objectives

The OA/IOT&E conducted during EMD, is characterized by testing performed by user organizations in a field exercise to examine the organization and doctrine, integrated logistics support, threat, communications, command and control, and tactics associated with the operational employment of the unit during tactical operations. This includes estimates which:

(1) Assess operational effectiveness and suitability;

(2) Assess the survivability of the system;

(3) Assess the systems reliability, maintainability and plans for integrated logistics support;

(4) Evaluate manpower, personnel, training and safety requirements;

(5) Validate organizational and employment concepts;

(6) Determine training and logistics requirements deficiencies;

(7) Assess the system's readiness to enter full-rate production.

12.5 OT&E DURING THE PRODUCTION AND DEPLOYMENT PHASE (MS III to MS IV)

After the MS III decision, the emphasis shifts towards procuring production quantities, repairing hardware deficiencies, managing changes, and phasing in full logistics support. During initial deployment of the system, the OT&E agency and/or the user may perform follow-on test and evaluation (FOT&E) to refine the effectiveness and suitability estimates made during earlier OT&E.

The FOT&E is performed with production articles in operational organizations. It is normally funded with operation and maintenance (O&M) funds. The first FOT&E conducted during this phase may be used to:

(1) Ensure that the production system performs as well as reported at the MS III review;

(2) Demonstrate expected performance and reliability improvements;

(3) Ensure that the correction of deficiencies identified during earlier testing have been completed;

(4) Evaluate performance not tested during IOT&E.

Additional objectives of FOT&E are to validate the operational effectiveness and suitability of a modified system during an operational assessment of the system in new environments. The FOT&E may look at different platform applications, new tactical applications or the impact of new threats.

12.5.1 FOT&E of Integrated Logistic Support

The testing objectives to evaluate postproduction logistics readiness and support are to:

(1) Assess the logistics readiness and sustainability;

(2) Evaluate the weapon support objectives;

(3) Assess the implementation of integrated logistics support plans;

(4) Evaluate the capability of the logistics support activities;

(5) Determine the disposition of displaced equipment;

(6) Evaluate the affordability and life-cycle cost of the system.

12.6 SUMMARY

Operational test and evaluation is that T&E (operational assessments, IOT&E or FOT&E) conducted to estimate a system's operational effectiveness and suitability. They will identify needed modifications; provide information on tactics, doctrine, organizations and personnel requirements; and evaluate the system's logistic supportability. The acquisition program structure should include operational assessments or evaluations beginning early in the development cycle and continuing throughout the system's life cycle.

13

PROGRAM MANAGEMENT OFFICE OPERATIONAL TEST RESPONSIBILITIES

13.1 INTRODUCTION

In the government program management office (PMO), there should be a section responsible for test and evaluation (T&E). Besides being responsible for development test and evaluation (DT&E) support to the program manager, this section should be responsible for program coordination with the operational test and evaluation (OT&E) agency. The offices of the systems engineer or the Deputy for T&E may be designated to provide this support to the program manager. In some Services, responsibilities of the Deputy for (T&E) include coordination of test resources for all phases of OT&E.

13.2 CONTRACT RESPONSIBILITIES

The Deputy for T&E or his representative ensures that certain sections of the Request for Proposal (RFP) contain sufficient allowance for T&E support by contractors. This applies whether the contract is for a development item, a production item (limited production, such as low-rate initial production (LRIP) or full-rate production) or the enhancement/upgrade of portions of a weapons system. Where allowed within the law, contractor support for OT&E should be considered to help resolve basic issues such as data collection requirements, test resources, contractor test support and funding.

In the overall portion of the RFP, government personnel, especially those in the operational test agencies, must be guaranteed access to the contractor's development facilities, particularly during the DT&E Phase. Government representatives must be allowed to observe all contractor in-house testing and have access to test data and reports.

13.2.1 Data Requirements

The contract must specify the data the contractor will supply the operational test agency (OTA). Unlike DT&E, the contractor will not be making the OT&E plans, procedures or reports. These documents are the responsibility of the OTA. The PMO Deputy for T&E should include the OTA on the distribution list for all test documents that are of concern during the DT&E phase of testing so they will be informed of test item progress and previous testing. In this way, the OTA will be informed when developing their own test plans and procedures for OT&E. In fact, OTA representatives should attend the Contract Data Requirements List (CDRL) Review Board and provide the PMO with a list of the types of documents the OTA will need. The Deputy for T&E should coordinate the test sections of this data list with the OTA and indicate concerns at that meeting. All contractor test reports should be made available to the OTA. In return, the Deputy for T&E must ensure that he is informed of all OTA activities, understands their test procedures and plans and receives their test reports. Unlike DT&E, the PMO Deputy for T&E will not have report or document approval authority as he does over contractor documentation. The Deputy for T&E is always responsible for keeping the program manager informed of OT&E results.

13.2.2 Test Schedule

Another important early activity the Deputy for T&E must accomplish is to coordinate the OT&E test schedule. Since the contractor may be required to provide support, the OT&E test support may need to be contractually agreed upon before contract award. Sometimes, the Deputy for T&E can formulate a strawman schedule (based on previous experience) and present this schedule to the operational test representative at the initial test planning working group for review; or he can contact the OTA and arrange a meeting to discuss the new program. In the meeting, time requirements envisioned by OTA can be discussed. Input from that meeting then goes into the RFP and to the program manager. The test schedule must allow time for DT&E testing and OT&E testing if testing is not combined or test assets are limited. Before set-up of initial operational test and evaluation (IOT&E), certification of readiness for IOT&E may require a time gap for review of DT&E test results and refurbishment or corrections of deficiencies discovered during DT&E, etc. The test schedule for DT&E should not be so "success-oriented" that the IOT&E test schedule is adversely impacted, not allowing enough time for adequate operational testing or the reporting of IOT&E results. For example, if the DT&E schedule slips six months, the OT&E schedule and milestone decision should slip also. The IOT&E should not be shortened just to make a milestone decision date.

13.2.3 Contractor Support

The Deputy for T&E provides all T&E input to the RFP/SOW; he must determine, before the beginning of the program acquisition phase, whether the contractor will be involved in supporting OT&E and, if so, to what extent. According to Title 10, USC, the system contractor can only be involved in the conduct of IOT&E if, once the item is fielded, tactics and doctrine say the contractor will be providing support or operating that item during combat. If not, no system contractor support is allowed during OT&E. Before IOT&E, however, the contractor may be tasked with providing training, training aids and handbooks to Service training cadre so they can train the IOT&E users and maintenance personnel. In addition, the contractor must be required to provide sufficient spare parts for the operational maintenance personnel to maintain the test item while undergoing operational testing. These support items must be agreed upon by the PMO and OTA and must contractually bind the contractor. If, however, the contractor will be required to provide higher-level maintenance of the item for the duration of the IOT&E, data collection on those functions will be delayed until a subsequent follow-on operational test and evaluation (FOT&E).

13.2.4 OT&E Funding

The Deputy for T&E provides the program manager estimates of PMO test program costs to conduct IOT&E. This funding includes contractor and government test support for which the program office directly or indirectly will be responsible. Since Service OTAs fund differently, program office funding for conducting OT&E varies. The Deputy for T&E must determine these costs and inform the program manager.

13.2.5 Statement of Work

One of the most important documents receiving input from the Deputy for T&E is the Statement of Work (SOW). He must outline all required or anticipated contractor support for DT&E and OT&E. This document outlines data requirements, contractor-conducted or supported testing, government involvement (access to contractor data, tests and results), operational test support, and any other specific test requirements the contractor will be tasked to perform during the duration of the contract.

13.3 TEST AND EVALUATION MASTER PLAN

The Test and Evaluation Master Plan (TEMP) should be updated regularly by the OTA. The Deputy for T&E is responsible for managing the TEMP throughout the test program. The operational test agency usually is tasked to complete the operational test section of the TEMP and outline their proposed test program through all phases of OT&E. It is important to keep the TEMP updated regularly so that test organizations involved in OT&E understand the scope of their test support. Further, if any upgrades, improvements or enhancements to the fielded weapon system occur, the TEMP must be updated or a new one created to outline new development test (DT) and operational test (OT) requirements.

13.4 PHASES OF OPERATIONAL TEST

For IOT&E, the Deputy for T&E must ensure the contract portions adequately cover the scope of testing as outlined by the operational test agency. The program office may want to provide an observer to represent the Deputy for T&E during the actual testing. The Deputy for T&E involvement in IOT&E will be to monitor and coordinate; he will keep the program manager informed of progress and problems that arise during testing and will monitor required PMO support to the test organization. Also, enough LRIP items must be manufactured to run a complete and adequate OT&E program. For problems requiring program office action, the Deputy for T&E will be the point of contact.

The Deputy for T&E will be concerned with IOT&E of the LRIP units after a limited number are produced. The IOT&E must be closely monitored so that a full-rate production decision can be made. As in the operational assessments, the Deputy for T&E will be monitoring test procedures and results and keeping the program manager informed. If the item does not succeed during IOT&E, a new process of DT&E or a modification may result; and the Deputy for T&E will be involved (as in any new programs inception). If the item passes IOT&E testing and is produced at full rate, the Deputy for T&E will be responsible for ensuring that testing of those production items is adequate to ensure that the end-items physically and functionally resemble the development items.

During FOT&E, the Deputy for T&E monitors the testing; the contractor is usually not involved. The Deputy for T&E should receive any reports generated by the operational testers during this time. Any deficiencies noted during FOT&E should be evaluated by the PMO, which may decide to incorporate upgrades, enhancements or additions to the current system. If the program manager and the engineering section of the program office design or develop modifications that are incorporated into the weapon system design, additional FOT&E may be required.

13.5 FOT&E FOR UPGRADES, ENHANCEMENTS, ADDITIONS

Once a weapon system is fielded, portions of that system may become obsolete, ineffective or deficient and may need replacing, upgrading or enhancing to ensure the weapon system meets current and future requirements. The Deputy for T&E plays a vital role in this process. Modifications to existing weapon systems may be managed as an entire newly acquired weapon system. However, since these are changes to existing systems, the Deputy for T&E is responsible for determining if these enhancements degrade the existing system, are compatible with its interfaces and functions and whether nondevelopment items (NDIs) require retest or the entire weapon system needs reverification. The Deputy for T&E must plan the test program's funding, schedule, test program and contract provisions with these items in mind. A new TEMP may have to be generated or the original weapon system TEMP modified and recoordinated with the test organizations. The design of the DT&E and FOT&E program usually requires coordination with the engineering, contracting and program management sections of the program office.

13.6 TEST RESOURCES

During all phases of OT, the Deputy for T&E must coordinate with the operational testers to ensure they have the test articles needed to accomplish their mission. Test resources will be either contractor provided or government provided. The contractor resources must be covered in the contract, whether in the development contract or the production contract. Government test resources needed are determined by the operational testers. They usually coordinate the test ranges, test support and the user personnel for testing. The program manager programs funding for his support of OT. Funding for Navy operational evaluation (OPEVAL) is identified in the TEMP and funded in the PMO's budget. Other Services allow the OTAs to develop and manage their own budget for operational testing. The OTAs then obligate funds for test ranges, instrumentation, etc., according to their operational test plans.

13.7 SUMMARY

The PMO should be proactive in its relations with the Service operational testing agency. There are many opportunities to educate the OTA on system characteristics and expected performance. Early OTA input to design considerations and requirements clarification can reduce downstream surprises. Operational testing is an essential component of the development and decision-making process. It can be used to facilitate system development or may become an impediment. In many cases, the PMO attitude toward operational testing and the OTA will influence which role the OTA assumes.

- Understand the policies
- Organize for T&E
- Keep system requirements documents current
- Agonize over system thresholds
- Work closely with the operational test director
- Don't forget about operational suitability
- Make final DT&E a rehearsal for IOT&E
- Prepare interfacing systems for your IOT&E
- Manage software testing closely
- Track availability of test resources and test support personnel/facilities

Source: NAVSEA T&E Office

Figure 13-1. Lessons Learned From OT&E

IV
MODULE
Test and Evaluation Planning

Many program managers face several test and evaluation issues that must be resolved to get their particular weapon system tested and ultimately fielded. These issues may include modeling and simulation support, combined and concurrent testing, test resources, survivability and lethality testing, multi-Service testing, or international T&E. Each issue presents a unique set of challenges for the program manager when he develops the integrated strategy for the test and evaluation program.

14

EVALUATION

14.1 INTRODUCTION

This chapter describes the evaluation portion of the test and evaluation process. It stresses the importance of establishing and maintaining a clear audit trail from system requirements through critical issues, evaluation criteria, test objectives and measures of effectiveness to the evaluation. The importance of the use of data from all sources is discussed as are the differences in approaches to evaluating technical and operational data.

14.2 DIFFERENCE BETWEEN "TEST" AND "EVALUATION"

The following distinction has been made between the functions of "test" and "evaluation":

> While the terms "test" and "evaluation" are most often found together, they actually denote clearly distinguishable functions in the RDT&E [research, development, test and evaluation] process.

"Test" denotes the actual testing of hardware/software -models, prototypes, production equipment, computer programs —to obtain data, including software, valuable in developing new capabilities, managing the process, or making decisions on the allocation of resources.

"Evaluation" denotes the process whereby data are logically assembled and analyzed to aid in making systematic decisions. (Reference 10)

To summarize, evaluation is "the review and analysis of qualitative or quantitative data obtained from design review, hardware inspection, testing or operational usage of equipment," (Reference 2).

14.3 THE EVALUATION PROCESS

The evaluation process requires a broad analytical approach with careful focus on the development of an overall test and evaluation (T&E plan that will provide timely answers to critical issues and questions required by decision authorities throughout all the acquisition phases. Evaluations should focus on critical system characteristics; i.e., "those design features that determine how well the proposed concept or system will function in its intended operational environment," (4-C, DODI 5000.2).

A functional block diagram of a generic (i.e., not Service-specific) evaluation process is shown in Figure 14-1. The process begins with the identification of a deficiency or need and the documentation of an operational requirement. It continues with the identification of critical issues that must be addressed to determine the degree to which the system meets user requirements. Objectives and thresholds must then

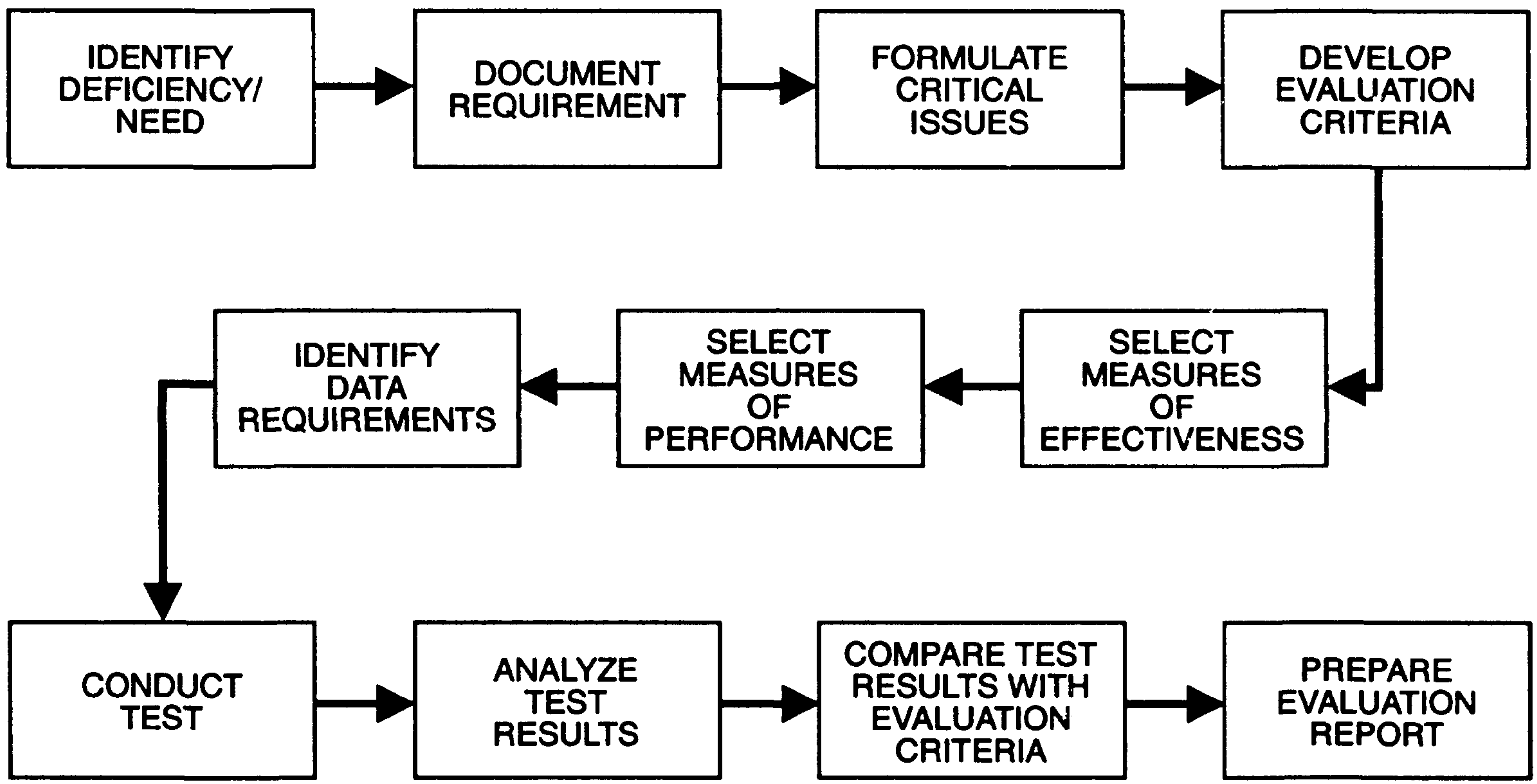


Figure 14-1. Functional Block Diagram of the Evaluation Process

be established to define required performance or supportability parameters and to evaluate progress in reaching them. Test and evaluation analysts then decompose the issues into measurable test elements, conduct the necessary testing, review and analyze the test data, weigh the test results against the evaluation criteria, and prepare an evaluation report for the decision authorities.

14.4 ISSUES AND CRITERIA

Issues are questions regarding a system that require answers during the acquisition process. Those answers may be needed to aid in the development of an acquisition strategy, to refine performance requirements and designs or to support milestone decision reviews. Evaluation criteria are the standards by which accomplishments of required technical and operational effectiveness and/or suitability characteristics or resolution of operational issues may be assessed. The evaluation program may be constructed using a structured approach identifying each issue.

(1) Issue — a statement of the question to be answered;

(2) Scope — detailed conditions and range of conditions that will guide the T&E process for this issue;

(3) Criteria — quantitative or qualitative standards that will answer the issue;

(4) Rationale — full justification to support the selected criteria.

(See Appendices for example.)

14.4.1 System Program Issues/ Critical Issues

System program issues often consist of a hierarchy of critical issues and less significant issues. Critical issues are those questions relating to a system's operational, technical, support or other capability. These issues must be answered before the system's overall worth can be estimated/ evaluated, and they are of primary importance to the decision authority in allowing the system to advance to the next acquisition phase. Critical issues in the TEMP may be derived from the critical system characteristics found in the operational requirement document. The system requirements and baseline documentation will provide many of the performance parameters required to develop the hierarchy of issues.

14.4.2 Evaluation Issues

Evaluation issues are those addressed in technical or operational evaluations during the acquisition process. Evaluation issues can be separated into technical or operational issues and addressed in the Test and Evaluation Master Plan (TEMP).

Technical issues primarily concern technical parameters or characteristics and engineering specifications normally assessed in development testing. Operational issues concern effectiveness and suitability characteristics for functions to be performed by equipment/personnel. They address the system's operational performance when examined in a realistic operational mission environment. Evaluation issues are answered by whatever means necessary (analysis/survey, modeling, simulation, demonstration or testing) to resolve the issue. Issues requiring test data are further referred to as test issues.

14.4.3 Test Issues

Test issues are a subset of evaluation issues and address areas of uncertainty that require test data to resolve the issue adequately. Test issues can be separated into technical issues that are addressed by the development test and evaluation (DT&E) community and operational issues that are addressed by the operational test and evaluation (OT&E) community. Test issues may be divided into critical and noncritical categories. All critical T&E issues, objectives, methodologies and evaluation criteria should be defined during the initial establishment of an acquisition program. Critical issues are documented in the TEMP. These evaluation issues serve to define the testing required for each phase of the acquisition process and serve as the structure to guide the testing program so these data may be compared against performance criteria.

14.4.4 Criteria

Criteria are statements of a system's required technical performance and operational effectiveness, suitability and supportability. Criteria are often expressed as "objectives and thresholds." (Some Services, however, specify performance and supportability requirements exclusively in terms of thresholds and avoid the use of the concept of objectives.) These performance measurements provide the basis for collecting data used to evaluate/answer test issues.

Criteria must be unambiguous and assessable whether stated qualitatively or quantitatively. They may compare the mission performance of the new system to the one being replaced, compare the new system to a predetermined standard, or compare mission performance results using the new system to not having the system. Criteria are the final values deemed necessary by the user. As such, they should be developed in close coordination with the system user, other testers and specialists in all other areas of operational effectiveness and suitability. These values may be changed as systems develop and associated testing and evaluation proceed. Every issue should have at least one criteria that is a concise measure of the function. Values must be realistic and achievable within the state of the art of engineering technology. A quantitative or qualitative criterion should have a clear definition, free of ambiguous or imprecise terminology, such as "adequate," "sufficient" or "acceptable."

14.4.4.1 Test and Thresholds and Objectives

An operational requirement document (ORD) *threshold* performance parameter lists a minimally acceptable requirement or a minimally acceptable level of performance required by a test article or system to provide a system capability that will satisfy the validated mission need. Thresholds are stated quantitatively whenever possible. Specification of minimally acceptable performance in measurable parameters is essential to selecting appropriate measures of effectiveness, which, in turn, heavily influence test design. Thresholds are of value only when they are testable; i.e., actual performance can be measured against them. The function of T&E is to verify the attainment of required thresholds. As stated in OPNAVINST 5000.42C, "T&E is the major control mechanism of the acquisition process. Programs advance from one phase to the next, not by the calendar or planned schedule, but by actual achievement of present thresholds, verified by T&E." (Reference 69)

Objectives are levels of performance (established by the user) above the threshold that, if achieved, will provide measurable benefits of additional operational capability, operations and support. Objectives are not normally addressed by the operational tester, whose primary concern is the requirement.

Going into Milestone II, thresholds and objectives are expanded along with the identification of more-detailed and refined performance capabilities and characteristics resulting from trade-off studies and testing conducted during the Demonstration and Validation Phase. Along with the ORD, they should remain relatively stable through production.

14.5 MEASURES OF EFFECTIVENESS

Requirements, thresholds and objectives established in early program documentation form the basis for evaluation criteria. If program documentation is incomplete, the tester may have to develop evaluation criteria in the absence of specific requirements. Evaluation criteria are associated with objectives, subobjectives and measures of effectiveness (MOEs). For example, an MOE (e.g., airspeed) may have an associated evaluation criterion (e.g., 450 knots) against which the actual performance (e.g., 425 knots) is compared to arrive at a rating. An MOE of a system is a parameter that evaluates the capacity of the system to accomplish its assigned missions under a given set of conditions. They are important because they determine how test results will be judged; and, since test planning is directed toward obtaining these measures, it is important that they be defined early. Generally, the resolution of each critical issue is in terms of the evaluation of some MOE (Reference 116). In this case, the operating, implementing, and supporting commands must agree with the criteria before the test organization makes use of them in assessing test results. Ensuring that MOEs can be related to the user's operational requirements is an important consideration when identifying and establishing evaluation criteria. Testers must ensure that evaluation criteria and MOEs are updated if requirements change. Measures of effectiveness should be so specific that the system's effectiveness during developmental and operational testing can be assessed using the same effectiveness criteria as the Cost and Operational Effectiveness Analysis (4-E, DODI 5000.2).

14.6 EVALUATION PLANNING

14.6.1 Evaluation Planning Techniques

Evaluation planning is an iterative process that requires formal and informal analyses of system operation (e.g., threat environment, system design, tactics and interoperability). Techniques that have been proven effective in evaluation planning include: process analysis, design or engineering analysis, matrix analysis and dendritic analysis (Reference 61).

14.6.1.1 Process Analysis Techniques

Process analysis techniques consist of thinking through how the system will be used in a variety of environments, threats, missions and scenarios in order to understand the events, actions, situations and results that are expected to occur. This technique aids in the identification and clarification of appropriate MOEs, test conditions and data requirements.

14.6.1.2 Design/Engineering Analysis Techniques

Design or engineering analysis techniques are used to examine all mechanical or functional operations that the system has been designed to perform. These techniques involve a systematic exploration of the system's hardware and software components, purpose, performance bounds, manpower and personnel considerations, known problem areas and impact on other components. Exploration of the way a system operates compared to intended performance functions often identifies issues, MOEs, specific data, test events and required instrumentation.

14.6.1.3 Matrix Analysis Techniques

Matrix analysis techniques are useful for analyzing any situation where two classifications must be cross-referenced. For example, a matrix of "types of data" vs. "means of data collection" can reveal not only types of data having no planned means of collection but also redundant or backup collection systems. Matrix techniques are useful as checklists, as organizational tools or as a way of identifying and characterizing problem areas. Matrix techniques are effective for tracing a system's operational requirements through contractual specification documents, issues and criteria to sources of individual data or specific test events.

14.6.1.4 Dendritic Analysis Techniques

Dendritic analysis techniques are an effective way of decomposing critical issues to the point where actual data requirements and test measurements can be identified. In these techniques, issues are successively broken down into objectives, measures of effectiveness, measures of performance and data requirements in a root-like structure as depicted in Figure 14-2. In this approach, objectives are used to clearly express the broad aspects of T&E related to the critical issues and the overall purpose of the test. Measures of effectiveness are developed as subsets of the objectives and are designed to treat specific and addressable parts of the objectives. Each MOE is traceable as a direct contributor to one objective and, through it, is identifiable as a direct contributor to addressing one or more critical issues (Reference 83). Each test objective and MOE is also linked to one or more measures of performance (quantitative or qualitative measures of system performance under specified conditions) that, in turn, are tied to specific data elements. The dendritic approach has become a standard military planning technique.

14.6.2 Sources of Data

As evaluation and analysis planning matures, focus turns toward identifying data sources as a means for obtaining each data element. Initial identification tends to be generic such as: engineering study, simulation, modeling or contractor test. Later identification reflects specific studies, models and/or tests. A data source matrix is a useful planning tool to show where data are expected to be obtained during the T&E of the system.

There are many sources of data that can contribute to the evaluation. Principal sources include: studies and analyses, models, simulations, war games, contractor testing, development testing, operational testing and comparable systems.

14.7 EVALUATING DEVELOPMENT AND OPERATIONAL TESTS

Technical and operational evaluations employ different techniques and have different evaluation criteria. Development test and evaluation is often considered technical evaluation while OT&E addresses the operational aspects of a system. Technical evaluation deals primarily with instrumented tests and statistically valid data. An operational evaluation deals with operational realism and the combat uncertainties (Reference 76). Development test and evaluation uses technical criteria for evaluating system performance. These criteria are usually parameters that can be

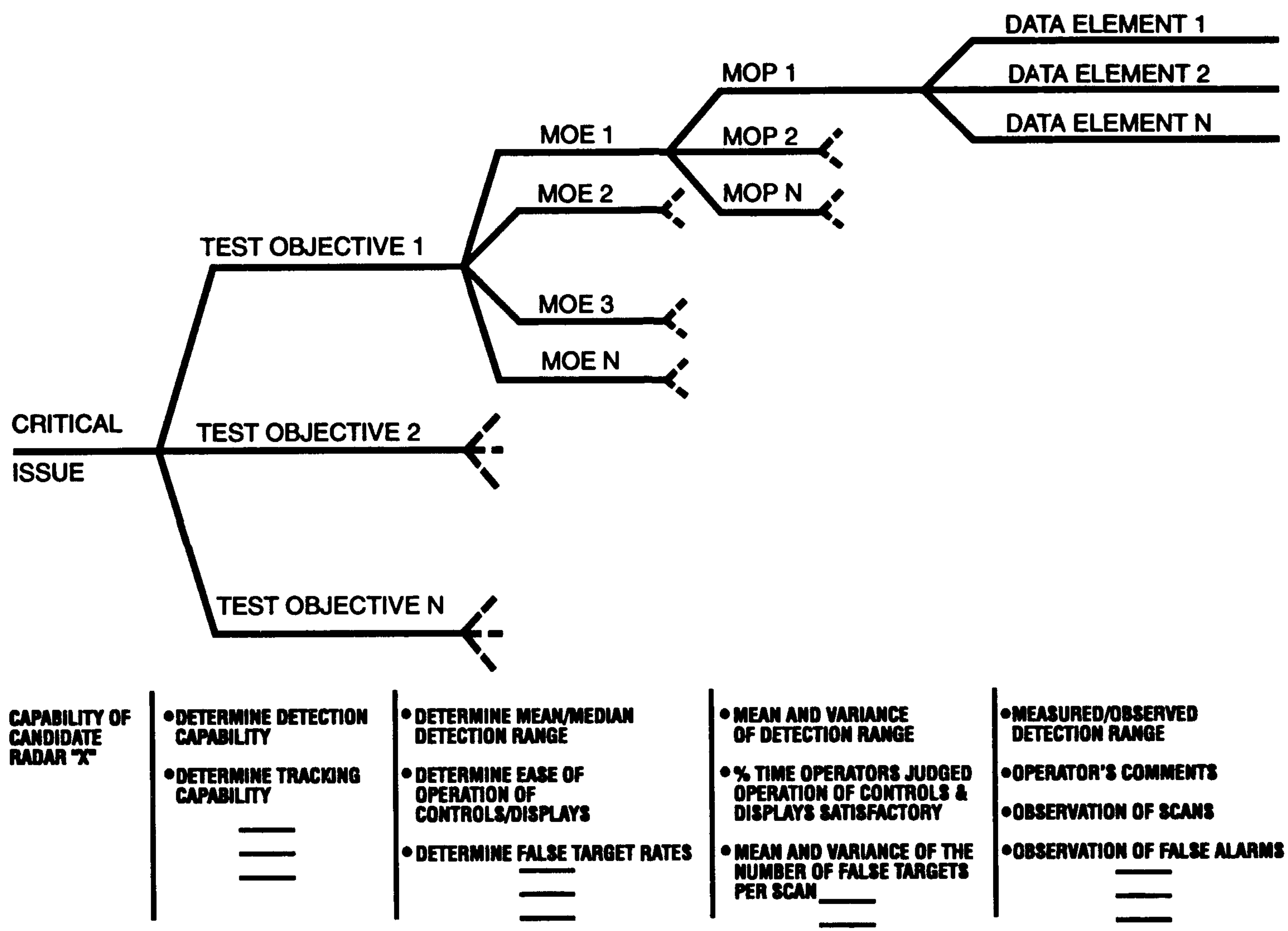


Figure 14-2. Dendritic Approach to Test and Evaluation

measured during controlled DT&E tests. They are particularly important to the developing organization and the contractor but are of less interest to the independent operational tester. The operational tester focuses on issues such as demonstrating target acquisition at useful ranges, air superiority in combat, or the probability of accomplishing a given mission. For example, during DT&E, firing may be conducted on a round-by-round basis with each shot designed to test an individual specification or parameter with other parameters held constant. Such testing is designed to measure the technical performance of the system. In contrast, in OT&E proper technical performance regarding individual specifications/parameters is de-emphasized and the environment is less controlled. The OT&E authority must assess whether, given this technical performance, the weapon system is operationally effective and operationally suitable when employed under realistic combat (with opposing force) and environmental conditions by typical personnel.

The emphasis in development test (DT) is strictly on the use of quantitative data to verify attainment of technical specifications. Quantitative data are usually analyzed using some form of statistics. Qualitative data takes on increasing importance in OT&E when effectiveness and suitability issues are being explored. Many techniques are used to analyze qualitative data. They range from converting expressions of preference or opinion into numerical values to establishing a consensus by committee. For example, a committee may assign values to parameters such as "feel," "ease of use," "friendliness to the user," and "will the user want to use it," on a scale of 1-to-10. Care should be exercised in the interpretation of the results of qualitative evaluations. For instance, when numbers are assigned to average evaluations and their standard deviations, meanings will differ from quantitative data averages and standard deviations.

14.7.1 Technical Evaluation

The Services' materiel development organizations are usually responsible for oversight of all aspects of DT&E including the technical evaluation. The objectives of a technical evaluation are:

- To assist the developers by providing information relative to technical performance; qualification of components; compatibility, inter-operability, vulnerability, lethality, transportability, reliability, availability and maintainability (RAM); manpower and personnel; system safety; integrated logistics support; correction of deficiencies; accuracy of environmental documentation; and refinement of requirements;

- To ensure the effectiveness of the manufacturing process of equipment and procedures through production qualification T&E;

- To confirm readiness for operational test (OT) by ensuring that the system is stressed beyond the levels expected in the OT environment;

- To provide information to the decision authority at each decision point regarding a system's technical performance and readiness to proceed to the next phase of acquisition;

- To determine the system's operability in the required climatic and realistic battlefield environments to include natural, induced, and countermeasure environments (Reference 54).

14.7.2 Operational Evaluation

The independent OT&E authority is responsible for the operational evaluation. The objectives of an operational evaluation are:

- To assist the developers by providing information relative to operational performance; doctrine, tactics, training, logistics; safety; survivability; manpower, technical publications; RAM; correction of deficiencies; accuracy of environmental documentation; and refinement of requirements;

- To assist decision-makers ensure that only systems that are operationally effective and suitable are delivered to the operating forces;

- To provide information to the decision authority at each decision point as to a system's operational effectiveness, suitability and readiness to proceed to the next phase of acquisition;

- To assess, from the user's viewpoint, a system's desirability, considering systems already fielded, and the benefits or burdens associated with the system (Reference 84).

14.8 SUMMARY

A primary consideration in identifying information to be generated by an evaluation program is having a clear understanding of the decisions the information will support. The importance of structuring the T&E program to support the resolution of critical issues cannot be overemphasized. It is the responsibility of those involved in the evaluation process to ensure that the proper focus is maintained on key issues, the T&E program yields information on critical technical and operational issues, all data sources necessary for a thorough evaluation are tapped and evaluation results are communicated in an effective and timely manner. The evaluation process should be evolutionary throughout the acquisition phases.

Table 14-1. Sample Evaluation Plan

CHAPTER 1 INTRODUCTION
- 1.1 Purpose
- 1.2 Scope
- 1.3 Background
- 1.4 System Description
- 1.5 Critical Operational Issues and Criteria (COIC)
- 1.6 Projected Threat
- 1.7 Test and Evaluation Milestones

CHAPTER 2 EVALUATION STRATEGY
- 2.1 Evaluation Concept
- 2.2 Operational Effectiveness
  - 2.2.1 Issue 1
    - 2.2.1.1 Scope
    - 2.2.1.2 Criteria
    - 2.2.1.3 Rationale
    - 2.2.1.4 Evaluation Approach
    - 2.2.1.5 Analysis of MOPs and Data Presentations
      - 2.2.1.5.1 MOP 1
      - through
      - 2.2.1.5.x MOP x
  - 2.2.2 Issue 2
  - through
  - 2.2.m Issue m
- 2.3 Operational Suitability
  - 2.3.1 Issue m+1
  - through
  - 2.3.n Issue n
- 2.4 Data Source Matrix
- 2.5 Description of Other Primary Data Sources
- 2.6 Test Approach
  - 2.6.1 Test Scope
  - 2.6.2 Factors and Conditions
  - 2.6.3 Sample Size and Other Test Design Considerations
  - 2.6.4 Data Authentication Group (DAG) Requirements
- 2.7 Evaluation Data Base Structure
  - 2.7.1 Identification of Required Files
  - 2.7.2 Description of File Relationships
  - 2.7.3 Data Element Definitions

APPENDICES:
- APPENDIX A IOT&E RESOURCE PLAN
- APPENDIX B PATTERN OF ANALYSIS
- APPENDIX C CONTROL CONCEPT
- APPENDIX D DATA COLLECTION CONCEPT
- APPENDIX E DATA REDUCTION CONCEPT
- APPENDIX F QUALITY CONTROL CONCEPT
- APPENDIX G DAG CHARTER AND SOP
- APPENDIX H TRAINING CONCEPT
- APPENDIX I TEST ENVIRONMENTAL ASSESSMENT AND ENVIRONMENTAL IMPACT STATEMENT
- APPENDIX J STATUS OF SUPPORT DOCUMENTS
- APPENDIX K SYSTEM DESCRIPTION
- APPENDIX L SCENARIO
- APPENDIX M INSTRUMENTATION
- APPENDIX N BASELINE CORRELATION MATRIX
- APPENDIX O STRAWMAN INDEPENDENT EVALUATION REPORT
- APPENDIX P GLOSSARY
- APPENDIX Q ABBREVIATIONS

Source: OT&E Methodology Guide, DA Pamphlet 71-3

15

MODELING AND SIMULATION SUPPORT TO T&E

15.1 INTRODUCTION

This chapter discusses the applications of modeling and simulation in test and evaluation (T&E). The need for modeling and simulation has long been recognized, as evidenced by this quotation from the USAF Scientific Advisory Board in June 1965:

> Prediction of combat effectiveness can only be, and therefore must be, made by using the test data in analytical procedures. This analysis usually involves some type of model, simulation, or game (i.e., the tools of operations or research analysis). It is the exception and rarely, that the 'end result' i.e., combat effectiveness, can be deduced directly from test measurements.

In mandating T&E early in the acquisition process (i.e., before Milestone II), DODI 5000.2 encourages the use of modeling and simulation as a source of T&E data. For instance, the Armored Family of Vehicles program used more than 60 models, simulations and other test data to support system concept exploration. The reliance on modeling and simulation by this and other acquisition programs provides the T&E community with valuable information which can increase confidence levels, decrease field test time and costs, and provide data for pretest prediction and post-test validation. The Defense Modeling and Simulation Office (DMSO), working for the Director Defense Research and Engineering, is developing Office of the Secretary of Defense (OSD) guidance on the application of modeling and simulation to the acquisition process.

This chapter discusses using modeling and simulation to increase the efficiency of the T&E process, reduce time and cost, provide otherwise unattainable and immeasurable data, and provide more timely and valid results.

15.2 TYPES OF MODELS AND SIMULATIONS

The term "modeling and simulation" is often associated with huge digital computer simulations; but it also includes manual and man-in-the-loop war games, test beds, hybrid laboratory simulators and prototypes.

A mathematical model is an abstract representation of a system that provides a means of developing quantitative performance requirements from which candidate designs can be developed. Static models are those that depict conditions of state while dynamic models depict conditions that vary with time, such as the action of an autopilot in controlling an aircraft. Simple dynamic models can be solved analytically, and the results represented graphically.

According to a former Director, Defense Test and Evaluation (Reference 121), simulations used in T&E can be divided into three categories:

> ...computer simulations, system test beds, and system prototypes. Computer simulations are strictly mathematical representations of systems and do not employ any actual hardware. They may, however, incorporate some of the actual software that might be used in a system. Early in a system's life cycle, computer simulations can be expected to provide the most system evaluation information. In many cases, computer simulations can be readily developed as modifications of existing simulations for similar systems. For example, successive generations of AIM-7 missile simulations have been effectively used in test and evaluation.

A system test bed usually differs from a computer simulation as it contains some, but not necessarily all, of the actual hardware that will be a part of the system. Other elements of the system are either not incorporated or, if they are incorporated, are in the form of computer simulations. The system operating environment (including threat) may either be physically simulated, as in the case of a flying test bed, or computer simulated, as in the case of a laboratory test bed. Aircraft cockpit simulators used to evaluate pilot performance are good examples of system test beds. As development of a system progresses, more subsystems become available in hardware form. These subsystems can be incorporated into system test beds that typically provide a great deal of the system evaluation information used during the middle part of a system's development cycle.

The third type of simulation used in T&E is the system prototype. Unlike the system test bed, all subsystems are physically incorporated in a system prototype. The system prototype may closely represent the final system configuration, depending on the state of development of the various subsystems that compose it. Preproduction prototype missiles and aircraft used in operational testing by the Services are examples of this class of simulation. As system development proceeds, eventually all subsystems will become available for incorporation in one or more system prototypes. Hardware-in-the-loop (HWIL) simulators or full-up system simulators may provide the foundation for continuous system testing and improvement. These simulators can provide the basis for transitioning hardware and software into operationally realistic training devices with mission rehearsal capabilities. Operational testing of these prototypes frequently provides much of the system evaluation information needed for a decision on full-scale production and deployment.

As illustrated in Figure 15-1, there is a continuous spectrum of simulation types with the pure computer simulation at one end and the pure hardware prototype at the other end.

15.3 VALIDITY OF MODELING AND SIMULATION

Simulations are not a substitute for live testing. There are many things that cannot be adequately simulated by computer programs; among them are the process of decision and the proficiency of personnel in the performance of their functions. Therefore, models and simulations are not a total substitution for physical tests and evaluations. Simulations, manual and computer-designed, can complement and increase the validity of live tests and evaluations by proper selection and application. Figure 15-2 contrasts the test criteria that are con-

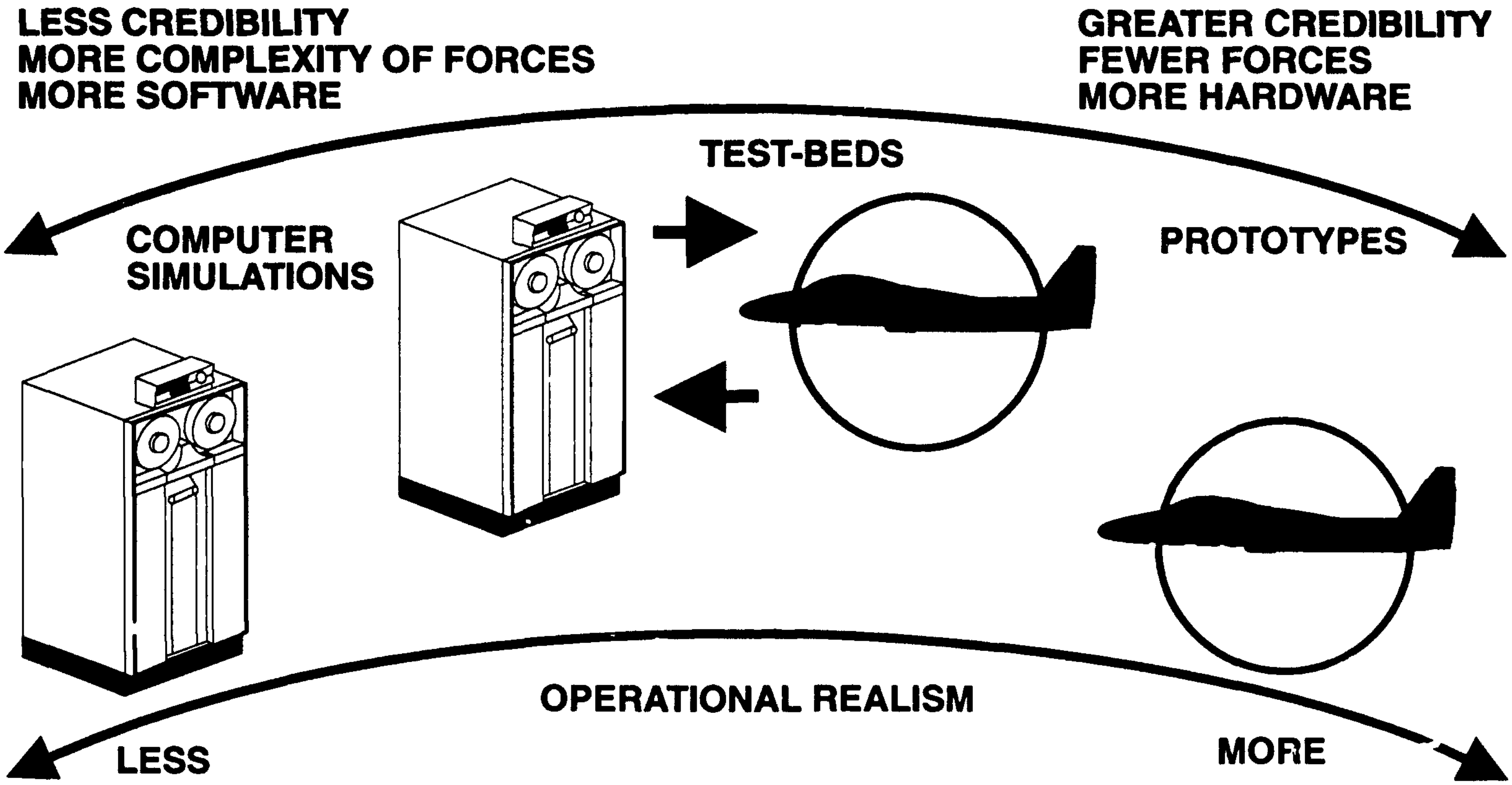


Figure 15-1. The Simulation Spectrum

ducive to modeling and simulation vs. physical testing. Careful selection of the simulation, knowledge of its application and operation and meticulous selection of input data will produce representative and valid results.

The important element in using a simulation is to select one that is representative and either addresses, or is capable of being modified to address, the level of detail (issues, thresholds and objectives) under investigation.

15.4 SUPPORT TO TEST DESIGN AND PLANNING

Modeling and simulation can assist in the T&E planning process and can reduce the cost of testing. Areas of particular application include scenario development and the timing of test events; the development of objectives, essential elements of analysis, and measures of effectiveness; the identification of variables for control and measurement; and the development of data collection, instrumentation and data analysis plans. For example, using simulation, the test designer can examine system sensitivities to changes in variables to determine the critical variables and their ranges of values to be tested. He can also predict the effects of various assumptions and constraints and evaluate candidate measures of effectiveness to help formulate the test design.

Caution must be exercised when planning to rely on simulations to obtain test data as they tend to be expensive to develop or modify, difficult to integrate with data from other sources, and often do not provide the level of realism required for operational tests. Although simulations are not a "cure-all," they should be used whenever feasible as another source of data for the evaluator to consider during the test evaluation.

Computer simulations may be used to test the planning for an exercise. By setting up and running the test exercise in a simulation, the timing and scenario may be tested and validated. Critical events may include interaction of various forces that test the measures of effectiveness and, in turn, test objectives. Further, the simulation may be used to verify the statistical test design and the instrumentation, data collection, and data analysis plans. Essentially, the purpose of computer simulation in pretest planning is to preview the test to evaluate ways to make test results more effective. Pretesting attempts to optimize test results by pointing out potential trouble spots. It constitutes a test setup analysis, which can encompass a multitude of areas.

As an example of simulations used in test planning, consider a model that portrays aircraft vs. air defenses. The model can be used to replicate typical scenarios and provide data on the number of engagements, air defense systems involved, aircraft target, length and quality of the engagement, and a rough approximation of the success of the mission (i.e., if the aircraft made it to the target). With such data available, a data collection plan can be developed to specify, in more detail, when and where data should be collected, from which systems, and in what quantity. The results of this analysis impact heavily on long lead-time items such as data collection devices and data processing systems. The more specificity available, the fewer the number of surprises that will occur downstream. As tactics are decided upon and typical flight paths are generated for the scenario, an analysis can be prepared on the flight paths over the terrain in question; and a determination can be made regarding whether the existing instrumentation can track the numbers of aircraft involved in their maneuvering envelopes. Alternative site arrangements can be examined and trade-offs can

THE SIMULATION SPECTRUM		
CRITERIA	VALUES CONDUCIVE TO:	
	PHYSICAL TESTING	MODELING AND SIMULATION
TEST SAMPLE SIZE/NUMBER OF VARIABLES	SMALL/FEW	LARGE/MANY
STATUS OF VARIABLES/UNKNOWNS	CONTROLLABLE	UNCONTROLLABLE
PHYSICAL SIZE OF PROBLEM	SMALL AREA/ FEW PLAYERS	LARGE AREA/ MANY PLAYERS
AVAILABILITY OF TEST EQUIPMENT	AVAILABLE	UNAVAILABLE
AVAILABILITY OF TEST FACILITIES	RANGES, OTHER TEST AVAILABLE	BENCHMARKED, VALIDATED COMPUTER MODELS AVAILABLE
TYPES OF VARIABLES/UNKNOWNS	SPATIAL/TERRAIN	LOW IMPORTANCE OF SPATIAL/TERRAIN
DIPLOMATIC/POLITICAL FACTORS	CONVENTIONAL CONFLICTS	NUCLEAR OR CHEMICAL CONFLICTS

Figure 15-2. Values of Selected Criteria Conducive to Modeling and Simulation

be made between the amount of equipment to be purchased and the types of profiles that can be tracked for this particular test. Use of such a model can also highlight numerous choices available to the threat air defense system in terms of opportunities for engagement and practical applications of doctrine to the specific situations.

15.5 SUPPORT TO TEST EXECUTION

Simulations can be useful in test execution and dynamic planning. With funds and other restrictions limiting the number of times that a test may be repeated and each test conducted over several days, it is mandatory that the test director exercises close control over the conduct of the test to ensure the specific types and quantities of data needed to meet the test objectives are being gathered and to ensure adequate safety. He must be able to make minor modifications to the test plan and scenario to force achievement of these goals. This calls for a dynamic (quick-look) analysis capability and a dynamic planning capability. Simulations may contribute to this capability. For example, using the same simulation(s) as used in pretest planning, the tester could input data gathered during the first day of the exercise to determine the adequacy of the data to fulfill the test objectives. Using this data, the entire test could be simulated. Projected inadequacies could be isolated, and the test plans could be modified to minimize the deficiencies.

Simulations may also be used to support test control and to ensure safety. For example, during missile test firings at White Sands Missile Range (WSMR), aerodynamic simulations of the proposed test were run on a computer during actual firings so that real-time missile position data could be compared continuously to the simulated missile position data. If any significant variations occurred and if the range safety officer was too slow (both types of position data were displayed on plotting boards), the computer issued a destruct command.

Simulations can be used to augment tests by simulating nontestable events and scenarios. Although operational testing should be accomplished in as realistic an operational environment as possible, pragmatically some environments are impossible to simulate for safety or other reasons. Some of these include the environment of a nuclear battlefield, to include the effects of nuclear bursts on friendly and enemy elements. Others include two-sided live firings and adequate representation of other forces to ascertain compatibility and interoperability data. Instrumentation, data collection and data reduction of large combined armed forces (e.g., brigade, division and larger-sized forces) become extremely difficult and costly. Simulations are not restricted by safety factors and can realistically replicate many environments that are otherwise unachievable in an operational test and evaluation (OT&E)—nuclear effects, large combined forces, electronic countermeasures (ECM), electronic counter-countermeasures (ECCM) and many engagements.

Usually, insufficient units are available to simulate the organizational relationships and interaction of the equipment with its operational environment, particularly during the early OT&E conducted using prototype or pilot production-type equipment. Simulations are not constrained by these limitations. Data obtained from a limited test can be plugged into a simulation that is capable of handling many of the types of equipment being tested. It can interface them with other elements of the blue forces and operate them against large elements of the red forces to obtain interactions.

End-item simulators can be used to evaluate design characteristics of equipment and can be used to augment the results obtained using prototype or pilot production-type equipment that is representative of the final item. The simulator may be used to expand test data to obtain the required iterations or to indicate that the human interface with the prototype equipment will not satisfy the design requirements.

It is often necessary to use substitute or surrogate equipment in testing; e.g., American equipment is used to represent threat-force equipment. In some cases the substitute equipment may have greater capabilities than the real equipment, in other cases it may have less. Simulations are capable of representing the real characteristics of equipment and, therefore, can be used as a means of modifying raw data collected during the test to reflect real characteristics.

As an example, if the substitute equipment is an AAA gun with a tracking rate of 30 degrees per second and the equipment for which it is substituted has a tracking rate of 45 degrees per second, the computer simulation could be used to augment the collected, measured data by determining how many rounds could have been fired against each target or whether targets that were missed because the tracking rate was too slow could have been engaged by the actual equipment. Consideration of other differing factors simultaneously could have a plus or minus synergistic effect on test results.

15.6 SUPPORT TO ANALYSIS AND TEST REPORTING

Modeling and simulation may be used in post-test analysis to extend and generalize results and to extrapolate to other conditions. The difficulty of instrumenting and controlling large exercises and collecting and reducing the data and resource costs, to some degree, limits the size of T&E. This makes the process of determining the suitability of equipment to include compatibility, interoperability, organization, etc., a difficult one. To a large degree the limited interactions, interrelationships and compatibility of large forces may be supplemented by using actual data collected during the test and applying it in the simulation.

Simulations can be used to extend test results, save considerable energy (fuel and manpower), and save money by reducing the need to repeat data points to improve the statistical sample or to determine overlooked or directly unmeasured parameters. Sensitivity analyses can be run using simulations to evaluate the robustness of the design.

In analyzing the test results, data can be compared to the results predicted by the simulations used early in the planning process. Thus, the simulation is validated by the actual live test results, but the test results are also validated by the simulation.

15.7 SUMMARY

Modeling and simulation in T&E can be used for concept evaluation, extrapolation, isolation of design effects, efficiency, representation of complex environments, and overcoming inherent limitations in actual testing. The use of modeling and simulation can validate test results, increase confidence levels, reduce test costs and provide opportunities to shorten the overall acquisition cycle by providing more data earlier for the decision-maker.

16

TEST RESOURCES

16.1 INTRODUCTION

This chapter describes the various types of resources available for testing, explains test resource planning in the Services, and discusses the ways in which test resources are funded.

According to DOD 5000.2-M, the term "test resources" is a collective term that encompasses elements necessary to plan, conduct, collect and analyze data from a test event or program. These elements include: funding (to develop new resources or use existing ones), manpower for test conduct and support, test articles, models, simulations, threat simulators, surrogates, replicas, test-beds, special instrumentation, test sites, targets, tracking and data acquisition instrumentation, equipment (for data reduction, communications, meteorology, utilities, photography, calibration, security, recovery, maintenance and repair), frequency management and control, and base/facility support services. "Testing shall be planned and conducted to take full advantage of existing investment in DOD ranges, facilities, and other resources, whenever practical, unless otherwise justified in the Test and Evaluation Master Plan," (part 8, DODI 5000.2).

Key DOD test resources are in great demand by competing acquisition programs. Often special, unique or one-of-a-kind test resources must be developed specifically for the test program. It is imperative that the requirements for these test resources be identified early in the acquisition process so adequate funding can be allotted for their development, and they will be available when the test is scheduled.

16.2 OBTAINING TEST RESOURCES

16.2.1 Major Range and Test Facility Base

All Services operate ranges and test facilities for test, evaluation and training purposes. Twenty-one of these activities constitute the DOD Major Range and Test Facility Base (MRTFB). This MRTFB is described as "a national asset which shall be sized, operated, and maintained primarily for DOD test and evaluation support missions, but also is available to all users having a valid requirement for its capabilities. The MRTFB consists of a broad base of T&E [test and evaluation] activities managed and operated under uniform guidelines to provide T&E support to DOD Components responsible for developing or operating materiel and weapon systems," (Reference 21). The list of MRTFB activities and their locations are shown on Figure 16-1. Summaries of the capabilities of each of these activities (with points of contact listed for further information) may be found in DOD 3200.11-D.

The MRTFB facilities are available for use by all the Services, other U.S. government

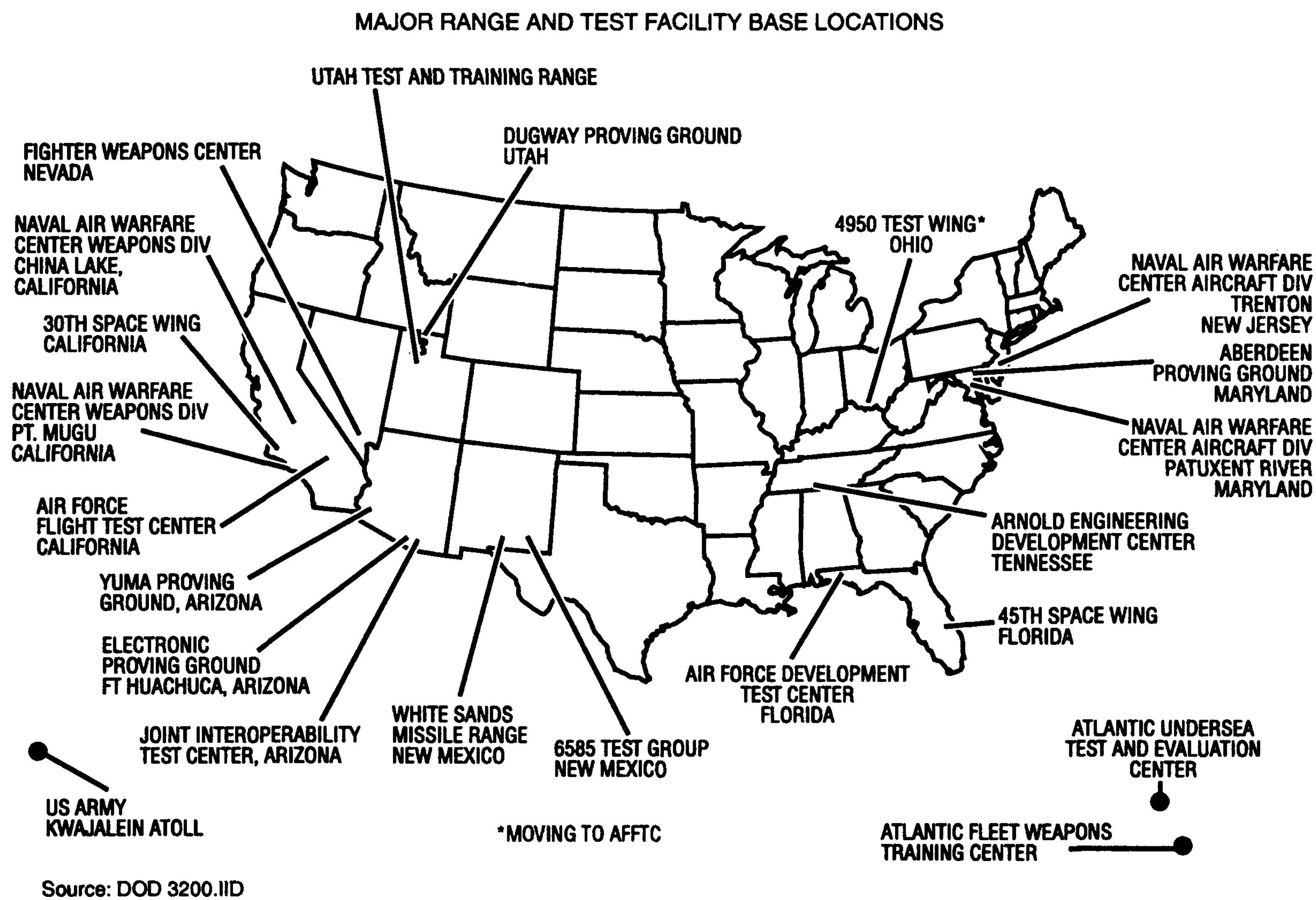


Figure 16-1. DOD Major Range and Test Facility Base

agencies and, in certain cases, allied foreign governments and contractor organizations. Scheduling is based on a priority system; and costs for usage are billed uniformly, as stated in DODD 3200.11. The Deputy Director, Test Facilities and Resources (DTE), sets policy for the composition, use and test program assignments of the MRTFB. In turn, the individual Services must fund, manage and operate their activities. They are reimbursed for direct costs by each user of the activity.

The DOD components wishing to use an MRTFB activity must provide timely and complete notification of their requirements, such as special instrumentation or ground-support equipment requirements, to the particular activity using the documentation formats prescribed by Document 501-84, *Universal Documentation System Handbook,* issued by the Range Commanders Council. The requirements must be stated in the Test and Evaluation Master Plan (TEMP) discussed below. Personnel at the MRTFB activity will coordinate with and assist prospective users with their T&E planning, to include conducting trade-off analyses and test scenario optimization based on test objectives and test support capabilities.

16.2.2 Project Reliance

In response to a stated need to consolidate DOD activities (Defense Management Review Directive 922), the Office of the Secretary of Defense (OSD) T&E organizations have initiated a process to review and centralize various types of system testing infrastructures at designated Service test facilities. Project Reliance is focused on more economical operations, allocating scarce funds for modernization and eliminating unwarranted duplication. The Defense Test and Evaluation Steering Group (DTESG) provides oversight guidance and approval of the Joint Commanders Group (T&E) (JCG(T&E)) recommendations. Various technically oriented panels (Multi-Service Test Investment Resource Committee (MSTIRC)) review Service test methodology areas and develop supporting studies to identify the Service test facility to host a particular type of system test activity (i.e., anechoic chambers, gun munitions testing, air breathing engines, etc.). The MSTIRC recommendations are reviewed by the JCG(T&E) and forwarded to the DTESG for final approval. This means all Services will be expected to use the designated test facility for that type of testing. Test planners must consider Project Reliance agreements when identifying future test sites; this will necessitate more cross-Service test support.

16.2.3 Service Test Facilities

There are other test resources available besides MRTFB. The tester can determine resources available by contacting his/her Service headquarters staff element or if within the Army, by consulting documents such as the Army Test and Evaluation Command (TECOM) Test Facilities Register, the Operational Test and Evaluation Command (OPTEC) Operational Test Instrumentation Guide and other Army test agency and range documents. Information on specific Navy test resources is found in user manuals published by each range and the Commander Operational Test and Evaluation Force (COMOPTEVFOR) catalog of available support.

16.3 TEST RESOURCE PLANNING

The development of special test resources to support a weapon system test can be costly and time-consuming. This, coupled with the competition for existing test resources and facilities, requires that early planning be accomplished to determine all test resource requirements for weapon system T&E. The tester must use government facilities whenever possible instead of funding construction of contractor test capabilities.

Problems associated with range and facility planning are that major systems tend to get top priority; i.e., B-1B, M-1, etc. Range schedules are often in conflict due to system problems, which cause schedule delays during testing; and there is often a shortage of funds to complete testing.

16.3.1 TEMP Requirements

The program manager must state all key test resource requirements in the TEMP and must include items such as unique instrumentation, threat simulators, surrogates, targets and test articles. Included in the TEMP are a critical analysis of anticipated resource shortfalls, their effect on system T&E and plans to correct resource deficiencies. As the preliminary TEMP must be prepared for Milestone I, initial test resource planning must be accomplished during the Concept Exploration and Definition Phase. Refinements and reassessments of test resource requirements are included in each TEMP update. The required content of the test resource summary section of the TEMP is in Part V - Test and Evaluation Resource Summary, DOD 5000.2-M.

16.3.2 Service Test Resource Planning

More-detailed listings of required test resources are generated in conjunction with the detailed test plans written by the materiel developer and operational tester. These test plans describe test objectives, measures of effectiveness (MOEs), scenarios and specific test resource requirements.

16.3.2.1 Army Test Resource Planning

In the Army, the tester prepares input to the TEMP and the Test and Evaluation Plan (TEP), the primary planning documents for operational test and evaluation (OT&E) of the weapon system. These documents should be prepared early in the acquisition cycle (at the beginning of the Concept Demonstration and Validation Phase). They describe the entire T&E approach including critical issues, test methodology, measures of effectiveness and all necessary test resources. The TEMP and TEP provide the primary input to the Outline Test Plan (OTP), which contains a detailed description of each identified required test resource, where and when it is to be provided, and the providing organization.

The tester must coordinate the OTP with all major commands or agencies expected to provide test resources. Then, the OTP is submitted to the Resource Management Division, HQ, OPTEC, for review by the Test Schedule and Review Committee (TSARC) and for incorporation into the Army's Five-Year Test Program (FYTP). The initial OTP for each test should be submitted to the TSARC as soon as testing is identified in the TEMP. Revised OTPs are submitted as more information becomes available or requirements change, but a final comprehensive version of the OTP should be submitted at least 18 months before the resources are required.

The TSARC is responsible for providing high-level, centralized management of T&E resource planning. The TSARC is chaired by the Commanding General OPTEC and consists of a general officer or equivalent representatives from the Army staff and major commands. The TSARC meets semi-annually to review all OTPs, resolve conflicts and coordinate all identified test resource requirements for inclusion in the FYTP. The FYTP is a formal resource tasking document for current and near-term tests and a planning document for tests scheduled for the out-years. All OTPs are reviewed during the semiannual reviews to ensure that any refinements or revisions are approved by the TSARC and reflected in the FYTP. The FYTP is produced as a hard-copy by OPTEC.

The TSARC-approved OTP is a tasking document by which the tester requests Army test resources. The TSARC coordinates resource requests, sets priorities, resolves conflicts and schedules resources. The resultant FYTP, when approved by the Deputy Chief of Staff for Operations and Plans (DCSOPS), HQ DA, is a formal tasking document that reflects the agreements made by the resource providers (Army Materiel Command (AMC), Training and Doctrine Command (TRADOC), Forces Command (FORSCOM), etc.) to make the required test resources available to the designated tests. If test resources from another Service, a non-DOD governmental agency (such as the Department of Energy (DOE) or NASA) or a contractor are required, the request is coordinated by the OPTEC Resource Management Division. For example, the request for a range must be made at least two years in advance to ensure availability. However, due to the long lead time required to schedule these non-Army resources, their availability cannot be guaranteed if testing is delayed or retesting is required. The use of resources outside the U.S., such as in Canada, Germany or other NATO countries, is also handled by OPTEC.

16.3.2.2 Navy Test Resource Planning

In the Navy, the developing agency and the operational tester are responsible for identifying the specific test resources required in testing the weapon system. In developing requirements for test resources, the program manager (PM) and operational test director (OTD) refer to documents such as the Mission Need Statement (MNS), Integrated Program Summary (IPS), Navy Decision Coordinating Paper (NDCP), Operational Requirement Document (ORD), threat assessments, Office of the Chief of Naval Operations Instruction (OPNAVINST) 5000.42D (Test and Evaluation), and the OTD Guide (Commander, Operation Test and Evaluation Force (Navy) (COMOPTEVFOR) Instruction 3960.1D). Upon Chief of Naval Operations (CNO) approval, the TEMP becomes the controlling management document for all T&E of the weapon system. It constitutes direction by the CNO to conduct the T&E program defined in the TEMP, including the commitment of research, development, test and evaluation (RDT&E) financial support and of fleet units and schedules. It is prepared by the PM, who is provided OT&E input by the COMOPTEVFOR Operational Test Director. The TEMP defines all T&E (development test and evaluation (DT&E), OT&E and production acceptance test and evaluation (PAT&E)) to be conducted for the system and describes, in as much detail as possible, the test resources required.

The Navy uses its operational naval forces to provide realistic T&E of new weapon systems. Each year, the CNO (N-091) compiles all Fleet support requirements for RDT&E program support from the TEMPs and publishes the CNO Long-Range RDT&E Support Requirements document for the budget and out-years. In addition, a quarterly forecast of support requirements is published approximately five months before the Fleet Employment Scheduling Conference for the quarter in which the support is required. These documents summarize OT&E requirements for Fleet services and are used by the Fleet for scheduling services and out-year budget projections.

Requests for use of range assets are usually initiated informally with a phone call from the PM and/or OTD to the range manager and followed by formal documentation. Requests for Fleet support are usually more formal. The COMOPTEVFOR, in coordination with the PM, forwards the TEMP and a Fleet RDT&E Support Request to the CNO. Upon approval of the request, the CNO tasks the Fleet Commander in Chief (CINC) by letter or message to coordinate with OPTEVFOR to provide the requested support.

Use of most Navy ranges must be scheduled at least a year in advance. Each range consolidates and prioritizes user requests, negotiates conflicts and attempts to schedule range services to satisfy all requests. If the desired range services cannot be made available when required, the test must wait; or the CNO resolves the conflict. Because ranges are fully scheduled in advance, it is difficult to accommodate a test that is delayed or requires additional range time beyond that originally scheduled. Again, the CNO can examine the effects of delays or retest requirements and issue revised priorities, as required.

Requests for use of non-Navy OT&E resources are initiated by COMOPTEVFOR. The Operational Test and Evaluation Force (OPTEVFOR) is authorized direct liaison with other Service-independent operational test agencies (OTAs) to obtain OTA-controlled resources. Requests for other government-owned resources are forwarded to the CNO (N-091) for formal submission to the Service Chief (for Service assets) or to the appropriate government agency (e.g., DOE or NASA). Use of contractor resources is usually handled by the PM, although contractor assets are seldom required in OT&E, since the Fleet is used to provide an operational environment. Requests for use of foreign ranges are handled by the N-091 Assistant for International Research and Development (R&D).

16.3.2.3 Air Force Test Resource Planning

The test resources required for T&E of an Air Force weapon system are identified in detail in the Test Program Outline (TPO), which is prepared by the responsible Air Force T&E organization. In general, the Air Force Operational Tests and Evaluation Center (AFOTEC) is the test organization for OT&E programs; it obtains support from a Service major command test agency for nonmajor programs, with AFOTEC directing and providing assistance, as required.

During the Advanced Planning Phase of a weapon system acquisition (five to six years before OT&E), AFOTEC prepares the OT&E section of the first full TPO, coordinates the TPO with all supporting organizations and assists the resource manager (RM) in programming required resources. The resource requirements listed in the Resource Information Network TPO are developed by the test manager, resource manager and test support group, using sources such as the ORD and threat assessments. The TPO should specify, in detail, all the resources necessary to successfully conduct a test when it is entered in the Test Resource Information Management System (TRIMS).

The TPO is the formal means by which test resource requirements are communicated to the Air Staff and to the appropriate commands and agencies tasked to supply the needed resources. Hence, if a required resource is not specified in the TPO, it is likely the resource will not be available for the test. The TPO is revised and updated on a continuous basis, since the test resource requirements become better defined as the OT&E plans mature. The initial TPO serves as a baseline for comparison of planned OT&E resources with actual expenditures. Comparisons of the initial TPO with subsequent updates provide an audit trail of changes in the test program and its testing requirements. The AFOTEC maintains all TPOs on TRIMS; this permits immediate response to all queries regarding test resource requirements.

The AFOTEC/RM consolidates the resource requirements from all TPOs coordinating with participating and supporting organizations and agencies outside AFOTEC. Twice yearly, the RM office prepares a draft of the USAF Program for Operational Test (PO). The PO is a master planning and programming document for resource requirements for all HQ USAF-directed OT&E and is distributed to all concerned commands, agencies and organizations for review and coordination. It is then submitted to the Air Staff for review and approval by the Operational Resource Management Assessment System for Test and Evaluation (ORMAS/TE), which operates under the authority of HQ AF/TE. The ORMAS Board is composed of HQ USAF action officers and senior officers from major commands (MAJCOMs) and agencies involved in OT&E; it meets to resolve impacts and conflicting requirements at the appropriate Air Staff level. Through the ORMAS process, HQ USAF approves the PO, which becomes a directive to participants for planning, programming and budgeting actions. Agreements made among ORMAS participants regarding TPO and PO resource requirements are considered binding.

All requests for test resources are coordinated by HQ AFOTEC as part of the TPO preparation process. When a new weapon system development is first identified, AFOTEC provides a test manager (TM) who begins long-term OT&E planning. The TM begins identifying needed test resources, such as instrumentation, simulators and models, and works with the resources directorate to obtain them. If the required resource does not belong to AFOTEC, it will negotiate with the commands having the resource. In the case of models and simulators, AFOTEC surveys what is available, assesses credibility, and then coordinates with the owner or developer to use it. The Joint Technical Coordinating Group publishes a document on electronic warfare (EW) models.

Range scheduling should be done early. At least a year is required, but often a test can be accommodated with a few months' notice if there is no requirement for special equipment or modifications to be provided at the range. Some of the Air Force ranges are scheduled well in advance and cannot accommodate tests that encounter delays or retest requirements.

The resource manager attempts to resolve conflicts among various systems competing for scarce test resources and elevates the request to the Commander, AFOTEC, if necessary. Decisions on resource utilization and scheduling are based on the weapon system's assigned priority.

The resource manager and the test manager also arrange for use of the resources of other Services, non-DOD government agencies and contractors. Use of non-U.S. resources, such as a Canadian range, are coordinated by AF/TE and based on formal Memoranda of Understanding (MOU). The USAFE/DOQ handles requests for European ranges. Use of a contractor-owned resource, such as a model, is often obtained through the System Program Office (SPO) or a general support contract.

16.4 TEST RESOURCE FUNDING

The Future Years Defense Program (FYDP), incorporating a biennial budgeting process, is the basic DOD programming document that records, summarizes and displays Secretary of Defense (SECDEF) decisions. In the FYDP, costs are divided into three categories for each acquisition program element: research and development costs, investment costs and operating costs. The Congress appropriates to the Office of Management and Budget (OMB), and OMB apportions funding through the SECDEF to the Services and to other defense agencies. The Services and defense agencies then allocate funds to others (claimants, subclaimants, administering offices, commanding generals, etc.).

The Planning, Programming, and Budgeting System (PPBS) is a DOD internal system used to develop input to the Congress for each year's budget while developing future-year budgets. The PPBS is calendar oriented. There are concurrent two-year PPBS cycles ongoing at one time. These cycles are: planning, programming and budgeting. At any one time there are three budgets being worked by the Services. The current two-year budget is being executed. The next six years of defense planning is being programmed, and long-range program plans and planning guidance are being reviewed for updating.

There are six types of funding in the PPBS: research funding for maintaining the technology base; exploratory development funding for conducting the Concept Exploration and Definition Phase; advanced development funding for conducting both the Concept Exploration and Definition Phase and the Demonstration and Validation Phase; engineering development funding for conducting the Engineering and Manufacturing Development Phase; operational systems development funding for conducting the Production and Deployment Phase; and RDT&E management and support funding, which is used throughout the development and production cycle until the system is operationally deployed when operations and maintenance (O&M) funding is used. The RDT&E appropriation funds the costs associated with research and development, including test items, DT&E and test support of OT&E of the system or equipment and the test items.

The funding that is planned, programmed and budgeted through the PPBS cycle is not always the same funding amount that the Congress appropriates or the PM receives. If the required funding for a test program is not authorized by the Congress, the PM has four ways to react. The PM can submit a supplemental budget (for unfunded portions of the program), request deficiency funding (for unforeseen program problems) or use transfer authority (from other programs within the Service); or the PM can try to reprogram the needed funds (to restructure the program).

Generally, testing that is accomplished for a specific system before the production decision is funded from RDT&E appropriations; and testing that is accomplished after the production decision is funded from other procurement or operations and maintenance appropriations. Testing of product improvements, block upgrades and major modifications is funded from the same appropriations as the program development. Follow-on Test and Evaluations (FOT&E) are normally funded from O&M funds.

Funding associated with T&E (including instrumentation, targets and simulations) are identified in the system acquisition cost estimates, Service acquisition plans and the TEMP. General funding information for development and operational tests follows:

<u>Development Test Funding</u>. *Funds required* to conduct engineering and development tests are programmed and budgeted by the materiel developer, based upon the requirements of the TEMP. These costs may in-

clude, but are not limited to, procuring test samples/prototypes; support equipment; transportation costs; technical data; training of test personnel; repair parts; and test-specific instrumentation, equipment and facilities. The DT&E funds are expended for contractor and government developmental test activities.

The Service PM may be required to pay for the use of test resources, such as the MRTFB, and for the development of specialized resources needed specifically for testing the weapon system being developed.

Operational Test (OT) Funding. Funds required to conduct OT are usually programmed and budgeted by the Service operational test agency or organization. The funds are programmed in the Service's long-range test program, and the funds requirements are obtained from the test resourcing documentation and TEMP.

16.4.1 Army Funding

Test resources are developed and funded under various Army appropriations. The Army Materiel Command and its commodity commands provide test items, spare parts, support items (such as diagnostic equipment) and ammunition. Soldiers, ranges, fuel, test support personnel and maneuver areas are provided by TRADOC or FORSCOM. The weapon system PM uses RDT&E funds to reimburse these supporting commands for costs directly related to his test. The weapon system materiel developer is also responsible for funding the development of new test resources specifically needed to test the weapon system. Examples of such special-purpose resources include models, simulations, special instrumentation and test equipment, range modifications, EW simulators and, sometimes, threat simulators. Although the Army has a separate budget and development plan for threat simulators, the Army Development and Acquisition of Threat Simulators (ADATS) program, many weapon system developers still have to fund the cost of new threat systems that are specifically needed to test their weapon system. Army OPTEC is funded through its own program element and has direct control of OT&E funds for all programs. Funding requirements are developed in consonance with the Outline Test Plan.

16.4.2 Navy Funding

In the Navy, the weapon system PM is responsible for funding the development of all required test-specific resources from the program's RDT&E funds. These resources include test articles, expendables, one-of-a-kind targets, data collection/reduction and instrumentation. The development of generic test resources that can be used in OT&E of multiple weapon systems such as targets, threat simulators and range capabilities, is funded from OPNAV generic accounts (such as target development) and not from weapon systems RDT&E. The PM's RDT&E funds pay for all DT and OT through OPEVAL. The PM pays for all post-production OT with program funds.

16.4.3 Air Force Funding

In the Air Force, direct-cost funding requires that test-peculiar (direct) costs associated with a particular test program be reimbursed by the System Program Office to the designated test agency. The RDT&E appropriation funds the cost associated with research and development, including test items, DT&E and Air Force Materiel Command (AFMC) support of OT&E of the system or equipment and the test items. Costs associated with initial operational test and evaluation (IOT&E) are RDT&E funded, and costs of qualification operational test and evaluation (QOT&E) are

O&M funded. The AFOTEC is funded through its own program element and has direct control of OT&E funds for all programs. The IOT&E manager prepares a TPO that summarizes the resource requirements for IOT&E and related test support. All pretest IOT&E planning is budgeted through and paid out of the O&M appropriation. The FOT&E costs are paid by AFOTEC and/or the MAJCOM operating the system and funded by the O&M appropriation.

16.5 SUMMARY

Test resources have many conflicting demands and their use must be scheduled well in advance of a test. Resources specific to a particular test must often be developed and funded from the PM's own RDT&E budget. Thus, the PM and his testers must ensure that test resource requirements are identified early in the acquisition cycle, that they are documented in the initial TEMP, and that modifications and refinements are reported in the TEMP updates.

Funds for testing are provided by congressional appropriation to the OMB, which apportions the funds to the Services through the SECDEF. The PPBS is the DOD process used to formulate budget requests to the Congress. Requests by PMs for test resources are usually outlined in the TEMP. Generally, system development is funded from RDT&E funds until the system is operationally deployed and maintained, and O&M funds are used for FOT&E and system maintenance. The weapon system materiel developer is also responsible for funding the development of new test resources specifically needed to test the weapon system. Army and Air Force operational test agencies develop and directly control OT&E funds for their Services.

Table 16-1. TEMP Test Resource Summary Section

Part V--Test and Evaluation Research Summary: Provide a summary of all key resources, government and contractor planned, to be used during the acquisition program. The initial TEMP should project those key resources, including major range and unique instrumentation requirements, threat simulators and targets, necessary to accomplish DT&E and OT&E objectives. As system development progresses, test resource requirements shall be reassessed, and subsequent TEMP updates shall reflect any changed system concepts or requirements and/or updated threat assessments. Specifically, the TEMP shall identify, as applicable, the following test resources:

- Test Articles
- Test Sites and Instrumentation
- Test Support Equipment
- Threat Systems/Simulators
- Test Targets and Expendables
- Operational Force Test Support
- Simulators, Models and Test-Beds
- Special Requirements
- Test and Evaluation Funding Requirements
- Manpower/Personnel Training

Source: DOD 5000.2-M

17

TEST AND EVALUATION MASTER PLAN

17.1 INTRODUCTION

Guidance contained in DODI 5000.2 stipulates "a Test and Evaluation Master Plan will be prepared for all acquisition programs." This reinforces the philosophy that good planning supports good operations. For effective engineering development and decision-making processes, an overall strategy must be developed integrating the collection and evaluation of test data on required performance parameters. The Test and Evaluation Master Plan (TEMP) "relates program schedule, test management strategy and structure, and required resources to: critical operational issues; critical technical parameters; minimum acceptable operational performance requirements; evaluation criteria, and, milestone decision points." Feedback about the degree of system performance maturity and its operational effectiveness and suitability during each phase is essential to the successful fielding of equipment that satisfies user requirements.

17.2 TEMP DEVELOPMENT

The development of program test and evaluation (T&E) strategy, codification in the TEMP, and effective management of the various test processes is one of the primary functions of a program management office. The T&E strategy is highly contingent on Phase 0 concept(s) that are deemed appropriate for satisfying user requirements. As outlined in DODD 5000.1, part 1, the priority for selecting a solution is:

(1) a non-materiel solution, such as changes to tactics, doctrine, operational concepts, training, or organization.

(2) the sequence of materiel alternatives is:

(a) use or modification of an existing US military system.

(b) use or modification of an existing commercially developed or Allied system that fosters a non-developmental acquisition strategy.

(c) a cooperative research and development program with one or more Allied nations.

(d) a new joint-Service development program.

(e) a new service-unique development program.

The quality of the test program may directly reflect the level of effort expended in its development and execution. This varies in direct relationship to the management imposed by the program manager (PM) and, to some extent, by the system engineer. The PM must evaluate the utility of dedicated T&E staff vs. matrix support from the development command. The levels of intensity for planning and executing T&E fluctuate with changes in phases of the acquisition process and in T&E staff support, as appropriate.

Early planning of long-range strategies can be supported with knowledgeable planning teams (Test Integration Working Groups (TIWGs), Test Planning Working Groups (TPWGs)) and reviews by panels of senior T&E management officials — "gray beards." As the tempo of actual test activities begins to build (late Demonstration/Validation Phase (Dem/Val) to pre-LRIP (low-rate initial production) Engineering and Manufacturing Development (EMD) Phase, internal T&E management staff is needed to control the processes and evaluate results.

17.2.1 Program Management Office Responsibilities

The Program Management Office (PMO) is the focal point of the development, review and approval process for the program TEMP. The DOD acquisition process requires a TEMP as one of the primary management strategy documents supporting the decision to start or terminate development efforts at Milestone I. This task is a "difficult do" during the Concept and Definition Phase since some Services do not formulate or staff a PMO until program start (Milestone (MS) I). An additional complicating factor is the nebulous condition of other program source documents (Operational Requirement Document (ORD), System Engineering Management Plan (SEMP), Acquisition Strategy, System Threat Assessment Report (STAR), Integrated Logistics Support Plan (ILSP), etc.) that are also in early stages of development/updating for the milestone review. Since the TEMP must conform to other program management documents, it frequently lags in the development process and does not receive the attention needed from PMO or matrix support personnel. Program Management Office emphasis on early formulation of the test planning teams (TIWG, TPWG) is critical to the successful development of the program TEMP. These teams should consist of the requisite players so a comprehensive and integrated strategy compatible with other engineering and decision-making processes is developed. The PMO will find that the number of parties desiring coordination on the TEMP far exceed the "streamlined" approval process signatories (part 7, DOD 5000.2-M). However, it must be coordinated; an early start in getting Service-level concurrence is important so the Defense Acquisition Board (DAB) document-submission schedule can be supported with the draft and final versions of the TEMP. Subsequent updates do not become easier, as each acquisition phase brings new planning, coordination and testing requirements.

17.2.2 T&E Planning

Developing an overall strategy provides the framework for incorporating phase-oriented T&E activities that will facilitate the acquisition process. The T&E strategy should be consistent with the program acquisition strategy, identifying requirements for contractor and government development test and evaluation (DT&E), interactions between DT&E and operational test and evaluation (OT&E), and provisions for the separate initial operational test and evaluation (IOT&E). An evolutionary acquisition strategy will generally include moderate to low-risk technologies that should reduce the intensity and duration of the T&E program. It does, however, include a requirement for postproduction test activities as the system is modified to accommodate previously unknown new technologies, new threats or other performance enhancements.

A revolutionary acquisition strategy incorporates all the latest technologies in the final production configuration and is generally a higher-risk approach. As the contractor works on maturing emerging technologies, the T&E work load increases in direct proportion to the difficulty in fixing problems. There is a much higher potential for extended schedules with iterative test-fix-test cycles.

The preplanned product improvements (P^3I) strategy is a variant of the evolutionary development process in which you recognize the high-risk technologies/subsystems and put them on a parallel development track. The testing strategy should anticipate the requirements to evaluate P^3I item maturity and then test the system during the integration of the additional capability.

Advanced Technology Demonstrations (ATD) may provide early insights into available technologies for incorporation into developmental or mature, post-MS III systems. Using proven, mature technology provides a lower-risk strategy and may significantly reduce the development testing work load (DODI 5000.2, 5-C, 5-D). "Test and Evaluation shall be used to determine system maturity and identify areas of technical risk," (part 1-C DODD 5000.1). The process for verifying contract technical specifications, MIL-SPEC and MIL-STD testing and evaluation of minimum performance requirements in the ORD, exit criteria or the acquisition program baseline performance should be addressed in the DT&E strategy. The DT&E is an iterative process starting at configuration item/software module levels and continuing throughout the component integration into subassemblies and, finally, system-level performance evaluations. Operational test and evaluation is interwoven into early DT&E for maximizing the efficient use of test articles and test schedules. However, OT&E must remain a distinct thread of activity that does not lose its identity in the tapestry of test events. Planning for test resources is driven by the sequence and intensity of development test (DT) and operational test (OT) events. Resource coordination is an equally arduous task, which frequently has lead times equal to major program development activities. Included in the program T&E strategy should be the overshadowing evaluation plan, outlining methodologies, models, simulations and test data required at periodic decision points. Part 6, Engineering and Manufacturing, DODI 5000.2 provides TEMP requirements:

The TEMP will: (a) address critical human issues to provide data to validate the results of human factors engineering analyses; and (b) require identification of mission critical operational and maintenance tasks. [DODI 5000.2, 6-H]

> A reliability growth (TAFT) program should be developed to satisfy the reliability levels required at MS III. Reliability tests and demonstrations (MIL-STD-785) will be based on actual or simulated operational conditions.
>
> Maintainability will be verified with a maintainability demonstration (MIL-STD-470) before MS III. [DODI 5000.2, 6-C]
>
> As early as practicable, developers and test agencies will assess survivability and validate critical survivability characteristics at as high a system level as possible. The TEMP will identify the means by which the survivability objectives are validated. [DODI 5000.2, 6-F]
>
> Field engineering test facilities and testing in the intended operational environments are required to (1) verify

electric or electronic systems predicted performance, (2) establish confidence in electromagnetic compatibility design based on standards and specifications, and (3) validate electromagnetic compatibility analysis methodology. [DODI 5000.2, 6-G]

The TEMP will address health hazard and safety critical issues to provide data to validate the results of system safety analyses. [DODI 5000.2, 6-I]

The TEMP strategy should directly support the development of more-detailed planning and resource documents needed to execute the actual test events and subsequent evaluations.

Test and Evaluation planning shall address measures of performance with appropriate quantitative criteria, test event or scenario description, resource requirements and test limitations. Test planning, at a minimum, must address all system components that are critical to the achievement and demonstration of contract technical performance specifications and minimum acceptable operational performance requirements specified in the Operation Requirements Document. [part 8, DODI 5000.2]

17.3 TEMP FORMAT

The format specified in DOD 5000.2-M, part 7, is required for all acquisition category I and Office of Secretary of Defense (OSD) designated oversight programs (Table 17-1). It may be tailored as needed for lesser category acquisition programs at the discretion of the milestone decision authority. The TEMP is intended to be a summary document (30 pages in main body) outlining DT&E and OT&E management responsibilities across all phases of the acquisition process. When the development is a multi-Service or joint acquisition program, one integrated TEMP is developed with Service annexes, as required. A Capstone TEMP may not be appropriate for a single major weapon platform but could be used to encompass testing of a collection of individual systems, each with its own annex (e.g., Strategic Defense Initiative Organization (SDIO), Family of Tactical Vehicles, FAADS). A program TEMP is updated at milestones, upon program baseline breach and for other significant program changes. Updates may consist of page changes and are no longer required when a program has no further development activities.

The TEMP is a living document that must address changes to critical issues associated with an acquisition program. Major changes in program requirements, schedule or funding usually result in a change in the test program. Thus, the TEMP must be reviewed and updated on program change, on baseline breach and before each milestone decision, to ensure that T&E requirements are current. As the primary document used in the OSD review and decision process to assess the adequacy of planned testing and evaluation, the TEMP must be of sufficient scope and content to explain the entire T&E program. The key topics in the TEMP are shown in Table 17-1.

Each TEMP submitted to OSD should be a summary document, detailed only to the extent necessary to show the rationale for the type, amount and schedules of the testing planned. It must relate the T&E effort clearly to technical risks, operational issues and concepts, system performance, reliability, availability, maintainability, logistic objectives and requirements, and major decision points. It should summarize the testing accomplished to date and explain the relationship of the various simulations, subsystem tests, integrated system devel-

Table 17-1. Test and Evaluation Master Plan Outline (Format)

PART I SYSTEM INTRODUCTION (2 pages suggested - refer to annexes)
- MISSION DESCRIPTION
- SYSTEM THREAT ASSESSMENT
- MINIMUM ACCEPTABLE OPERATIONAL PERFORMANCE REQUIREMENTS
- SYSTEM DESCRIPTION
- CRITICAL TECHNICAL PARAMETERS

PART II INTEGRATED TEST PROGRAM SUMMARY (2 pages suggested)
- INTEGRATED TEST PROGRAM SCHEDULE
- MANAGEMENT

PART III DEVELOPMENT TEST AND EVALUATION OUTLINE (10 pages suggested)
- DEVELOPMENT TEST AND EVALUATION OVERVIEW
- DEVELOPMENT TEST AND EVALUATION TO DATE
- FUTURE DEVELOPMENTAL TEST AND EVALUATION
- LIVE-FIRE TEST AND EVALUATION

PART IV OPERATIONAL TEST AND EVALUATION OUTLINE (10 pages suggested)
- OPERATIONAL TEST AND EVALUATION OVERVIEW
- CRITICAL OPERATIONAL ISSUES
- OPERATIONAL TEST AND EVALUATION TO DATE
- FUTURE OPERATIONAL TEST AND EVALUATION

PART V TEST AND EVALUATION RESOURCE SUMMARY (6 pages suggested)
- TEST ARTICLES
- TEST SITES AND INSTRUMENTATION
- TEST SUPPORT EQUIPMENT
- THREAT SYSTEMS/SIMULATORS
- TEST TARGETS AND EXPENDABLES
- OPERATIONAL FORCE TEST SUPPORT
- SIMULATIONS, MODELS AND TEST BEDS
- SPECIAL REQUIREMENTS
- TEST AND EVALUATION FUNDING REQUIREMENTS
- MANPOWER/TRAINING

APPENDIX A BIBLIOGRAPHY
APPENDIX B ACRONYMS
APPENDIX C POINTS OF CONTACT

ANNEXES or ATTACHMENTS (if appropriate)

opment tests and initial operational tests that, when analyzed in combination, provide confidence in the system's readiness to proceed into the next acquisition phase. The TEMP must address the T&E to be accomplished in each program phase, with the next phase addressed in the most detail. The TEMP is also used as a coordination document to outline each test and support organization's role in the T&E program and identify major test facilities and resources. The TEMPs supporting the production and initial deployment decision must include the T&E planned to verify the correction of deficiencies and to complete production qualification testing and follow-on OT&E.

The objective of the OSD TEMP review process is to ensure successful T&E programs that will support decisions to commit resources at major milestones. Some of the T&E issues considered during the TEMP review process include:

(1) Are DT&E and OT&E initiated early to assess performance, identify risks and estimate operational potential?

(2) Are critical issues, test directives and evaluation criteria related to mission need and operational requirements established well before testing begins?

(3) Are provisions made for collecting sufficient test data with appropriate test instrumentation to minimize subjective judgment?

(4) Is OT&E conducted by an organization independent of the developer and user?

(5) Do the test methodology and instrumentation provide a mature and flexible network of resources that stress (as early as possible) the weapon system in a variety of realistic environments?

17.4 SUMMARY

The PMO is directly responsible for the content and quality of the test strategy and planning document. The TEMP, as an integrated summary management tool, requires an extensive commitment of man-hours and PM guidance. The interactions of the various T&E players and support agencies must be woven into the fabric of the total system acquisition strategy. Cost and schedule implications must be negotiated to ensure a viable test and evaluation program that provides timely and accurate data to the engineering and management decision-makers.

V

MODULE

Specialized Testing

The nature of a weapon system sometimes requires the use of a specially tailored test and evaluation program. In some cases, hazardous testing must be performed. In other cases, testing must be conducted by specialized organizations or at special times in the development life cycle.

This module addresses the testing of special weapons (such as chemical, laser and space systems), embedded computer systems, electronic warfare and command-and-control systems, logistics infrastructure test and evaluation, and production-related testing activities.

18

EMBEDDED COMPUTER SYSTEMS TESTING

18.1 INTRODUCTION

Software components present a major development risk for military computer systems. They escalate the cost and reduce the reliability of military systems. Embedded computer systems are physically incorporated into larger systems, neither having a major function of data processing. The output of the systems are normally information, control signals or computer data required by the host system to complete its mission. Although hardware and software contribute in equal measure to successful implementation of embedded computer system functions, there have been relative imbalances in their treatment during system development.

The development of embedded systems involves a series of activities in which there are frequent opportunities for errors. Errors may occur at the inception of the process, when the requirements of the system may be erroneously specified, or later in the development cycle, when system specifications are implemented. This chapter addresses the use of testing to control the development risk of embedded computer systems, particularly as it pertains to the software development process.

18.2 MISSION CRITICAL COMPUTER RESOURCES

The term Mission Critical Computer Resources (MCCR) is defined as automated data processing equipment, software or services; and the function, operation or use of the equipment software or services involves:

(1) Intelligence activities;

(2) Cryptologic activities related to national security;

(3) Command and control of military *forces;*

(4) Equipment that is an integral part of a weapons system;

(5) Critical, direct fulfillment of military or intelligence missions.

Acquisition of MCCR is described in part 6, section D, DODI 5000.2, and directs MCCR development planning to follow DOD-STD-2167 and DOD-STD-2168.

18.3 PURPOSE OF SOFTWARE TEST AND EVALUATION

A major problem in software development is a lack of well-defined requirements. If requirements are not well-defined, errors can multiply throughout the development process. As illustrated in Figure 18-1, errors may occur at the inception of the process. These errors may occur during requirements definition, when objectives may

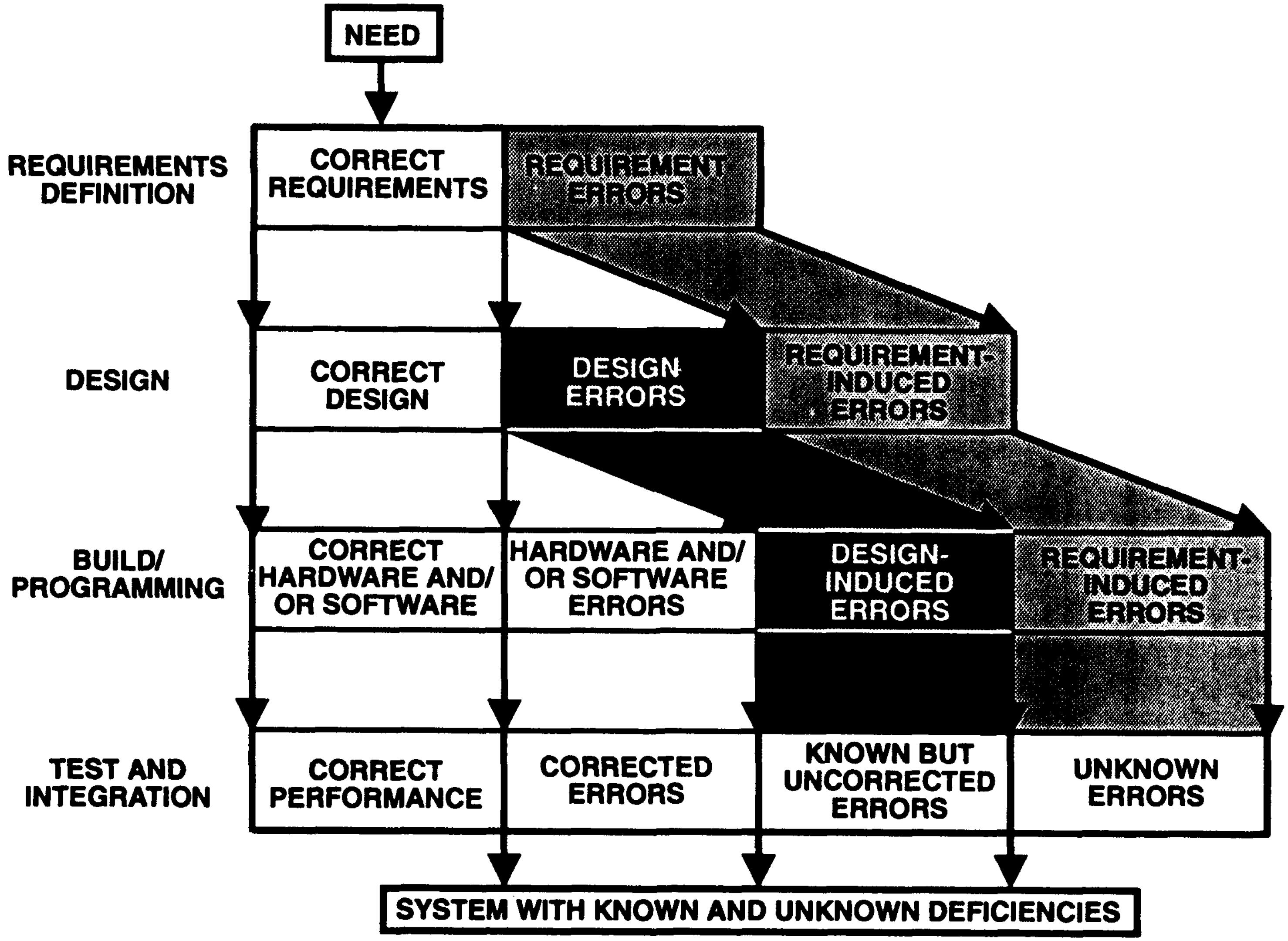


Figure 18-1. The Error Avalanche

be erroneously or imperfectly specified; during the later design and development stages, when these objectives are implemented; and during software maintenance and operational phases, when software changes are needed to eliminate errors or enhance performance. Estimates of increased software costs arising from incomplete testing help to illustrate the dimension of software life-cycle costs. Averaged over the operational life cycle of a computer system, development costs encompass approximately 30 percent of total system costs. The remaining 70 percent of life-cycle costs are associated with maintenance, which includes system enhancements and error correction. Complete testing during earlier development phases may have detected these errors. The relative costs of error correction increase as a function of time from the start of the development process. Relative costs of error correction rise dramatically between requirements and design phases and more dramatically during code implementation.

Previous research in the area of software T&E reveals that half of all maintenance costs are incurred in the correction of previously undetected errors. Approximately one-half of the operational life cycle costs can be traced directly to inadequate or incomplete testing activities. In addition to cost increases, operational implications of software errors in weapon systems can result in mission critical software failures that may impact mission success and personnel safety.

A more systematic and rigorous approach to software testing is required. To be effective, this approach must be applied to all phases of the development process in a planned and coordinated manner, beginning at the earliest design stages and proceeding through operational testing of the integrated system. Early, detailed software test and evaluation (T&E) planning is critical to the successful development of a computer system.

18.4 SOFTWARE DEVELOPMENT PROCESS

Software engineering technologies used to produce operational software are key risk factors in a development program. The T&E program should help determine which of these technologies increase risk and have a life-cycle impact. A principal source of risk is the support software required to develop operational software. In terms of life-cycle impact, operational software problems are commonly associated with the difficulty in maintaining and supporting the software once deployed. Software assessment requires an analysis of the life-cycle impact, which varies depending on the technology used to design and implement the software. One approach to reducing long-term life-cycle risks is to use ADA language and common hardware throughout the development and operation of the software. These life-cycle characteristics that affect operational capabilities must be addressed in the Test and Evaluation Master Plan (TEMP), and tests should be developed to identify problems caused by these characteristics. The technology used to design and implement the software may significantly affect software supportability and maintainability.

The TEMP must sufficiently describe the acceptance criteria or software metrics from the written specifications for operational effectiveness and suitability. The specifications must define the required software metrics to set objectives and thresholds for mission critical functions. Additionally, these metrics should be evaluated at the appropriate stage of system development rather than at some arbitrarily imposed milestone.

18.5 T&E IN THE SOFTWARE LIFE CYCLE

Software testing is an iterative process executed at all development stages to examine program design and code to expose errors. Software test planning should be described in the TEMP with the same care as test planning for other system components.

18.5.1 Testing Approach

The integration of software/MCCR development into the overall acquisition process dictates a testing process consistent with the bottom-up approach taken with hardware development. The earliest stage of software testing is characterized by heavy human involvement in basic design and coding processes. Thus, human testing is defined as informal, noncomputer-based methods of evaluating architectures, designs and interfaces. It can consist of:

- Inspections — The programmer explains his work to a small group of peers with discussion and direct feedback on errors, inconsistencies and omissions.

- Walk-through — A group of peers develop test cases to evaluate work to date and give direct feedback to the programmer.

- Desk Checking — A self evaluation is made by the programmer of his work. There is a low probability of identifying his errors of logic or coding.

- Peer Ratings — Mutually supportive, anonymous reviews are performed by groups of peers with collaborative evaluations and feedback.

- Design Reviews — Preliminary design reviews (PDRs) and critical design reviews (CDRs) provide milestones in the development efforts that review development and evaluations to date. An independent verification and validation (IV&V) contractor may facilitate the government's ability to give meaningful feedback.

Once the development effort has matured beyond the benefits of human testing, computerized-software-only testing may be appropriate. It is performed to determine the functionality of the software when tested as an entity or "build." Documentation control is essential so that test results are correlated with the appropriate version of the build. Software testing is usually conducted using some combination of "black box" and "white box" testing.

- **Black Box** — Functional testing of a software unit without knowledge of how the internal structure or logic will process the input to obtain the specified output. Within-boundary and out-of-boundary stimulants test the software's ability to handle abnormal events. Most likely cases are tested to provide a reasonable assurance that the software will demonstrate specified performance. Even the simplest software designs rapidly exceed our capacity to test all alternatives.

- **White Box** — Structural testing of the internal logic and software structure provides an opportunity for more extensive identification and testing of critical paths. The process and objectives are otherwise very similar to black box testing.

Testing should be performed from the bottom up. The smallest controlled software modules — computer software units — are tested individually. They are then combined or integrated and tested in larger aggregate groups or builds. When this process is complete, the software system is tested in its entirety. Obviously, as errors are found in the latter stages of the test program, a return to earlier portions of the development program to provide corrections is required. The cost impact of error detection and correction can be diminished using the bottom-up testing approach.

System level testing can begin once all modules in the computer software configuration item (CSCI) have been coded and individually tested. A software integration lab (SIL), with adequate machine time and appropriate simulations, will facilitate hardware simulation/emulation and the operating environment. If data analysis indicates proper software functioning, it is time to advance to a more complex and realistic test environment.

• **Hot Bench Testing** — Integration of the software released from the SIL for full-up testing with actual system hardware in a hardware-in-the-loop (HWIL) facility marks a significant advance in the development process. Close approximation of the actual operating environment should provide test sequences and stress needed to evaluate the effectiveness of the software system(s). Problems stimulated by the "noisy environment," interface problems, electromagnetic interference (EMI) and different electrical transients should surface. Good hardware and software test programs leading up to HWIL testing should aid in isolating problems to the hardware or software side of the system. Caution should be taken to avoid any outside stimuli that might trigger unrealistic responses.

• **Field Testing** — Development test and evaluation (DT&E) and operational test and evaluation (OT&E) events must be designed to provide for data collection processes and instrumentation that will measure system responses and allow data analysts to identify the appropriate causes of malfunctions. Field testing should be rigorous, providing environmental stresses and mission profiles likely to be encountered in operational scenarios. Government software support facilities tasked for future maintenance of the software system should be brought on board to familiarize them with the system operating characteristics and documentation. Their expertise should be included in the software test and evaluation process to assist in the selection of stimuli likely to expose software problems.

It is critical that adequate software T&E information be contained in documents such as TEMPs and test plans. The TEMP must define characteristics of critical software components that effectively address goals and thresholds for mission critical functions. The measures of effectiveness (MOEs) must support the critical software issues. The test plan should specify the test methodologies that will be applied. Test methodologies consist of two components. The first is the test strategy that guides the overall testing effort, and the second is the *testing technique that is applied* within the framework of a test strategy.

Effective test methodologies require realistic software test environments and scenarios. The test scenarios must be appropriate for the test objectives; i.e., the test results must be interpretable in terms of software test objectives. The test scenarios and analysis should actually verify and validate accomplishment of requirements. The test environments must be chosen on a careful analysis of characteristics to be demonstrated and its relationship to the development, operational and support environments. In addition, environment must be representative of that in which the software will be maintained.

18.5.2 Independent Verification and Validation

Independent verification and validation are risk-reducing techniques that are applied to major software development efforts. The primary purpose of IV&V is to ensure that software meets requirements and is reliable and maintainable. The IV&V is effective only if implemented early in the software development schedule. Requirements analysis and risk assessment are the most critical activities performed by IV&V organizations; their effectiveness is limited if brought on board a project after the fact. Often, there is a reluctance to implement IV&V because of the costs involved, but early implementation of IV&V will result in lower overall costs of error correction and software maintenance. As development efforts progress, IV&V involvement typically decreases. This is due more to the expense of continued involvement than to a lack of need. For an IV&V program to be effective, it must be the responsibility of an individual or organization external to the software development program manager.

The application of the IV&V process to software development maximizes the maintainability of the fielded software system, while minimizing the cost of developing and fielding it. Maintenance of a software system falls into several major categories: corrective maintenance, modifying software to correct errors in operation; adaptive maintenance, modifying the software to meet changing requirements; and perfective maintenance, modifying the software to incorporate new features or improvements.

The IV&V process maximizes the reliability of the software product, which eases the performance of and minimizes the need for corrective maintenance. It attempts to maximize the flexibility of the software product, which eases the performance of adaptive and perfective maintenance. These goals are achieved primarily by determining at each step of the software development process that the software product completely and correctly meets the specific requirements determined at the previous step of development. This step-by-step, iterative process continues from the initial definition of system performance requirements through final acceptance testing.

The review of software documentation at each stage of development is a major portion of the verification process. The current documentation is a description of the software product at the present stage of development and will define the requirements laid on the software product at the following stage. Careful examination and analysis of the development documentation ensure that each step in the software design process is consistent with the previous step. Omissions, inconsistencies or design errors can then be identified and corrected early in the development process.

Continuing participation in formal and informal design reviews by the IV&V organization maintains the communication flow between software system "customers" and developers, ensuring that software design and production proceed with minimal delays and misunderstandings. Frequent informal reviews, design and code walkthrough and audits ensure that the programming standards, software engineering standards, software quality assurance and configuration management procedures designed to produce a reliable, maintainable operational software system are followed throughout the process. Continuous monitoring of computer hardware resource allocation throughout the software development process also ensures that the fielded system has adequate capacity to meet operation and maintainability requirements.

The entire testing process, from the planning stage through final acceptance test, is also approached in a step-by-step manner by the IV&V process. At each stage of development, the functional requirements determine test criteria as well as design criteria for the next stage. An important function of the IV&V process is to ensure that the test requirements are derived directly from the performance requirements and are independent of design implementation. Monitoring of, participation in and performance of the various testing and inspection activities by the IV&V contractor ensure that the developed software meets requirements at each stage of development.

Throughout the software development process, the IV&V contractor reviews any proposals for software enhancement or change, proposed changes in development baselines, and proposed solutions to design or implementation problems to ensure that the original performance requirements are not forgotten. An important facet of the IV&V contractor's role is to act as the objective third party, continuously maintaining the "audit trail" from the initial performance requirements to the final operational system.

18.6 SUMMARY

There is a useful body of software testing technologies that can be applied to testing of embedded systems. As a technical discipline, though, software testing is still maturing. There is little to guide the program manager in choosing one testing technique over another. It is apparent that systematic T&E techniques are far superior to *ad-hoc* testing techniques. Implementation of an effective T&E plan requires a set of strong technical and management controls. Given the increasing number of embedded computer systems being acquired, there will be an increased emphasis on tools and techniques for T&E. For more-detailed information on MCCR development and testing, review the DSMC Mission Critical Computer Resource Management Guide.

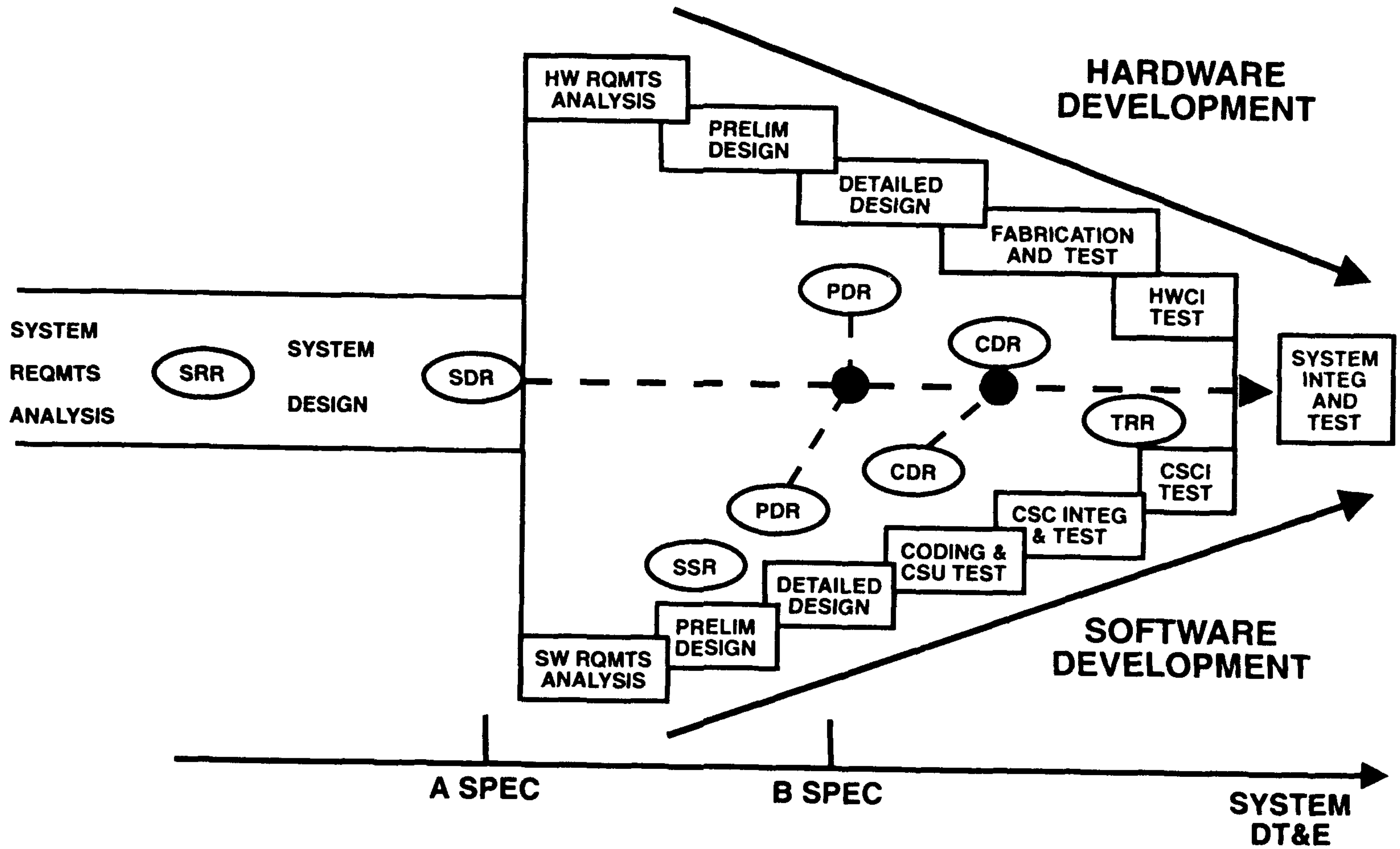


Figure 18-2. System Development Process

19

TESTING FOR VULNERABILITY AND LETHALITY

19.1 INTRODUCTION

This chapter addresses the need to explore the vulnerability and lethality aspects of a system through test and evaluation (T&E) practices and procedures. In particular, this chapter describes the legislatively-mandated Live Fire Test Program, which has been established to evaluate the vulnerability and lethality of developing systems. It also discusses the role of T&E in assessing a system's ability to perform in a nuclear combat environment. The discussion of testing for nuclear survivability is based primarily on information contained in the "Nuclear Survivability Handbook for OT&E," prepared by the Air Force Operational Test and Evaluation Center (Reference 91).

19.2 LIVE FIRE TESTING

19.2.1 Background

In March 1984, OSD chartered a joint T&E program designated "The Joint Live Fire Program." This program was to assess the vulnerabilities and lethalities of selected U.S. and threat systems already fielded. The controversy over joint live fire testing of the Army's Bradley Fighting Vehicle System, subsequent congressional hearings and media exposure resulted in provisions being incorporated in the National Defense Authorization Act of FY 1987. This act required an Office of the Secretary of Defense (OSD)-managed Live Fire Testing (LFT) program for major acquisition programs fitting certain criteria. Subsequent amendments to legislative guidance have dictated the current program. The DOD implementation of congressional guidance in DODI 5000.2, part 8, requires that "covered major systems and munitions programs," (i.e., Acquisition Category (ACAT) I and II programs) must execute survivability and lethality testing before Milestone (MS) III, full-rate production. Additionally, product improvements to those systems may reinitiate live fire testing requirements. The Secretary of Defense has the authority to waive these requirements before the system passes MS II. Programs subject to LFT are listed on the OSD annual oversight list. The USD(A) agent for management of the DOD Live Fire Test program is the Director, Development Test. This type of development test and evaluation (DT&E) must be planned to start early enough in the development process to impact design and to provide timely test data for the OSD Live Fire Test Report required for the MS III Defense Acquisition Board (DAB) and congressional committees. The Service-detailed Live Fire Test Plan must be reviewed and approved by the Director, Test and Evaluation (DTE), and live fire testing must be addressed in Part III of the program Test and Evaluation Master Plan (TEMP). The OSD has published guidelines, which can be obtained from the DTE.

19.2.2 Live Fire Tests

There are varying types and degrees of live fire tests. The matrix in Table 19-1 illustrates the various possible combinations. Full-scale, full-up testing is usually considered to be the most realistic and is the type of testing called for in the National Defense Authorization Act for FY 1987.

The importance of full-scale testing has been well demonstrated by the Joint Live Fire (JLF) tests. In one case, these tests contradicted earlier conclusions concerning the flammability of a new hydraulic fluid used in F-15 and F-16 aircraft. Laboratory tests had demonstrated that the new fluid was less flammable than the standard fluid. However, during the JLF tests, 30 percent of the shots on the new fluid resulted in fires contrasted with 15 percent of the shots on the standard fluid (Reference 100).

While much insight and valuable wisdom are to be obtained through the testing of components or subsystems, some phenomena are only observable when full-up systems are tested. The interaction of such phenomena has been termed "cascading damage." Such damage is a result of the synergistic damage mechanisms that are at work in the "real world" and likely to be found during actual combat. Live Fire Testing provides a way of examining the damages inflicted not only on materiel but also on personnel. The crew casualty problem is an important issue that the LFT program is addressing. The program provides an opportunity to assess the effects of the complex environments that crews are likely to encounter in combat (e.g., fire, toxic fumes, blunt injury shock and acoustic injuries) (Reference 91).

19.2.3 Use of Modeling and Simulation

Survivability and lethality assessments have traditionally relied largely on the use of modeling and simulation techniques. The Live Fire Test Program does not replace the need for such techniques; in fact, the Live Fire Test Guidelines issued by OSD in May 1987 require that no shots be conducted until pre-shot model predictions are made concerning the expected damage. Such predictions are useful for several reasons. First, they assist in the test planning process. If a model predicts that no damage will be inflicted, test designers and planners should reexamine the selection of the shotlines and/or reassess the accuracy of the threat representation. Second, pre-shot model predictions provide the Services with the opportunity to validate the accuracy of the models by comparing them with actual LFT results. At the same time, the LFT program reveals areas of damage that may be absent from existing models and simulations. Third, pre-shot model predictions can be used to help conserve scarce target resources. For example, models can be used to determine a sequence of shots that provides for the less-damaging shots to be conducted first, followed by the more-catastrophic shots resulting in maximum target damage.

19.2.4 Live Fire Test Guidelines

The Live Fire Test Planning Guide was updated by OSD in June 1989. The guidelines state that plans for live fire testing must be included in the TEMP. Key points covered in the LFT guidelines include the following:

- The Live Fire Test and Evaluation (LFT&E) plan is the basic planning document used by OSD and the Services to plan, review and approve LFT&E.

- The LFT&E plan must contain general information on the system's required performance, operational and technical characteristics, critical test objectives and the evaluation process.

Table 19-1. Types of Live Fire Testing

	LOADING	
	FULL-UP	INERT [a]
Full Scale	Complete system: with combustibles (e.g., Bradley Phase II tests, aircraft "proof" tests)	Complete system: no combustibles (e.g., tests of new armor on actual tanks, aircraft flight control tests)
Subscale	Components, subcomponents: with combustibles (e.g., fuel cell tests, behind armor, mock-up aircraft, engine fire tests)	Components, subcomponents, structures, terminal ballistics, munitions performance, behind-armor tests, warhead characterization (e.g., armor/warhead interaction tests, aircraft component structural tests)

[a] In some cases, targets are "semi-inert," meaning some combustibles are on board but not all. (Example: tests of complete tanks with fuel and hydraulic fluid, but dummy ammunition)

Source: "Live Fire Testing: Evaluating DoD's Programs," General Accounting Office, GAO/PEMD-87-17, August 1987.

• Each LFT&E plan must include testing of complete systems. A limited set of live fire tests may involve production components configured as a subsystem before full-up testing.

• A Service report must be submitted within 60 days of the completion of the live fire test. The report must include the firing results, test conditions, limitations and conclusions and be submitted in classified and unclassified form.

• Within 45 days of receipt of the Service report, a separate Live Fire Test Report (part 10, DOD 5000.2-M) will be produced by OSD. The conclusions of the OSD report will be independent of the conclusions of the Service report.

• The Congress shall have access to all live fire test data and all live fire test reports held by or produced by the Secretary of the concerned Service or by OSD.

• The costs of all live fire tests shall be paid from funding for the system being tested. In some instances, the DTE may elect to supplement such funds for the acquisition of targets or target simulators, although the ultimate responsibility rests on the concerned Service.

19.3 TESTING FOR NUCLEAR HARDNESS AND SURVIVABILITY

19.3.1 Background

Nuclear survivability must be incorporated into the design, acquisition and operation of all systems that must perform critical missions in a nuclear environment. Nuclear survivability is achieved through a combination of four methods: hardness, avoidance, proliferation and reconstitution. Hardness allows a system to physically withstand a nuclear attack. Avoidance encompasses measures taken to avoid encountering a nuclear environment. Proliferation involves having sufficient systems to compensate for probable losses. Reconstitution includes the actions taken to repair or resupply damaged units in time to complete a mission satisfactorily.

There is a wide variety of possible effects from a nuclear detonation. They include: electromagnetic pulse (EMP), ionizing radiation, thermal radiation, blast, shock, dust, debris, blackout and scintillation. Each weapon system is susceptible to some but not all of these effects. The program manager and his staff must identify the effects that may have an impact on the system under development and manage the design, development and testing of the system in a manner that minimizes degradation. The variety of possible nuclear effects is described more fully in the "Nuclear Survivability Handbook for Air Force OT&E" (Reference 91).

19.3.2 Assessing Nuclear Survivability Throughout The System Acquisition Cycle

The program manager must ensure that nuclear survivability issues are addressed throughout the system acquisition cycle. During the Concept Exploration and Definition Phase, the survivability requirements stated in the Service requirements document should be verified, refined or further defined. "Critical survivability characteristics will be used to evolve survivability design criteria which will be included in appropriate configuration baselines" (part 6-F, DODI 5000.2). During the Demonstration and Validation Phase, trade-offs between hardness levels and other system characteristics (such as weight, decontaminability and compatibility) should be described quantitatively. Trade-offs, between hardness, avoidance, proliferation and reconstitution as a method for achieving survivability, should also be considered at this time. During the Engineering and Manufacturing Development Phase, the system must be adequately tested to confirm that hardness objectives, criteria, requirements and specifications are met. Plans for nuclear hardness and survivability testing must be outlined in the TEMP (part 6-F, DODI 5000.2). The appropriate commands must make provision for test and hardness surveillance equipment and procedures so required hardness levels can be maintained once the system is operational.

During the Production and Deployment Phase, system hardness is maintained through an active hardness assurance program. Such a program ensures that the end product conforms to hardness design specifications and that hardness aspects are re-evaluated before any retrofit changes are made to existing systems.

Once a system is operational, a hardness surveillance program may be implemented to maintain system hardness and to identify any further evaluation, testing or retrofit changes required to ensure survivability. A hardness surveillance program consists of a set of scheduled tests and inspections to ensure that a system's designed hardness is not degraded through operational use, logistic support, maintenance actions or natural causes.

19.3.3 Test Planning

The "Nuclear Survivability Handbook for Air Force OT&E" describes the following challenges associated with nuclear hardness and survivability testing:

(1) The magnitude and range of effects from a nuclear burst are much greater than those from conventional explosions that may be used to simulate nuclear bursts. Nuclear detonations have effects not found in conventional explosions. The intense nuclear radiation, blast, shock, thermal and EMP fields are difficult to simulate. In addition, systems are often tested at stress levels that are either lower than those established by the criteria or lower than the level needed to cause damage to the system.

(2) The yields and configurations for underground testing are limited. It is generally not possible to test all relevant effects simultaneously or to observe possibly important synergism between effects.

(3) System-level testing for nuclear effects is normally expensive, takes years to plan and conduct and requires specialized expertise. Often, classes of tests conducted early in the program are not repeated later. Therefore, operational requirements should be folded into these tests from the start, often early in the acquisition process. This mandates a more-extensive, combined DT&E/OT&E test program than normally found in other types of testing.

Program managers and test managers must remain sensitive to the ambiguities involved in testing for nuclear survivability. For example, there is no universal quantitative measure of survivability; and statements of survivability may lend themselves to a variety of interpretations. Moreover, it can be difficult to combine system vulnerability estimates for various nuclear effects into an assessment of overall survivability. As a result, program/test managers must exercise caution when developing test objectives and specifying measures of merit related to nuclear survivability.

19.3.4 Test Execution

For nuclear hardness and survivability testing, development test (DT) and operational test (OT) efforts are often combined because it is not possible to test in an operational nuclear environment. The use of an integrated DT/OT program requires early and continuous dialogue between the two test communities so each understands the needs of the other and maximum cooperation in meeting objectives is obtained.

Test and evaluation techniques available to validate the nuclear survivability aspects of systems and subsystems include underground nuclear testing, environmental simulation (system level, subsystem level and component level) and analytical simulation. Table 19-3 outlines the major activities relevant to the assessment of nuclear hardness and survivability and the phases of the acquisition cycle in which they occur.

19.4 SUMMARY

The vulnerability and lethality aspects of a system can be evaluated through live fire tests. These tests are used to provide an insight into the system's ability to protect its crew and to continue to operate/fight after being hit by enemy weapons. It provides a way of examining the damages inflicted not only on materiel but also on personnel. Live fire testing also provides an opportunity to assess the effects of complex environments that crews are likely to encounter in combat.

Nuclear survivability must be carefully evaluated during the system acquisition cycle. Trade-offs between hardness levels and other system characteristics, such as weight, speed, range, cost, etc., must be evaluated. Nuclear survivability testing is difficult, and the evaluation of test results may lend itself to a variety of interpretations. Therefore, program managers must exercise caution when developing test objectives related to nuclear survivability.

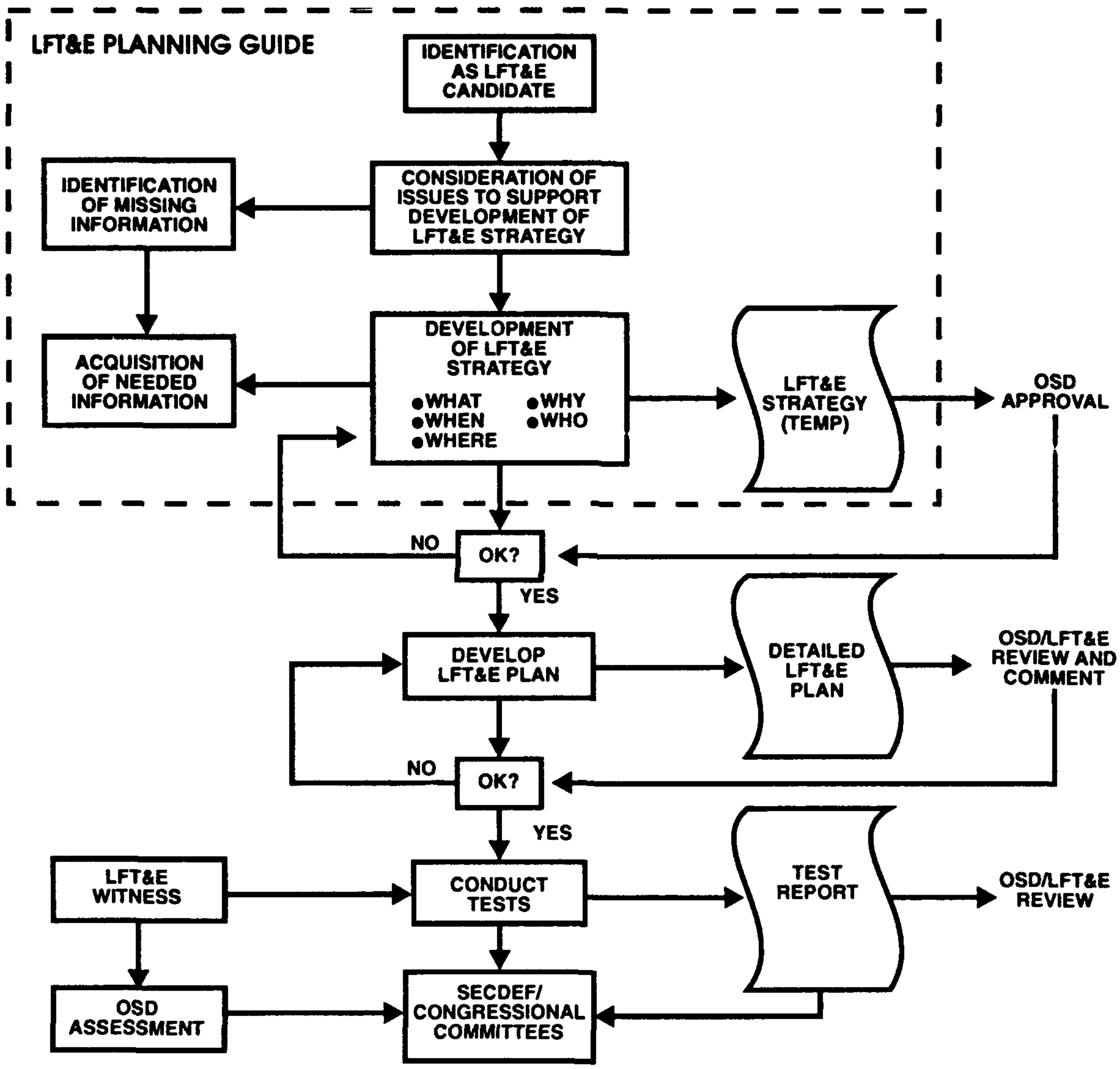


Figure 19-1. Live Fire T&E Planning Guide

Table 19-2. Relationships Between Key Concepts

TERMINOLOGY	PERSPECTIVE		MEANING
	DEFENSIVE	OFFENSIVE	
SURVIVABILITY	X		PROBABILITY OF ENGAGEMENT
EFFECTIVENESS		X	
VULNERABILITY	X		PROBABILITY OF KILL GIVEN A HIT
LETHALITY		X	
SUSCEPTIBILITY	X		PROBABILITY OF ENGAGEMENT

Source: Adapted from "Live Fire Testing: Evaluating DoD's Programs,"
U.S. General Accounting Office, GAP/PEMD-87-17, August 1987, page 15.

Table 19-3. Nuclear Hardness and Survivability Assessment Activities

CONCEPT EXPLORATION/DEFINITION PHASE

- Preparation of Test and Evaluation Master Plan (TEMP) that includes initial plans for nuclear hardness and survivability (NH&S) tests
  - Identification of NH&S requirements in verifiable terms
  - Identification of special NH&S test facility requirements with emphasis on long lead time items
- Development of nuclear criteria

DEMONSTRATION/VALIDATION PHASE

- Increased test planning
- TEMP update
- Conduct of NH&S trade studies
  - NH&S requirements vs. other system requirements
  - Alternate methods for achieving NH&S
- Conduct of limited testing
  - Piece-part hardness testing
  - Design concept trade-off testing
  - Technology demonstration testing
- Development of system specifications that include quantitative hardness levels

Table 19-3. Nuclear Hardness and Survivability Assessment Activities (Continued)

ENGINEERING AND MANUFACTURING DEVELOPMENT PHASE

- **First opportunity to test prototype hardware**
- **TEMP update**
- **Development of Nuclear Hardness Design Handbook**
  - Prior to Preliminary Design Review
  - Usually prepared by nuclear effects specialty contractor
- **Conduct of testing**
  - Pre-Critical Design Review (CDR) development and qualification tests
  - Development testing on nuclear-hardened piece parts, materials, cabling, and circuits
  - NH&S box and subsystem qualification tests (post-CDR)
  - Acceptance tests to verify hardware meets specifications (post-CDR, prior to first delivery)
  - System-level hardness analysis (using box and subsystem test results)
  - System-level NH&S test

PRODUCTION AND DEPLOYMENT PHASE

- **Implementation of program to ensure system retains its NH&S properties throughout production and deployment**
- **Screening of production hardware for hardness**

OPERATIONS AND SUPPORT PHASE

- **Implementation by user of procedures to ensure system's operation, logistic support and maintenance do not degrade hardness features**
- **Reassessment of survivability throughout system life cycle**

20

LOGISTICS INFRASTRUCTURE T&E

20.1 INTRODUCTION

In all materiel acquisition programs, the integrated logistics support (ILS) effort begins in the Mission Area Analysis Phase before program initiation, continues throughout the acquisition cycle and extends past the deployment phase. Logistics testing must, therefore, extend over the entire acquisition cycle of the system and be carefully planned and executed to ensure the readiness and supportability of the system. This chapter covers the development of logistics support test requirements and the conduct of supportability assessments to ensure that readiness and supportability objectives are identified and achieved. The importance of the ILS manager's participation in the Test and Evaluation Master Plan (TEMP) development process should be stressed. He must ensure the ILS test and evaluation (T&E) objectives are considered and that adequate resources are available for ILS T&E.

Integrated logistic support is defined as a disciplined, unified and iterative approach to the management and technical activities necessary to integrate support considerations into system and equipment design; develop support requirements that are related consistently to readiness objectives, design and each other; acquire the required support; and provide the required support during the Operational Phase at minimum cost (DODI 5000.2, part 15).

Integrated logistic support consists of 10 specific components, or elements:

(1) Maintenance planning

(2) Manpower and personnel

(3) Supply support

(4) Support equipment

(5) Technical data

(6) Training and training support

(7) Computer resources support

(8) Facilities

(9) Packaging, handling, storage and transportation

(10) Design interface.

20.2 PLANNING FOR ILS T&E

20.2.1 Objectives of ILS T&E

The main objective of ILS T&E is to verify that the logistic support being developed for the materiel system is capable of meeting the required objectives for peacetime and wartime employment. The ILS T&E consists of the usual development test and evaluation (DT&E) and operational test and evaluation (OT&E) but also includes postdeployment supportability assessments. The formal DT&E and OT&E begin in the Concept Exploration and Definition

Phase and continue into the Production and Deployment Phase. Figure 20-1, which appears in the DSMC Integrated Logistics Support Guide, describes the specific development (DT), operational test (OT) and supportability assessment objectives for each acquisition phase.

20.2.2 Planning Documentation for ILS T&E

20.2.2.1 Integrated Logistic Support Plan

The ILS manager for a materiel acquisition system is responsible for developing the Integrated Logistic Support Plan (ILSP), which is the primary document for planning and implementing the support of the fielded system. It is initially prepared during the Concept Exploration and Definition Phase, and progressively developed in more detail as the system moves through the acquisition phases. Identification of the specific ILS test issues related to the individual ILS elements and the overall system support and readiness objectives are included in the ILSP.

The ILS manager is assisted throughout the system's development by the Integrated Logistics Support Management Team (ILSMT), which is formed early in the acquisition cycle. The ILSMT is a coordination/advisory group composed of personnel from the program management office, the using command and other commands concerned with acquisition activities such as logistics, testing and training.

20.2.2.2 Supportability Assessment Plan

Based upon the ILSP objectives, the ILS manager, in conjunction with the system's test manager, develops the Supportability Assessment Plan (data item description DI-5-7120). This plan identifies the testing approach and the evaluation criteria that will be used to assess the supportability-related design requirements (e.g., reliability and maintainability) and adequacy of the planned logistic support resources for the materiel system. Development of the Supportability Assessment Plan begins in the Concept Exploration and Definition Phase; the plan is then updated and refined in each successive acquisition phase. The ILS manager applies the techniques of logistic support analysis as described in MIL-STD-1388-1A. Test and evaluation strategy is formulated, T&E program objectives and criteria are established and required test resources are identified. The ILS manager ensures that T&E strategy is based upon quantified supportability requirements and addresses supportability issues including those with a high degree of associated risk. Also, the ILS manager ensures that the necessary quantities and types of data will be collected during system development and after deployment of the system to validate the various T&E objectives. The T&E objectives and criteria must provide a basis that ensures critical supportability issues and requirements are resolved or achieved within acceptable confidence levels.

20.2.2.3 Test and Evaluation Master Plan (TEMP)

The program manager must include ILS T&E information in the TEMP as specified in DOD 5000.2-M. The input, which is derived from the Supportability Assessment Plan with the assistance of the ILS manager and the tester, includes descriptions of required operational suitability, specific plans for testing logistics supportability and required testing resources. It is of critical importance that all key test resources required for ILS testing (DT, OT, and postdeployment supportability) be identified in the TEMP because the TEMP provides a long-range alert upon which test resources are budgeted and obtained for testing.

TEST TYPE / ACQUISITION PHASE	CONCEPT EXPLORATION/ DEFINITION	DEMONSTRATION/ VALIDATION	ENGINEERING AND MANUFACTURING DEVELOPMENT	FULL-RATE PRODUCTION/ DEPLOYMENT	OPERATIONS AND SUPPORT
DEVELOPMENT T&E	• SELECT PREFERRED SYSTEM AND SUPPORT CONCEPTS	• IDENTIFY PREFERRED TECHNICAL APPROACH, LOGISTIC RISKS, AND PREFERRED SOLUTIONS	• IDENTIFY DESIGN PROBLEMS AND SOLUTIONS IN RE: - SURVIVABILITY - COMPATIBILITY - TRANSPORTATION - R&M - SAFETY - HUMAN FACTORS	• ENSURE PRODUCTION ITEMS MEET DESIGN REQUIREMENTS AND SPECIFICATIONS	• ENSURE ADEQUACY OF SYSTEM DESIGN CHANGES
OPERATIONAL T&E AND SUPPORTABILITY ASSESSMENT	• ASSESS OPERATIONAL IMPACT OF CANDIDATE TECHNICAL APPROACHES • ASSIST IN SELECTING PREFERRED SYSTEM AND SUPPORT CONCEPTS • ESTIMATE OPERATIONAL COMPATIBILITY AND SUITABILITY	• EXAMINE OPERATIONAL ASPECTS OF ALTERNATIVE TECHNICAL APPROACHES • ESTIMATE POTENTIAL OPERATIONAL SUITABILITY OF CANDIDATE SYSTEMS	• ASSESS OPERATIONAL SUITABILITY - OPERATIONAL R&M - BUILT-IN DIAGNOSTIC CAPABILITY - TRANSPORTABILITY • EVALUATE LOGISTICS SUPPORTABILITY - EFFECTIVENESS OF MAINTENANCE PLANNING - APPROPRIATE PERSONNEL SKILLS/ GRADES - APPROPRIATE SPARES, REPAIR PARTS, BULK SUPPLIES - ADEQUATE SUPPORT EQUIPMENT, INCLUDING EFFECTIVE ATE AND SOFTWARE - ACCURATE AND EFFECTIVE TECHNICAL DATA; VALIDATION/ VERIFICATION OF TECHNICAL MANUALS - ADEQUATE FACILITIES (SPACE, ENVIRONMENTAL SYSTEMS, STORAGE) - EFFECTIVE PACKAGEING, LIFTING DEVICES, TIE-DOWN POINTS, TRANSPORTATION INSTRUCTIONS	• ENSURE PRODUCTION ITEMS MEET OPERATIONAL SUITABILITY REQUIREMENTS	• DEMONSTRATE ATTAINMENT OF SYSTEM READI-NESS OBJECTIVES • UPDATE O&S COST ESTIMATES • EVALUATE OPERATIONAL SUITABILITY AND SUPPORTABILITY OF DESIGN CHANGES • IDENTIFY IMPROVEMENT REQUIRED IN SUPPORTABILITY PARAMETERS • PROVIDE DATA REQUIRED TO ADJUST ILS ELEMENTS

Source: DSMC, Integrated Logistics Support Guide, May 1986.

ATE: Automatic Test Equipment

Figure 20-1. ILS Objectives in the T&E Program

20.2.3 Planning Guidelines for Logistic T&E

The following guidelines for ILS T&E were selected from those listed in the DSMC ILS Guide:

(1) Develop a test strategy for each ILS-related objective. Ensure that OT&E planning encompasses all ILS elements. The general ILS objectives shown in Figure 20-1 must be translated into detailed quantitative and qualitative requirements for each acquisition phase and each T&E program.

(2) Incorporate ILS testing requirements (where feasible) into the formal DT&E/OT&E plans.

(3) Identify ILS T&E that will be performed outside of the normal DT&E and OT&E. Include subsystems that require off-system evaluation.

(4) Identify all required resources, including test articles and logistic support items for formal DT/OT and separate ILS testing (participate with test planner).

(5) Ensure establishment of an operationally realistic test environment, to include personnel representatives of those who will eventually operate and maintain the fielded system. These personnel should be trained for the test using prototypes of the actual training courses and devices. They should be supplied with drafts of all technical manuals and documentation that will be used with the fielded system.

(6) Ensure planned OT&E will provide sufficient data on high-cost and high-maintenance burden items (e.g., for high-cost critical spares, early test results can be used to reevaluate selection).

(7) Participate early and effectively in the TEMP development process to ensure the TEMP includes critical logistic T&E and needed ILS test funds from program and budget documents.

(8) Identify the planned utilization of all data collected during the assessments to avoid mismatching of data collection and information requirements.

Detailed evaluation criteria for each of the 10 ILS elements listed above are presented in Department of the Army Pamphlet 700-50, "Integrated Logistic Support: Developmental Supportability Test and Evaluation Guide."

20.3 CONDUCTING ILS T&E

20.3.1 The Process

The purposes of ILS T&E are to measure the supportability of a developing system throughout the acquisition process, to identify supportability deficiencies and potential corrections/improvements as test data becomes available, and to assess the operational suitability of the planned support system. The ILS T&E also evaluates the system's ability to achieve planned readiness objectives for the system/equipment being developed. Specific ILS T&E tasks (as prescribed in MIL-STD-1388-1A) include:

- Analysis of test results to verify achievement of specified supportability requirements;

- Determination of improvements in supportability and supportability-related design parameters needed for the system to meet established goals and thresholds;

- Identification of areas where established goals and thresholds have not been demonstrated within acceptable confidence levels;

• Development of corrections for identified supportability problems such as modifications to hardware, software, support plans, logistic support resources or operational tactics;

• Projection of changes in costs, readiness and logistic support resources due to implementation of corrections;

• Analysis of supportability data from the deployed system to verify achievement of the established goals and thresholds and where operational results deviate from projections, determination of the causes and corrective actions.

Integrated logistics support T&E may consist of a series of ILS demonstrations and assessments that are usually conducted as part of system performance tests. Special end-item equipment tests are rarely conducted solely for ILS evaluation.

20.3.2 Reliability and Maintainability

System availability is generally considered to be composed of two major system characteristics — reliability and maintainability. The DODI 5000.2 states:

Reliability (R) is the ability of a system and its parts to perform its mission without failure, degradation, or demand on the support system.

Maintainability (M) is the ability of an item to be retained in or restored to specified condition when maintenance is performed by personnel having specific skill levels, using prescribed procedures and resources, at each prescribed level of maintenance and repair.

Operational Reliability and Maintainability Value is any measure of reliability or maintainability that includes the combined effects of item design, quality, installation, environment, operation, maintenance, and repair.

The R and M program objectives are to be defined as system parameters early in the development process. They will be used as evaluation criteria throughout the design, development and production processes. "Reliability and maintainability objectives will be translated into quantifiable contractual terms and allocated through the system design hierarchy." An understanding of how this allocation affects testing operating characteristics below system level can be found in DOD 3235.1-H, "T&E of System Reliability, Availability and Maintainability." This is especially important to testing organizations expected to make early predictions of system performance. Guidance on testing reliability may also be found in MIL-STD-78I, "Reliability Testing for Engineering Development, Qualification, and Production."

20.3.2.1 Reliability

In MIL-STD-785, the following is discussed:

Environmental Stress Screening (ESS) is a test, or series of tests during engineering development, specifically designed to identify weak parts or manufacturing defects. Test conditions should stimulate failures typical of early field service rather than provide an operational life profile.

Reliability Development/Growth Testing (RDT/RGT) is a systematic engineering process of test-analyze-fix-retest (TAFT) where equipment is tested under actual, simulated, or accelerated environments. It is an iterative methodology intended to rapidly and steadily improve reliability.

Reliability Qualification Test (RQT) is to verify that minimum acceptable reliability requirements have been met before items are committed to production. A statistical test plan is used to predefine criteria which will limit government risk. Test conditions must be operationally realistic.

Production Reliability Acceptance Test (PRAT) is intended to simulate in-service use of the delivered item or production lot. "Because it must provide a basis for determining contractual compliance, and because it applies to the items actually delivered to operational forces, PRAT must be independent of the supplier if at all possible". PRAT may require expensive test facilities, so 100% sampling is not recommended.

20.3.2.2 Maintainability

Maintainability design factors and test/demonstration requirements used to evaluate maintainability characteristics must be based on program objectives and thresholds. Areas for evaluation might include (DOD 3235.1-H):

Accessibility: Assess how easily the item can be repaired or adjusted.

Visibility: Assess the ability/need to see the item being repaired.

Testability: Assess ability to detect and isolate system faults to the faulty replaceable assembly level.

Complexity: Assess the impact of the number, location and characteristic (standard or special purpose) on system maintenance.

Interchangeability: Assess the level of difficulty encountered when failed or malfunctioning parts are removed or replaced with an identical part not requiring recalibration.

"A true assessment of system maintainability generally must be developed at the system level under operating conditions and using production configuration hardware." Therefore, DODI 5000.2, part 6-C, requires that a maintainability demonstration (MIL-STD-470) be conducted prior to Milestone III.

20.3.3 T&E of System Support Package

The T&E of the support for a materiel system requires a system support package consisting of spares, support equipment, technical documents and publications, representative personnel, any peculiar support requirements and the test article itself, in short, all of the items that would eventually be required when the system is operational. This complete support package must be at the test site before the test is scheduled to begin. Delays in the availability of certain support items could prevent the test from proceeding on schedule. This could be costly due to on-site support personnel on hold or tightly scheduled system ranges and expensive test resources not being properly utilized. Also, it could result in the test proceeding without conducting the complete evaluation of the support system. The ILS test planner must ensure that the required personnel are trained and available, the test facility scheduling is flexible enough to permit normal delays, and the test support package is on site on time.

20.3.4 Data Collection and Analysis

The ILS manager must coordinate with the testers to ensure that the methods used for collection, storage and extraction of ILS T&E data are compatible with those used in testing the materiel system. As with any testing, the ILS test planning must ensure that all required data is identified; it is sufficient to evaluate a system's readiness and supportability; and plans are made for a data-management system that is capable of the data classification, storage, retrieval and reduction necessary for statistical analysis. Large statistical sample sizes may require a common database that integrates contractor, DT&E and OT&E data so required performance parameters can be demonstrated.

20.3.5 Use of ILS Test Results

The emphasis on the use of the results of testing changes as the program moves from the CED Phase to postdeployment. During early phases of a program, the evaluation results are used primarily to verify analysis and develop future projections. As the program moves into EMD and hardware becomes available, the evaluation addresses design, particularly the reliability and maintainability aspects; training programs; support equipment adequacy; personnel skills and availability; and technical publications.

The ILS manager must make the program manager (PM) aware of the impact on the program of logistical shortcomings that are identified during the T&E process. The PM, in turn, must ensure that the solutions to any shortcomings are identified and reflected in the revised specifications and that the revised test requirements are included in the updated TEMP as the program proceeds through the various acquisition stages.

20.4 LIMITATIONS TO ILS T&E

Concurrent testing or tests that have accelerated schedules frequently do not have sufficient test articles, equipment or hardware to achieve statistical confidence in the testing conducted. The shortage of equipment is often the reason that shelf-life and service-life testing is incomplete, leaving the ILS evaluator with insufficient data to predict future performance of the test item. Some evaluations must measure performance against a point on the parameter's growth curve. The ILS testing will continue postproduction to obtain required sample sizes for verifying performance criteria. Many aspects of the logistic support system may not be available for IOT&E and become testing limitations. The PMO must develop enough logistic support to ensure the user can maintain the system during IOT&E without requiring system contractor involvement. Any ILS limitations upon IOT&E will likely be evaluated during FOT&E.

20.5 SUMMARY

Test and evaluation are the logisticians' tools for measuring the ability of the planned support system to fulfill the materiel system's readiness and supportability objectives. The effectiveness of ILS T&E is based upon the completeness and timeliness of the planning effort.

The ILS T&E requirements must be an integral part of the TEMP to ensure budgeting and scheduling of required test resources. Data requirements must be completely identified, with adequate plans made for collection, storage, retrieval and reduction of test data. At MS II, decision-makers can expect that some ILS performance parameters will not have finished testing because of the large sample sizes required for statistical analysis.

21

EC/C3 TEST AND EVALUATION

21.1 INTRODUCTION

Testing of electronic combat (EC) and command, control and communications (C3) systems pose unique problems for the tester because of the difficulty in measuring their operational performance. Special testing techniques and facilities are normally required in EC and C3 testing. This chapter discusses the problems associated with EC and C3 testing and presents methodologies the tester can consider using to overcome the problems.

21.2 TESTING EC SYSTEMS

21.2.1 Special Consideration When Testing EC Systems

Electronic combat systems operate across the electromagnetic spectrum, performing offensive and defensive support roles. Configurations vary from subsystem components to full-up independent systems. The EC systems are used to increase survivability, degrade enemy capability and contribute to the overall success of the combat mission. Decision-makers want to know the incremental contribution to total force effectiveness made by a new EC system when measured in a force-on-force engagement. However, the contractual specifications for EC systems are usually stated in terms of engineering parameters such as effective radiated power, reduction in communications intelligibility and jamming-to-signal ratio. These measures are of little use by themselves in assessing contribution to mission success. The decision-makers require that testing be conducted under realistic operational conditions; but the major field test ranges, such as the shoreline at Eglin AFB or the desert at Nellis AFB, cannot provide the signal density or realism of threats that would be presented by regional conflicts in central Europe. In field testing, the tester can achieve one-on-one or, at best, few-on-few testing conditions. To do this he needs a methodology that will permit extrapolation of engineering measurements and one-on-one test events to create more operationally meaningful measures of mission success in a force-on-force context, usually under simulated conditions.

21.2.2 Integrated Test Approach

An integrated approach to EC testing using a combination of large-scale models, computer simulations, hybrid man-in-the-loop simulators and field test ranges is a solution for the EC tester. No tool by itself is adequate to provide a comprehensive evaluation. Simulation, both digital and hybrid, can provide a means for efficient test execution. Computer models can be used to simulate many different test cases to aid the tester in assessing the critical test issues (i.e., sensitivity analysis) and produce a comprehensive set of predicted results. As digital simulation models are validated with empirical data from testing, they can be used to evaluate the system under test in a more dense and complex threat environment and at expected wartime levels. In addition, the field test results are used to validate the model; and the model is used to validate the field tests, thus lending more credibility to both results. Hybrid man-in-the-loop simulators, such as the Real-Time Electromagnetic Digitally Controlled Analyzer and Processor (REDCAP) and the Air Force Electronic Warfare Evaluation Simulator (AFEWES) can provide a capability to test against new threats. Hybrid simulators cost less and are safer than field testing. The field test ranges are used when a wider range of actions and reactions by aircraft and ground threat system operations is required.

Where one tool is weak, another may be strong. By using all the tools, an EC tester can do a complete job of testing. An example of an integrated methodology is shown in Figure 21-1. The EC integrated testing can be summarized as:

(1) Initial modeling phase for sensitivity analysis and test planning,

(2) Active test phases at hybrid laboratory simulator and field range facilities,

(3) Test data reduction and analysis,

(4) Post-test modeling phase repeating the first step using test data for extrapolation,

(5) Force effectiveness modeling and analysis phase to determine the incremental contribution of the new system to total force effectiveness.

Another alternative is the electronic combat test process proposed in the "Air Force Electronic Combat Development Test Process Guide," May 1991, issued by what is now the Air Staff T&E Element, AF/TE. The six step process described here is graphically represented by Figure 21-2:.

(1) Deriving test requirements,

(2) Conducting pretest analysis to predict EC system performance,

(3) Conducting test sequences under progressively more rigorous ground- and flight-test conditions,

(4) Processing test data,

(5) Conducting post-test analysis and evaluation of operational effectiveness and suitability,

(6) Feeding results back to the system; development employment process.

As can be seen from Figure 21-3, assuming a limited budget and field test being the most expensive per number of trials, the cost of test trials forces the developer and tester to make trade-offs to obtain the necessary test data. Many more iterations of a computer simulation can be run for the cost of an open-air test.

21.3 TESTING OF C3 SYSTEMS

21.3.1 Special Considerations When Testing C3 Systems

The purpose of a C3 system is to provide a commander with timely and relevant information to support sound decision-making. A variety of problems face the C3 system tester. However, in evaluating command effectiveness, it is difficult to separate the contribution made by the C3 system from the contribution made by the commander's innate, cognitive processes. To assess a C3 system in its operational

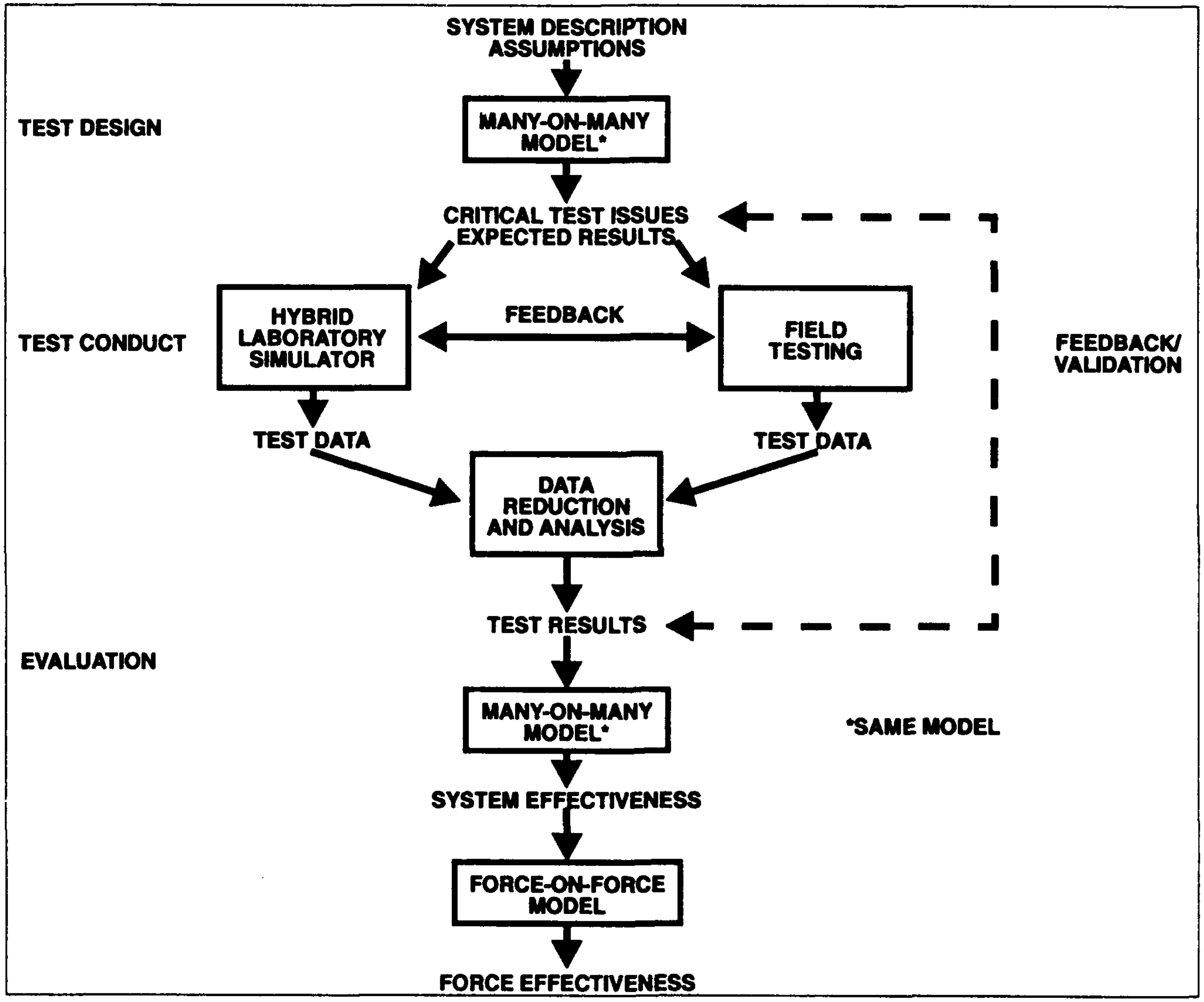
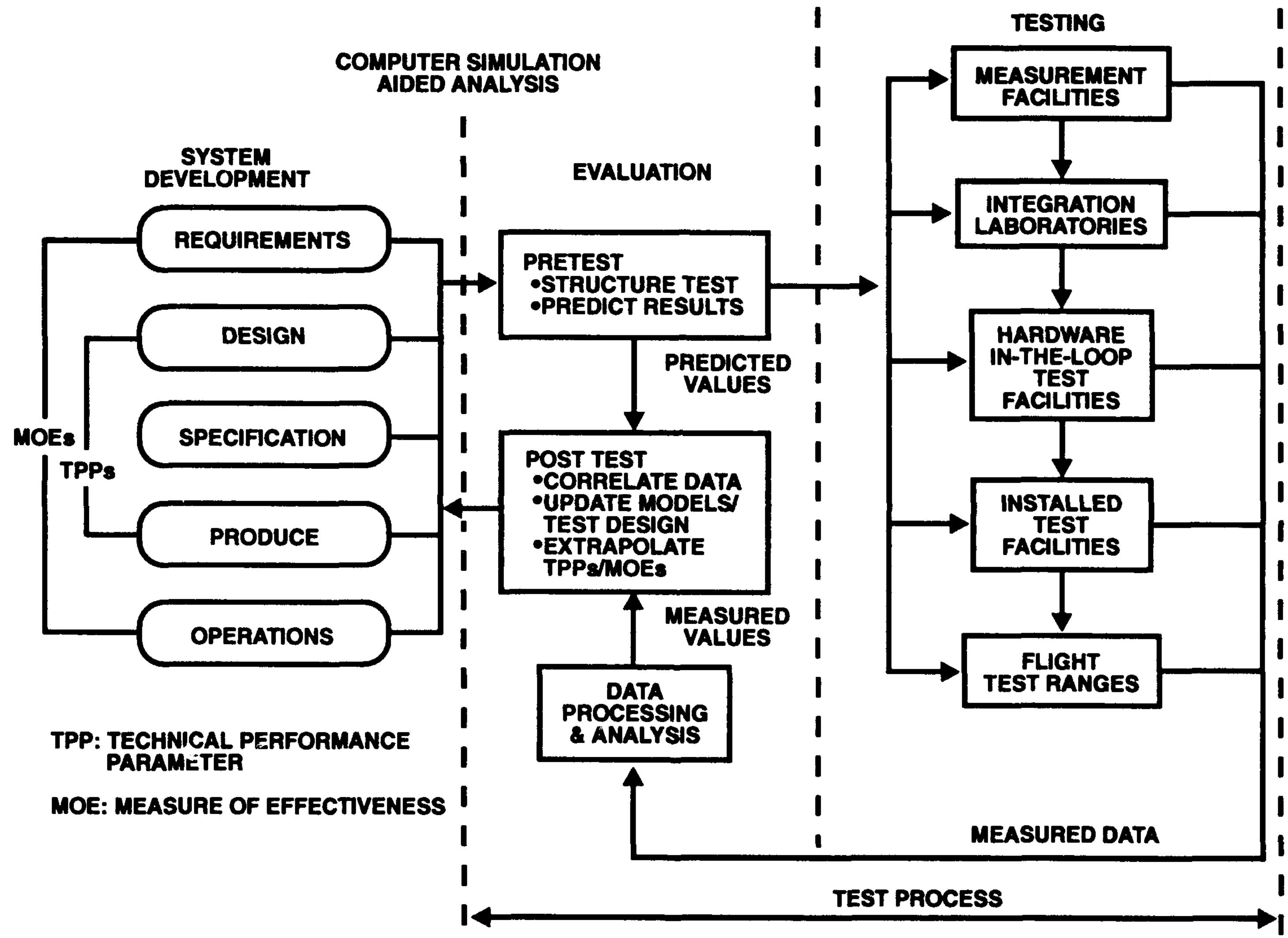
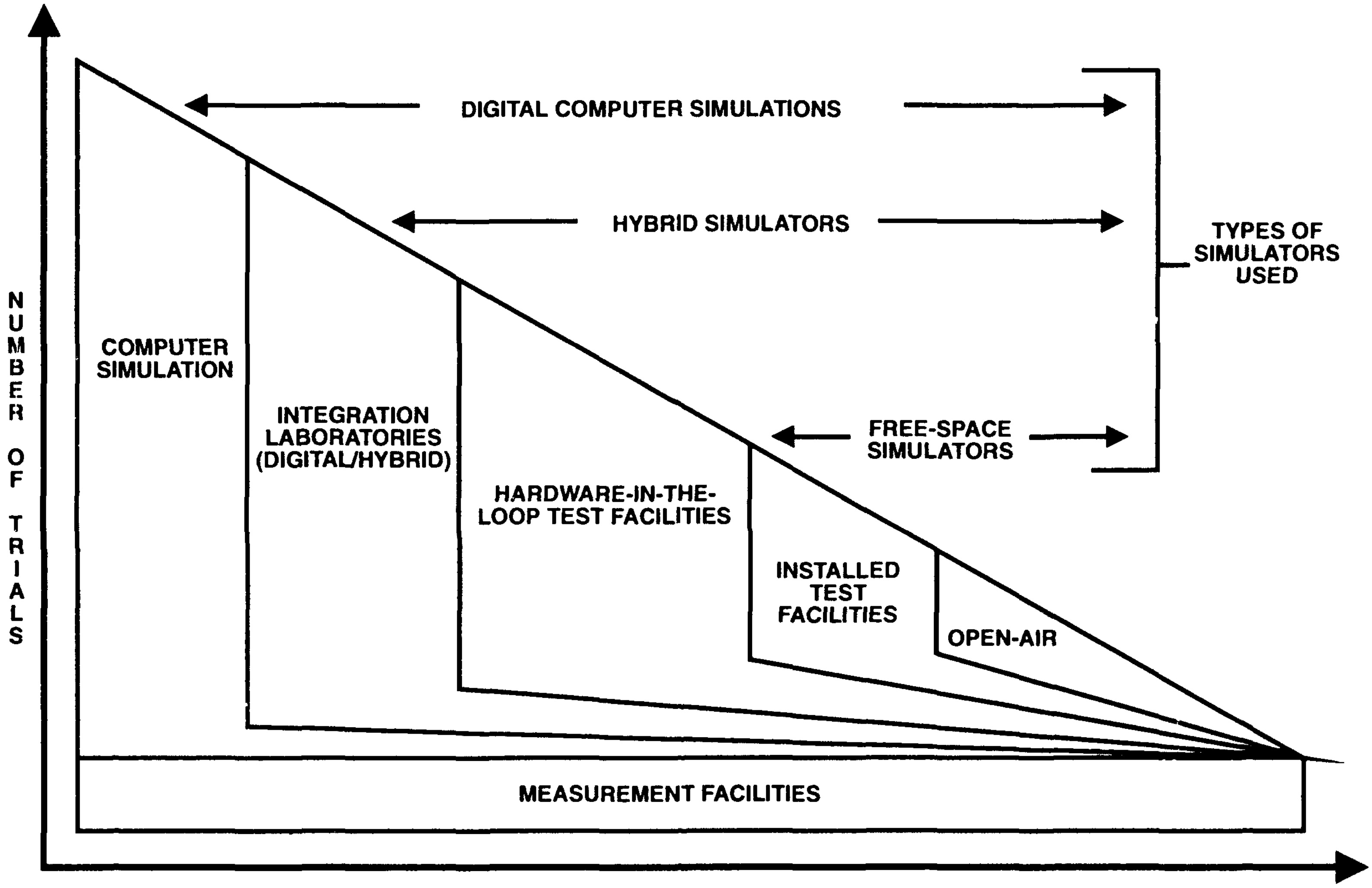


Figure 21-1. Integrated EC Testing Approach



Source: USAF Electronic Combat Development Test Process Guide

Figure 21-2. EC Test Process Concept



Source: USAF Electronic Combat Development Test Process Guide

Figure 21-3. EC Test Resource Categories environment, it must be connected to the other systems with which it would normally operate, making traceability of test results difficult. Additionally, modern C3 systems are software intensive and highly interactive, with complex man-machine interfaces. Measuring C3 system effectiveness thus requires the tester to use properly trained user troops during the test and to closely monitor software test and evaluation (T&E). The C3 systems of the Army, Navy, Air Force and Marines are expected to interoperate with each other and with those of the NATO forces; hence, the tester must also ensure inter-Service and NATO compatibility and interoperability.

21.3.2 C3 Test Facilities

Testing of C3 systems will have to rely more on the use of computer simulations and command, control, communication, intelligence (C3I) test-beds to assess their overall effectiveness. The Joint Tactical Command, Control, and Communications Agency (JTC3A), which is responsible for ensuring interoperability among all U.S. tactical C3 systems that would be used in joint or combined operations, operates the Joint Interoperability Test Center in Ft. Huachuca, Ariz. The center is a test-bed for C3I systems interoperability. Another facility, the huge test-bed developed at Kirtland AFB, N.M., for the Identification Friend, Foe or Neutral (IFFN) Joint Test, will be operated by the Air Force and be available for use by the development and operational communities of all the Services for their C3I testing needs.

21.4 TRENDS IN TESTING C3 SYSTEMS

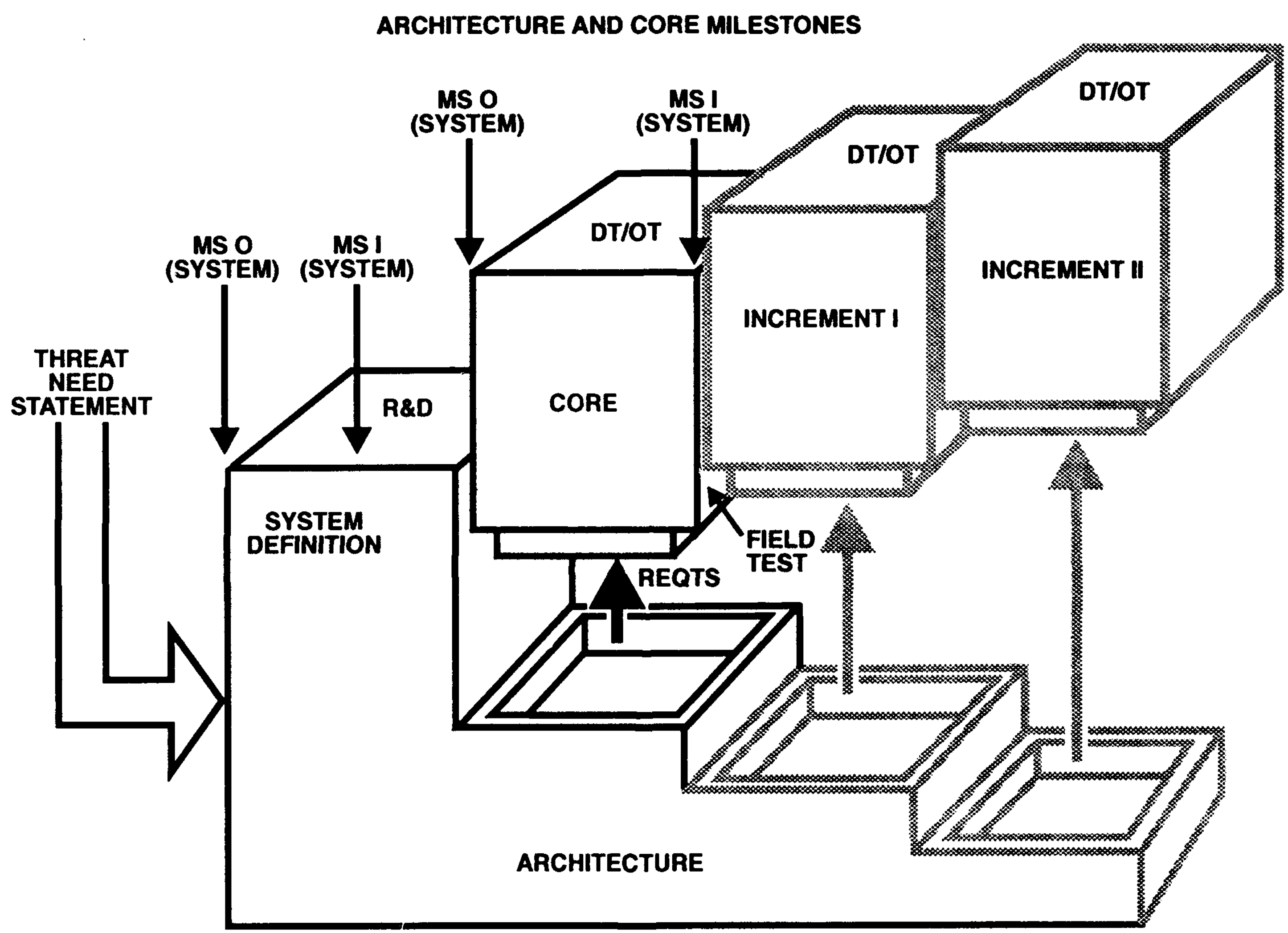
21.4.1 Evolutionary Acquisition of C3 Systems

Evolutionary Acquisition (EA) is a strategy designed to provide an early, useful capability even though detailed overall system requirements cannot be fully defined at the program's inception. The EA strategy contributes to a reduction in the risks involved in system acquisition, since the system is developed and tested in manageable increments. The C3 systems are likely candidates for EA because they are characterized by system requirements that are difficult to quantify or even articulate and that are expected to change as a function of scenario, mission, theater, threat and emerging technology. Therefore, the risk associated with developing these systems can be very great.

Studies by the Defense Systems Management College and the International Test and Evaluation Association (ITEA) have addressed the issues involved in the evolutionary acquisition and testing of C3 systems. The ITEA study illustrated EA in Figure 21-4 and stated that:

> With regard to the tester's role in EA, the study group concluded that iterative test and evaluation is essential for success in an evolutionary acquisition. The tester must become involved early in the acquisition process and contribute throughout the development and fielding of the core and the subsequent increments....The testers contribute to the requirements process through feedback of test results to the user...and...must judge the ability of the system to evolve. [Reference 4]

The testing of EA systems presents the tester with a unique challenge as the core system must be tested during fielding and the first increment before the core testing is completed. This could lead to a situation where the tester has three or four tests ongoing on various increments of the same system. The program manager must insist that the testing for EA systems be carefully



*Brassboard Rapid Prototyping System Simulation, etc.

Source: ITEA Study, "Test & Evaluation of C2 Systems Developed by Evolutionary Acquisition"

Figure 21-4. The Evolutionary Acquisition Process planned to ensure the test data is shared by all and there is a minimum of repetition or duplication in testing.

21.4.2 Radio Vulnerability

The Radio Vulnerability Analysis (RVAN) methodology is for assessing the anti-jam capability and limitations of radio frequency data links when operating in a hostile electronic countermeasures environment. The RVAN evolved from the test methodologies developed for an Office of the Secretary of Defense (OSD)-chartered Joint Test on Data Link Vulnerability Analysis (DVAL). In 1983, OSD directed the Services to apply the DVAL methodology to all new data links being developed.

The purpose of the DVAL methodology is to identify and quantify the antijam capabilities and vulnerabilities of a radio frequency (RF) data link operating in a hostile electronic countermeasures (ECM) environment. The methodology is applied throughout the acquisition process and permits early identification of needed design modifications to reduce identified ECM vulnerabilities. The following four components determine a data link's electronic warfare (EW) vulnerability:

(1) The susceptibility of a data link; i.e., the receiver's performance when subjected to intentional threat ECM;

(2) The interceptibility of the data link; i.e., the degree to which the transmitter could be intercepted by enemy intercept equipment;

(3) The accessibility of the data link; i.e., the likelihood that a threat jammer could degrade the data link's performance;

(4) The feasibility that the enemy would intercept and jam the data link and successfully degrade its performance.

The analyst applying the DVAL methodology will require test data; and the test manager of the C3I system, of which the data link is a component, will be required to provide this data. The DVAL joint test methodologies and test results are on file as part of the Joint Test Library being maintained by the USAF Operational Test and Evaluation Center, Kirtland AFB, N.M.

21.5 SUMMARY

The EC systems must be tested under conditions representative of the dense threat signal environments in which they will operate. The C3 systems must be tested in representative environments where their interaction and responsiveness can be demonstrated. The solution for the tester is an integrated approach using a combination of analytical models, computer simulations, hybrid laboratory simulators and test beds, and actual field testing. The tester must understand these test techniques and resources and apply them in EC and C3 test and evaluation.

22

MULTI-SERVICE TESTS

22.1 INTRODUCTION

This chapter discusses the planning and management of a multi-Service test program. A multi-Service test program is conducted when a system is to be acquired for use by more than one Service or when a system must interface with equipment of another Service. A multi-Service test program should not be confused with the OSD-sponsored, nonacquisition-oriented Joint Test and Evaluation (JT&E) program. A brief description of the JT&E program is provided in Chapter 6.

22.2 BACKGROUND

A definition of multi-Service test and evaluation is contained in DODI 5000.2. It designates the participants in the program and gives a Lead Service responsibility for preparing a single report concerning a system's operational effectiveness and suitability. (The Lead Service is the Service responsible for the overall management of a multi-Service program. A "Supporting Service" is a Service designated to assist the Lead Service.)

A multi-Service test and evaluation (T&E) program may include either development test and evaluation (DT&E) or operational test and evaluation (OT&E) or both. The Service's operational test agencies have executed a formal Memorandum of Agreement on multi-Service OT&E (Reference 35) that provides a framework for the conduct of a multi-Service operational test program.

Air Force Regulation 80-14 describes the procedures followed in a multi-Service T&E program as follows:

(1) In a multiservice acquisition program, T&E is planned and conducted according to Lead Service regulations. The designated Lead Service will have *the overall responsibility for management* of the multi-Service program and will ensure that supporting service requirements are included. If another Service has certain unique T&E requirements, testing for these unique requirements may be planned, funded, and conducted according to that Service's regulations.

(2) Participating Services will prepare reports in accordance with their respective regulations. The Lead Service will prepare and coordinate a single DT&E report and a single OT&E report, which will summarize the conclusions and recommendations of each Service's reports. Rationale will be provided to explain any significant differences. The individual Service reports will be attached to this single report.

(3) Deviations from the Lead Service T&E regulations may be accommo-

AD-A271 595 TEST AND EVALUATION MANAGEMENT GUIDE(U) DEFENSE SYSTEMS
MANAGEMENT COLL FORT BELVOIR VA AUG 93 XD-DSMC

UNCLASSIFIED NL

END
FILMED
DTIC

AIIM

Association for Information and Image Management

1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910

301/587-8202

Centimeter

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 mm

1 2 3 4 5

Inches

1.0 1.1 1.25 1.4 1.6 1.8 2.0 2.2 2.5 2.8 3.2 3.6 4.0 4.5 5.0 5.6 6.3

MANUFACTURED TO AIIM STANDARDS
BY APPLIED IMAGE, INC.

dated by mutual agreement among the Services involved.

22.3 TEST PROGRAM RESPONSIBILITIES

The Lead Service has overall management responsibility for the program. It must ensure that supporting Service requirements are included in the formulation of the basic resource and planning documents.

A Test Management Council (TMC) should be established for each multi-Service test program. Its membership consists of one senior representative from each participating Service or agency headquarters. The TMC works closely with the program management office (PMO) and is responsible for arbitrating all disagreements among Services that cannot be resolved at the working level.

Resource requirements are documented in the Test and Evaluation Master Plan (TEMP). Each participating Service is directed to budget for the testing necessary to accomplish its assigned test objectives and for the participation of its personnel and equipment in the entire test program. Separate annexes may be used to address each Service's test requirements.

22.4 TEST TEAM STRUCTURE

A sample test team structure is shown in Figure 22-1. As shown in the figure, Service test teams work through a Service deputy test director or senior representative. The test director exercises test management authority but not operational control over the test teams. The responsibilities include integration of test requirements and efficient scheduling of test events. The deputy test directors exercise operational control or test management authority over their Service test teams in accordance with their Service directives. Additionally, they act as advisers to the test director; represent their Service's interests; and are responsible, at least administratively, for resources and personnel provided by their Services.

22.5 TEST PLANNING

Test planning for multi-Service T&E is accomplished in the manner prescribed by Lead Service directions and in accordance with the following general procedures extracted from the "Memorandum of Agreement on Multi-Service OT&E and Joint T&E":

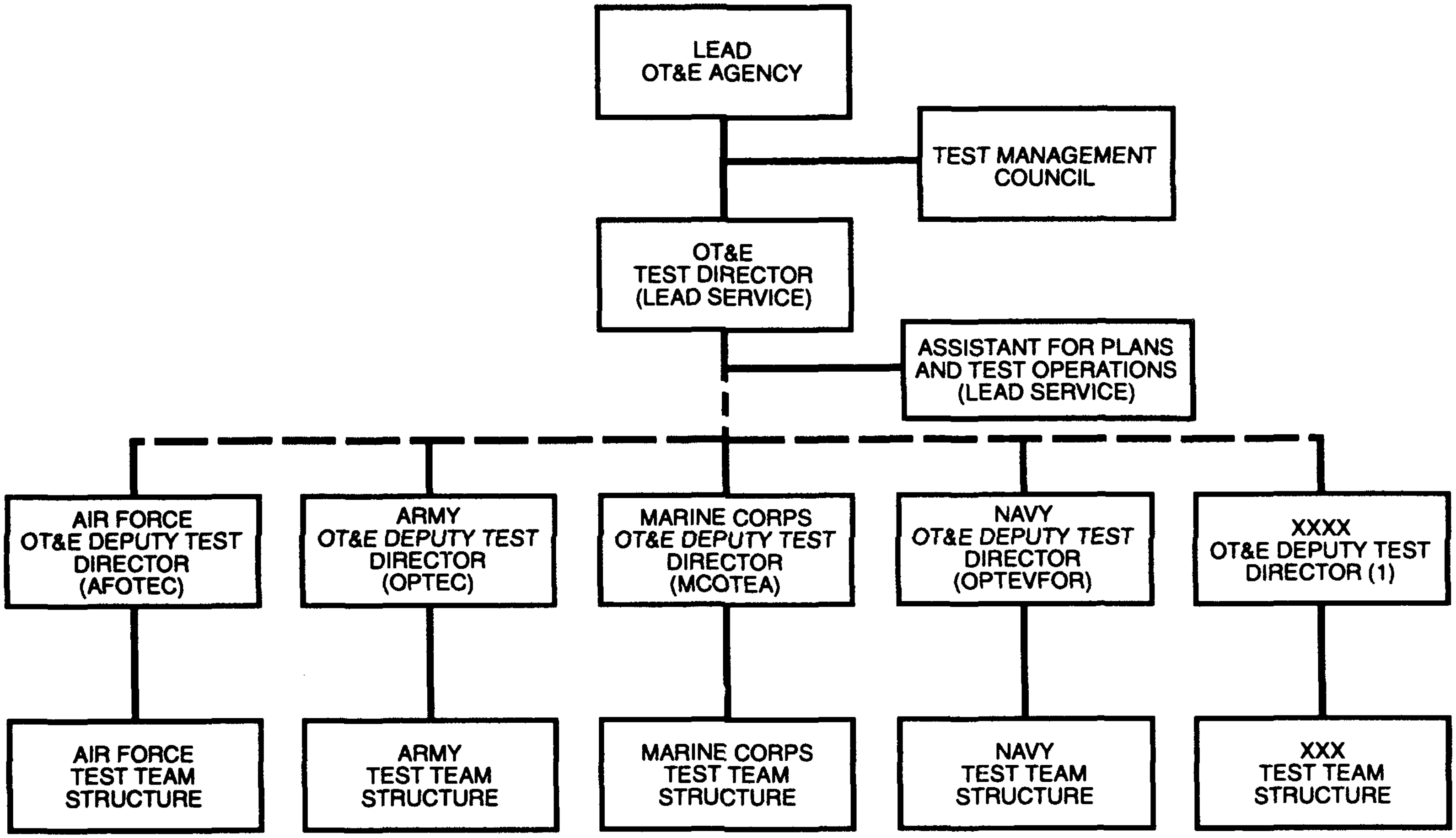
(1) The Lead Service T&E agency begins the planning process by issuing a call to the supporting Service T&E agencies for critical issues and test objectives.

(2) The Lead Service T&E agency consolidates the objectives into a list and coordinates the list with the supporting Service T&E agencies.

(3) The Lead Service T&E agency accommodates supporting Service T&E requirements and input in the formal coordination action of the TEMP.

(4) Participating T&E agency project officers assign responsibility for the accomplishment of test objectives (from the consolidated list) to each T&E agency. These assignments are made in a mutually agreeable manner. Each agency is then responsible for resource identification and accomplishment of its assigned test objectives under the direction of the Lead Service T&E agency.

(5) Each participating agency prepares the portion of the overall test plan(s) for its assigned objectives, in the Lead Service test plan(s) format, and identifies its data needs.



(1) USED FOR COMPLEX PROGRAMS WITH MANY PARTICIPANTS

Source: "Memorandum of Agreement On Multi-Service OT&E and Joint T&E"

Figure 22-1. Simple Multi-Service OT&E Test Team Composition

(6) The Lead Service T&E agency prepares the multi-Service T&E test plan(s), consolidating the input from all participating agencies.

22.6 DISCREPANCY REPORTING

In a multi-Service T&E program, a discrepancy report is a report of any condition that reflects adversely on the item being tested and that must be reported outside the test team for corrective action. The discrepancy reporting system of the Lead Service is normally used. All members of the multi-Service test team will report discrepancies through their Service's system.

Items undergoing test will not necessarily be used by each of the Services for identical purposes. As a result, a discrepancy considered disqualifying by one Service is not necessarily disqualifying for all Services. Discrepancy reports of a disqualifying nature must include a statement by the concerned Service of why the discrepancy has been so classified. It also includes statements by the other Services as to whether or not the discrepancy affects them significantly.

If one of the participating Services identifies a discrepancy that it considers as warranting termination of the test, the circumstances are reported immediately to the test director.

22.7 TEST REPORTING

The following test-reporting policy applies to multi-Service OT&E programs:

(1) Interim test reports are not normally prepared. If they are required on a particular program; they are prepared in accordance with Lead Service directives and coordinated with all participating OT&E agencies prior to release.

(2) Within 60 days of the end of testing, the multi-Service OT&E team must present a factual report of the test to all participating OT&E agencies. (This factual report presents the data collected but no evaluation, conclusions or recommendations concerning the data.)

(3) Each participating OT&E agency prepares an independent evaluation report in its own format and forwards that report through its normal Service channels.

(4) Approved independent evaluation reports are distributed to all participating OT&E agencies.

(5) The Lead Service OT&E agency is responsible for preparing the Defense Acquisition Board (DAB) briefing(s) which is (are) coordinated with all participating OT&E agencies.

22.8 SUMMARY

Multi-Service test programs are conducted by two or more Services when a system is to be acquired by more than one Service or when a system must interface with equipment of another Service. Test procedures for multi-Service T&E follow those of the designated Lead Service, with mutual agreements resolving areas where deviations are necessary. Care must be exercised when integrating test results and reporting discrepancies since items undergoing testing may be used for different purposes in different Services. Close coordination is required to ensure that an accurate summary of the developing system's capabilities is provided to Service and DOD decision authorities.

23

INTERNATIONAL TEST AND EVALUATION PROGRAMS

23.1 INTRODUCTION

This chapter discusses test and evaluation (T&E) from an international perspective. It describes the Office of the Secretary of Defense (OSD)-sponsored Foreign Comparative Test Program. Factors that bear on the T&E of multinational acquisition programs are discussed also.

23.2 FOREIGN COMPARATIVE TEST PROGRAM

23.2.1 Program Objective

The Foreign Comparative Test (FCT) Program is designed to support the evaluation of a foreign nation's weapons system, equipment or technology in terms of its potential to meet a valid requirement of one or more of the U.S. Armed Services. Additional goals of the FCT program include avoiding unnecessary duplication in development, enhancing standardization and interoperability and promoting international technology exchanges. The FCT program is not intended for use in exploiting threat systems or for intelligence gathering. The primary objective of the program is to reduce the costs of research and development, while leading to the acquisition of foreign equipment for U.S. use. Policy and procedures for the execution of the program are documented in DOD 5000.3-M-2.

23.2.2 Program Administration

Foreign weapons evaluation activities and responsibilities are assigned to the Director, Test and Evaluation (DTE) by direction of the Congress in 1980. Each year, sponsoring military services forward Candidate Nomination Proposals (CNPs) for systems to be evaluated under the FCT program to the DTE. The Services are encouraged to prepare and submit a CNP whenever a promising candidate that appears to satisfy a current or potential Service requirement is found. A CNP must contain the information as required by DOD 5000.3-M-2.

The fundamental criterion for FCT program selection is the candidate system's potential to satisfy operational or training requirements that exist or are projected. Its possible contribution to the U.S. technology base is considered also. Additional factors influencing candidate selection include: candidate maturity, available test data, multi-Service interest, existence of a statement of operational requirement need, potential for subsequent procurement, sponsorship by U.S.-based licensee, realistic evaluation schedule cost, DOD component OSD evaluation cost-sharing proposal, and preprogrammed procurement funds. For technology evaluation programs within the FCT program, the candidate nomination proposal must address the specific arrangements under which the United States and foreign participants (governments, armed forces, corporations) will operate. These may include government-to-government Memoranda of Agreement, private industry licensing agreements, data exchange agreements and/or cooperative technology exchange programs.

Foreign weapons evaluation activities are funded by OSD and executed by the Service with the potential need for the system. Points of contact at the headquarters level in each Service monitor the conduct of the programs. Work is performed in laboratories and test centers throughout the country. Systems evaluated recently under the FCT program include millimeter wave communications equipment, chemical defense equipment, gunnery devices, maritime decoys and navigational systems.

23.3 NATO COMPARATIVE TEST PROGRAM

The NATO Comparative Test Program has been integrated with the FCT program. It was created by an act of the Congress in the FY 1986 Defense Authorization Bill. The program supports the evaluation of NATO nations' weapons systems, equipment and technology and assesses their suitability for use by U.S. forces. The selection criteria for the NATO Comparative Test Program are essentially the same as for the FCT program. The exception is that the equipment must be produced by a NATO member nation and be considered as an alternative to a system that is either in a late stage of development in the United States or that offers a cost, schedule or performance advantage over U.S. equipment. In addition, the NATO Comparative Test Program requires that notification be sent to the Armed Services and Appropriations Committees of the House of Representatives and Senate before funds are obligated. With this exception, the NATO Comparative Test Program follows the same nomination process and administrative procedures. Guidelines for the program will also be contained in DOD 5000.3-M-2.

Examples of proposals funded under the NATO Comparative Test Program include T&E of a German mine reconnaissance and detection system for the Army, a United Kingdom-designed minehunter for the Navy, and the Norwegian Penguin missile system for the Air Force. According to the FY 1988 Report of the Secretary of Defense to the Congress, the program has generated considerable interest among NATO allied nations and has become a primary way of promoting armaments cooperation within NATO.

Problems associated with testing foreign weapons normally stem from politics, national pride and a lack of previous test data. When foreign companies introduce weapon systems for testing, they often will attempt to align the U.S. military/congressional organizations with their systems. For example, when a foreign nation introduced an antitank weapon to the Army, they did so by having a U.S. Senator write the Army stating a need for the system. Attached to the letter was a document containing doctrine to employ the system and a test concept to use when evaluating the system. Systems tested in the NATO Comparative Test Program often become involved in national pride. The test community must be careful not to allow national pride to be a driving force in the evaluation. At times, the 9mm pistol competition in NATO resembled an international soccer match, with each competing nation cheering for their pistol and many other nations selecting sides. Evaluating the 9mm pistol was difficult because of these forces. Thus, U.S. testers must make every effort to obtain all available test data on foreign systems. These data can be used to help validate the evolving test data and additional test data during the evaluation.

23.4 T&E MANAGEMENT IN MULTINATIONAL PROGRAMS

Rationalization, standardization and interoperability have become increasingly important elements in the materiel acquisition process. Public Law 94-361, passed on July 14, 1976, requires that "equipment for use of personnel of the Armed Forces of the United States stationed in Europe under the terms of the North Atlantic Treaty should be standardized or at least interoperable with equipment of other members of the North Atlantic Treaty Organization" (Reference 4, pages 1-2). Program Managers and test managers must, therefore, be fully aware of any potential international applications of the systems for which they are responsible. The *Joint Logistics Commanders Guide for the Management of Multinational Programs* published by the Defense Systems Management College (Reference 47) is a valuable compendium of information for the program manager of a developing system with potential multinational applications.

Representatives of the United States, United Kingdom, France and Germany have signed a Memorandum of Agreement concerning the mutual acceptability of each country's T&E data. This agreement seeks to avoid redundant testing by documenting the extent of understanding among involved governments concerning mutual acceptability of respective T&E procedures for systems that are developed in one country and are candidates for procurement by one or more of the other countries. Focal points for development and operational testing in each of the countries are identified, and procedures governing generation and release of T&E data are described in the Memorandum of Understanding (MOU).

Early and thorough planning is an important element of any successful T&E program but is even more critical in a multinational program. Agreement must be reached concerning T&E procedures, data requirements and methodology. Differences in tactics, battlefield representations and military organizations may make it difficult for one nation to accept another's test data. Therefore, agreement must be reached in advance concerning the operational test scenario and battlefield representation that will be used.

23.5 U.S. AND NATO ACQUISITION PROGRAMS

Some test programs involve combined development and test of new weapon systems for U.S. and other NATO countries. In these programs, some differences from the regular "way of doing things" occur. For example, the formulation of the Request for Proposal (RFP) must be coordinated with the North Atlantic Program Management Agency (NAPMA); and their input to the Statement of Work, data requirements, operational test planning and test schedule formulation must be included. Also, the U.S. Army operational user, Forces Command, must be involved in the operational test program. Usually, a Multinational Memorandum of Understanding (MMOU) is created concerning test program and production funding, test resources, test team composition, use of national assets for testing, etc.

Nations are encouraged to use the data that another nation has gathered on similar test programs to avoid duplication of effort. For example, during the U.S. and NATO Airborne Warning and Control System (AWACS) ESM Program, both U.S. and NATO E-3As will be used for test aircraft in combined development test and evaluation (DT&E) and subsequent operational test and evaluation (OT&E). Testing will be conducted in the U.S. and European theaters. The Joint Test Force will be composed of program management office, contractor, U.S. operational users, Air Force Operational Test and Evaluation Center (AFOTEC), Force Command (NATO users), and logistics personnel for this program. A Multinational Memorandum of Agreement for this program was created. The U.S. program is managed by the AWACS System Program Office, and the NATO program is managed by the NAPMA.

23.6 SUMMARY

The procurement of weapon systems from foreign nations for use by U.S. Armed Forces can provide the following advantages: reduced research and development costs, faster initial operational capability, improved interoperability with friendly nations, and lower procurement costs because of economies of scale. Testing such systems presents specific challenges to accommodate the needs of all users. Such testing requires careful advance planning and systematic execution. Expectations and understandings must be well documented at an early stage to ensure that the test results have utility for all concerned.

24

NONDEVELOPMENT ITEMS

24.1 INTRODUCTION

Many options are available when an acquisition strategy for a new system is chosen. They range from the last option of a traditional new research and development program to modification of the existing system. Between these two extremes are other acquisition strategies that call for using nondevelopment items to various extents. Figure 24-1, an adaptation of an illustration found in Army Materiel Command Pamphlet 70-2, shows the broad spectrum of approaches that can be taken in a system acquisition and provides examples of systems that have been developed using each approach.

24.1.1 Definition of NDI

A nondevelopmental item (NDI) refers to materiel coming from a variety of sources but involving little or no development effort. It includes commercial products, materiel developed by other U.S. government sources, or materiel developed in other countries. All such systems are required to undergo technical and operational test and evaluation (T&E) before the procurement decision, unless the decision authority makes a definitive decision that previous testing or other data (such as user/market investigations) provide sufficient evidence of acceptability (Reference 54). A nondevelopmental item is:

- Any item available in the commercial marketplace;
- Any previously developed item in use by a federal, state or local agency of the United States or a foreign government with which the United States has a mutual defense cooperation agreement;
- Any item described above that requires only minor modification to meet the requirements of the procuring agency;
- Any item currently being produced that does not meet the three requirements above solely because it is not yet in use or available in the commercial marketplace. (DODI 5000.2, 6-L)

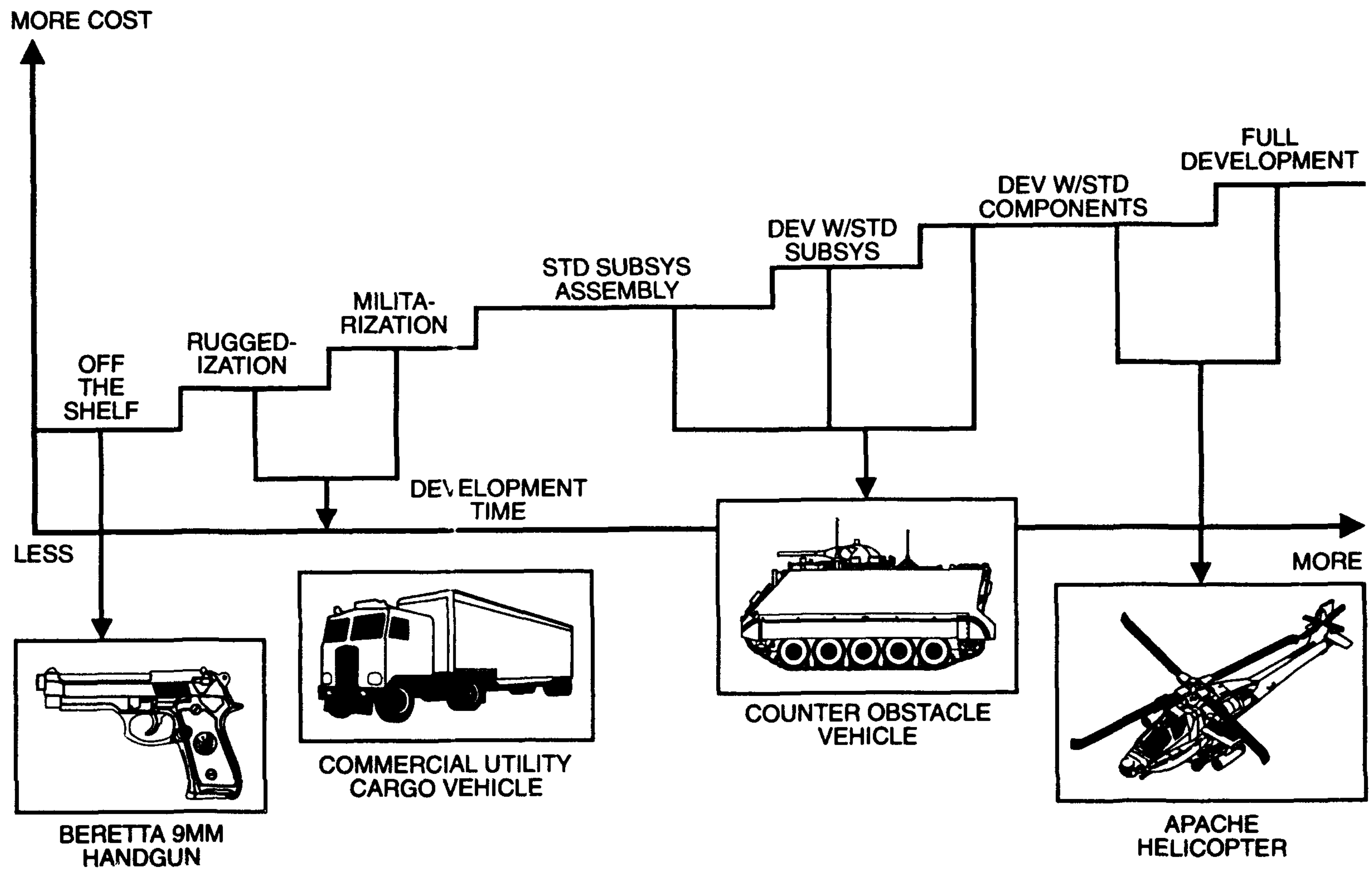
24.1.2 Advantages and Disadvantages of the NDI Approach

The use of NDI offers the following advantages:

- The time to field a system is greatly reduced, and a quick response is provided to the user's needs;
- Research and development costs are reduced;
- State-of-the-art technology is available immediately.

NDI offers the following disadvantages:

- NDI acquisitions are difficult to standardize with the current fleet equipment;



Source: Army Materiel Command Pamphlet 70-2, "AMC-TRADOC Materiel Acquisition Handbook," 26 March 1987.

Figure 24-1. The Spectrum of Acquisition Strategies

- NDI acquisitions create logistics support difficulties;

- NDI acquisitions tend not to have competition; therefore, the availability of second source is not present;

- With NDI acquisitions, engineering and test data often is not available.

24.1.3 Types of NDI

Nondevelopment items can be separated into two general categories; each requires a modified testing approach. The categories are:

(1) Commercial off-the-shelf items for use in the same environment for which the items were designed. Such items normally do not require development testing prior to the production qualification test except in those cases where a contract may be awarded to a contractor who has not previously produced an acceptable finished product and the item is assessed as high risk. In that case, preproduction qualification testing would be required (Reference 54).

(2) Commercial off-the-shelf items for use in an environment other than that for which the items were designed. Such items may require modifications in hardware and/or software. These items require testing in an operational environment, preproduction qualification testing (if previous testing resulted in item redesign), and production qualification testing.

Existing components that must be integrated with a new system configuration may be purchased off the shelf. These would not be classified as NDIs, but many of the testing and evaluation methods would still apply. This type of NDI effort requires more extensive research, development and testing to achieve the desired system configuration. Testing required includes: feasibility testing in a military environment, preproduction qualification testing, hardware/software integration testing, operational testing and production qualification testing.

Given the variety of NDI approaches that may be employed, it is imperative that the acquisition strategy clearly specifies, with the agreement of the testing authority, the level of testing that will be performed on NDI systems and the environment in which those systems will be tested.

24.2 MARKET INVESTIGATION AND PROCUREMENT

A market investigation is the central activity leading to the Milestone I review decision regarding the use of an NDI acquisition strategy. The purpose of the market investigation is to determine the nature of available products and the number of potential vendors. Market investigations may vary from informal telephone inquiries to comprehensive industry-wide reviews. During the market investigation, sufficient data must be gathered to support a definitive NDI decision, to finalize the requirements and to develop an acquisition strategy that is responsive to these requirements.

During the Market Investigation Phase, a formal "request for information" process may be followed wherein a brief narrative description of the requirement is published and interested vendors are invited to respond. Test samples or test items may be leased or purchased at this time to support the conduct of operational suitability tests, to evaluate the ability of the equipment to satisfy the requirements and to help build the functional purchase description or system specification. This type of preliminary testing should not be used to select or eliminate any particular vendor or product unless it is preceded by competitive contracting procedures (Reference 61).

It is imperative that technical and operational evaluators become involved during this early stage of an NDI procurement and that they perform an early assessment of the initial issues. The evaluator must also relate these issues to test and evaluation criteria and provide their independent evaluation plans and reports to the decision authorities before the Milestone I decision review.

24.3 NDI TESTING

24.3.1 General Considerations

Test and evaluation must be considered throughout the acquisition of a system that involves NDI. The program manager and his staff must ensure that the testing community is fully involved in the acquisition from the start. The amount and level of testing required depends on the nature of the NDI and its anticipated use; it should be planned to support the design and decision process. At a minimum, T&E of NDI will be conducted to verify integration and interoperability with other system elements. All NDI modifications necessary to adapt them to the weapon system environment will also be subject to T&E. Available test results from all commercial and government sources will determine the actual extent of testing necessary. There are some inherent advantages in NDI testing. For example, an NDI usually encompasses a mature design. The availability of this mature design contributes to the rapid development of the logistics support system that will be needed. In addition, there are more "production" items available for use in a test program. The program manager and his staff must remember that NDI systems also require activity in areas associated with traditional development and acquisition programs. For example, training and maintenance programs and manuals must be developed; and sufficient time should be allowed for their preparation.

When the solicitation package for an NDI acquisition is assembled, the program manager must ensure that it includes the following T&E-related items:

(1) Approved T&E issues and criteria;

(2) A requirement that the offerer provide a description of the testing performed by the contractor on the system, including test procedures followed, data and results achieved;

(3) Production qualification test and quality conformance requirements;

(4) Acceptance test plans for the system and its components.

24.3.2 Testing Before Milestone I

An important advantage of using an NDI acquisition strategy is reduced acquisition time. Consequently, it is important that testing not be redundant and that it is limited to the minimum effort necessary to obtain the required data. Testing can be minimized by:

(1) Obtaining and assessing contractor test results;

(2) Obtaining usage/failure data from other customers;

(3) Observing contractor testing;

(4) Obtaining test results from independent test organizations (e.g., Underwriter's Laboratory);

(5) Verifying selected contractor test data.

If it is determined that more information is needed after the initial data collection from the above sources, NDI candidates may be bought or leased, and technical and operational tests may be conducted.

24.3.3 Testing After Milestone I

All testing to be conducted after the initial milestone decision to proceed with the NDI acquisition should be described in the Acquisition Strategy and the Test and Evaluation Master Plan. Development testing is conducted only if specific information that cannot be satisfied by contractor or other test data sources is needed. Operational testing is conducted as needed. The independent operational test and evaluation agency should concur in any decisions to limit or eliminate operational testing.

Test and evaluation continue even after the system has been fielded. This testing takes the form of a follow-on evaluation to validate and refine: operating and support cost data; reliability, availability, and maintainability characteristics; logistic support plans; and training requirements, doctrine and tactics.

24.4 RESOURCES AND FUNDING

Programming and budgeting for an NDI acquisition present a special challenge. Because of the short duration of the NDI acquisition process, the standard lead times required in the normal Planning, Programming, and Budgetary System cycle may be unduly restrictive. This situation can be minimized through careful, advanced planning and, in the case of urgent requirements, reprogramming/supplemental funding techniques.

Research, Development, Test and Evaluation (RDT&E) funds are normally used to support the conduct of the Market Investigation Phase and the purchase or lease of NDI candidates required for T&E purposes. The RDT&E funds are also used to support T&E activities such as: modification of the NDI; purchase of specifications, manufacturer's publications, repair parts, special tools and equipment; transportation of the NDI to and from the test site; and training, salaries and temporary duty costs of T&E personnel. Procurement, operations and maintenance funds are usually used to support production and deployment costs.

One chief advantage of using an NDI acquisition strategy is reduced overall cost. Additional cost savings can be achieved after a contract has been awarded if the program manager ensures that incentives are provided to contractors to submit value engineering change proposals to the government when unnecessary costs are identified.

24.5 SUMMARY

The use of nondevelopment items in a system acquisition can provide considerable time and cost savings. The testing approach used for an NDI acquisition must be carefully tailored to the type of system and the amount of test data already available. The T&E community must get involved early in the process so that all test issues are adequately addressed and timely comprehensive evaluations are provided to decision authorities.

25

TESTING THE SPECIAL CASES

25.1 INTRODUCTION

This chapter covers the special factors and alternative test strategies the tester must consider in testing dangerous or lethal weapons, systems that involve one-of-a-kind or limited production and systems with high-cost and/or special security considerations. Examples include chemical and laser weapons; ships; space weapons; missile systems; and electronic warfare (EW), command and control (C2) and intelligence systems.

25.2 TESTING OF HAZARDOUS WEAPONS

The tester of dangerous or lethal systems, like chemical and laser weapons, must consider various safety, health and medical factors in deveioping test plans, such as:

(1) Provision of medical facilities for pre- and post-test checkups and emergency treatment;

(2) Need for protective gear for participating/observer personnel;

(3) Approval of the test plan by the Surgeon General;

(4) Restrictions in selection of test participants (e.g., medical criteria or use of only volunteer troops);

(5) Restricted test locations;

(6) Environmental Impact Statements.

Also, the tester must allow for additional planning time, test funds and test resources to accommodate such factors.

25.2.1 Chemical Weapons Testing

The testing of chemical weapons poses unique problems, because the tester cannot perform actual open-air field testing with real nerve agents or other toxic chemicals. Since the United States signed and ratified the Geneva Protocol of 1925, U.S. policy has been that the United States will never be the first to use lethal chemical weapons; it may, however, retaliate with chemical weapons if so attacked. In addition to the health and safety factors discussed in the last paragraph, test issues the chemical weapons tester must address include:

(1) All possible chemical reactions due to variations such as moisture, temperature, pressure and contamination;

(2) Physical behavior of the chemical; i.e., droplet size, dispersion density and ground contamination pattern when used operationally;

(3) Toxicity of the chemical; i.e., lethality and duration of contamination when used operationally;

(4) Safety of the chemical weapon during storage, handling and delivery;

(5) Decontamination process.

Addressing all of these issues requires a combination of laboratory toxic chamber tests and open-air field testing. The latter must be performed using "simulants," which are substances that replicate the physical and chemical properties of the agent but with no toxicity.

The development and use of simulants for testing will require increased attention as more chemical weapons are developed. Chemical agents can demonstrate a wide variety of effects depending on such factors as moisture, temperature and contamination. Consequently, the simulants must be able to replicate all possible agent reactions; it is likely that several simulants would have to be used in a test to produce all predicted agent behaviors. In developing and selecting simulants, the tester must thoroughly understand all chemical and physical properties and possible reactions of the agent.

Studies of the anticipated reactions can be performed in toxic-chamber tests using the real agent. Here, factors such as changes in moisture, temperature, pressure and levels of impurity can be controlled to assess the agent's behavior. But, the tester must think through all possible environmental conditions in which the weapon could operate so all cases can be tested in the laboratory chamber with the real agent. For example, during development testing of the BIGEYE chemical weapon, it was found that higher-than-expected temperatures due to aerodynamic heating caused pressure buildup in the bomb body that resulted in the bomb exploding. This caused the operational concept for the BIGEYE to be changed from on-board mixing of the two chemicals to mixing after release of the bomb.

Tests to confirm toxicity must be conducted using simulants in the actual environment. Since the agent's toxicity is dependent on factors such as droplet size, dispersion density, ground contamination pattern and degradation rate, a simulant that behaves as the agent does must be used in actual field testing. Agent toxicity is determined in the lab.

The Services publish a variety of technical documents on specific chemical test procedures. Documents such as the U.S. Army Test and Evaluation Command (TECOM) Pamphlet 310-4, a bibliography that includes numerous reports on chemical testing issues and procedures, can be consulted for specific documentation on chemical testing.

25.2.2 Laser Weapons Testing

Many new weapon systems are being designed with embedded laser range finders and laser designators. Because of the danger to the human eye posed by lasers, the tester must adhere to special safety requirements and utilize special locations during test and evaluation (T&E). For instance, the only Army installation in the continental United States permitting free-play airborne laser testing is Fort Hunter-Liggett, Calif. During tests involving lasers, the airspace must be restricted; and guards must be posted to prevent anyone from accidentally venturing into the area. A potential solution to the safety issue is to develop and use an "eye-safe" laser for testing. The tester must ensure that eye-safe lasers produce the same laser energy as the real laser system.

Another concern of the laser energy weapons tester is the accurate determination of laser energy level and location on the target. Measurements of the laser energy on the target are usually conducted in the laboratory as part of development test (DT). In the field, video cameras are often used to verify that the laser designator did indeed illuminate the target. Such determinations are important when the tester is trying to attribute weapon performance to behavior of the laser, behavior of the guidance system, or some other factor.

A bibliography of Army test procedures, TECOM Pamphlet 310-4, lists several documents that cover the special issues associated with laser testing.

25.3 SPACE-SYSTEM TESTING

From a historical perspective, space-system acquisition has posed several unique problems to the test process (especially the operational test process) that generally fall into four categories: limited quantities/high cost, "block upgrade" approach to acquisition, operating environment (peacetime and wartime), and test environment.

(1) Limited quantities/high cost - Space systems have traditionally involved the acquisition of relatively few (historically, less than 20) systems at extremely "high per-unit costs" (in comparison with more traditional military systems). The high per-unit costs are driven by a combination of high transportation costs (launch to orbit), high life-cycle reliability requirements and associated costs because of the lack of an "on-orbit" maintenance capability and the high costs associated with "leading edge" technologies that tend to be a part of spacecraft design. From a test perspective, this serves to drive space-system acquisition strategy into the "nonstandard" approach addressed below. The problem is compounded by the "block upgrade" approach to acquisition.

(2) Block upgrade approach to acquisition - Due to the "limited buy" and "high per-unit cost" nature of spacecraft acquisition, these systems tend to be procured using a "block upgrade" acquisition strategy. Under this concept, "the decision to deploy" is often made at the front end of the acquisition cycle; and the first prototype to be placed in orbit becomes the first operational asset. As early and follow-on systems undergo ground and on-orbit testing (either development test and evaluation (DT&E) or operatonal test and evaluation (OT&E)), discrepancies are corrected by "block changes" to the next system in the pipeline. This approach to acquisition can perturb the test process as the tester may have no formal milestone decisions to test toward. The focus must change toward being able to influence the design of (and block changes to) systems further downstream in the pipeline. As the first "on-orbit" asset usually becomes the first operational asset, pressure is created from the operational community to expedite (and sometimes limit) testing so a limited operational capability can be declared and the system can begin fulfilling mission requirements. Once the asset "goes operational," any use of it for testing must compete with operational mission needs — a situation potentially placing the tester in a position of relatively low priority. Recognition of these realities and careful "early-on" test planning can overcome many of these problems, but the tester needs to be involved and ready much earlier in the cycle than with traditional systems.

(3) Operating environment (peacetime and wartime) - Most currently deployed space systems and near-term future space systems operate in the military support arena such as tactical warning/attack assessment, communications, navigation, weather and intelligence; and their day-to-day peacetime operating environment is not much different from the war-

time operating environment except for activity level (i.e., message throughput, more objects to track/see, etc.). Historically, space has been a relatively benign battlefield environment because of technology limitations in the capability of potential adversaries to reach into space with weapons. This combination of support-type missions and a battlefield environment that is not much different from the peacetime environment has played a definite role in allowing systems to reach limited operational capability without as much dedicated prototype system-level testing as seen on other systems. However, this situation is likely to change with the advent of systems like the Strategic Defense Initiative (SDI) where actual weapons systems are on alert in space, and day-to-day peacetime operations will not mirror the anticipated battlefield environment as closely. Likewise, the elevation of the battlefield into space and the advancing technologies that allow potential adversaries to reach into space may also change the thrust of how space systems need to be tested in space. An increased need for dedicated on-orbit testing on a type of space range where the battlefield environment will be replicated can be anticipated — a situation similar to the dedicated testing done today on test ranges with Army, Navy and Air Force weapons.

(4) Test environment - The location of space assets in "remote" orbits also compounds the test problem. Space systems do not have the ready access (as with ground or aircraft systems) to correct deficiencies identified during testing. This situation has driven the main thrust of testing into the "prelaunch" ground simulation environment where discrepancies can be corrected before the system becomes inaccessible. However, as mentioned previously, when space-system missions change from a war-support focus to a war-fighting focus and the number of systems required to do the mission increases from the "high reliability/limited number" mode to a more traditional "fairly large number buy" mode, future space-system testing could be expected to become more like the testing associated with current ground, sea and air systems. From a test perspective, this could also create unique "test technology" requirements; i.e., with these systems we will have to bring the test range to the operating system as opposed to bringing the system to the range. Also, because the space environment tends to be "visible to the world" (others can observe our tests as readily as we can), unique test operations security methodologies may be required to allow us to achieve test realism without giving away system vulnerabilities.

In summary, current and near-term future space systems have unique test methodologies. However, in the future, space operations might entail development/deployment of weapon platforms on orbit with lower design-life reliability (because of cost); and day-to-day peacetime operations will not mirror the wartime environment. Thus, space-system testing requirements may begin to more closely parallel those of traditional weapon systems.

25.4 TESTING WITH LIMITATIONS

Certain types of systems cannot be tested using relatively standard T&E approaches for reasons such as a nonstandard acquisition strategy, resource limitations, cost, safety or security constraints. The Test and Evaluation Master Plan (TEMP) must contain a statement that identifies "those factors that will preclude a full and completely realistic operational test...(IOT&E and FOT&E)," such as inability to realistically portray the entire threat, limited resources or locations, safety and system maturity. The impact of these limitations on the test's critical operational issues must also be addressed in the TEMP.

Nonstandard acquisition strategies are often used for one-of-a-kind or limited production systems. Examples of these include space systems; missiles; ships; and EW, C2 and intelligence systems. For one-of-a-kind systems, the production decision is often made prior to system design; hence, testing does not support the traditional decision process. In limited production systems, there are often no prototypes available for test; consequently, the tester must develop innovative test strategies.

25.5 OPERATIONS SECURITY AND T&E

Operations security (OPSEC) issues must be considered in all test planning. The DODI 5000.2 requires the protection of "sensitive design information and test data" throughout the acquisition cycle by:

(1) Protecting sensitive technology;

(2) Eliminating nonsecure transmittal data on and from test ranges;

(3) Providing secure communications linking DOD agencies to each other and to their contractors.

Such protection is obviously costly and will require additional planning time, test resources and test constraints. The test planner must determine all possible ways in which the system could be susceptible to hostile exploitation during testing. For example, announcement of test schedule and location could allow monitoring by unauthorized persons. Knowledge of the locations of systems and instrumentation or test concepts could reveal classified system capabilities or military concepts. Compilations of unclassified data could, as a whole, reveal classified information as could surveillance (electronic or photographic) of test activities or interception of unencrypted transmissions. The T&E regulations of each Service require an operational security plan for a test. A detailed list of questions the test planner can use to identify the potential threat of exploitation is provided in AFR 55-43.

25.6 SUMMARY

All weapon systems tests are limited to some degree, but certain systems face major limitations that could preclude a full and realistic test. The test planners of these special systems must allow additional planning time, budget for extra test resources and devise alternative test strategies to work around testing limitations caused by such factors as security restrictions, resource availability, environmental safety factors and nonstandard acquisition strategies.

26

T&E OF WEAPON SYSTEMS TYPES

26.1 INTRODUCTION

This chapter offers guidance to Department of Defense personnel who plan, monitor and execute test and evaluation (T&E). Checklists for the chapter were obtained from the Defense Science Board Study, *Report of Task Force on Test and Evaluation,* dated April 2, 1974. This excellent study is highly regarded in the T&E community but has become dated; consequently, the Defense Systems Management College decided to update the study findings and include those findings and summary checklists in this management guide.

26.2 GENERAL TEST AND EVALUATION ISSUES

The Defense Science Board (DSB) report presented guidance on T&E at two levels. On a general level it discussed a number of issues that were appropriate to all weapon acquisition programs. These issues, along with a summary discussion, are given below.

26.2.1 Effects of Test Requirements on System Acquisition

The acquisition strategy for the system should allow sufficient time between the end of demonstration testing and procurement, as contracted with limited production decisions, to allow flexibility for modification of plans that will be required. It should ensure that sufficient dollars are available not only to conduct T&E but to allow for additional T&E that is always required due to failure, design changes, etc.; and, it should be evaluated relative to constraints imposed by:

- The level of system testing at various stages of the research, development, test and evaluation (RDT&E) cycle;

- The number of test items available and the schedule interface with other systems needed in the tests, such as aircraft, electronics, etc.;

- The support required to assist in preparing for and conducting tests and analyzing the test results;

- Being evaluated to minimize the so-called T&E gap caused by lack of hardware during the test phase.

26.2.2 Test Requirements and Restrictions

Tests should:

- Have specific objectives;

- List, in advance, actions to be taken as a consequence of the test results;

- Be instrumented to permit diagnosis of the cause of lack of performance including random, design induced, wear out and operator error failure;

- If failures occur, not be repeated without a detailed analysis of the failure.

("Most likely the failure will not go away.")

26.2.3 Trouble Indicators

Establish an early detection scheme to identify program illness.

When a program begins to have trouble, there are indicators that will show up during testing. Some of these indicators are:

- A test failure;
- Any repetitive failure;
- A revision of schedule or incremental funding that exceeds the original plan;
- Any relaxation of the basic requirements such as lower performance.

26.2.4 Requirement For Test Rehearsals

Test rehearsals should be conducted for each new phase of testing.

26.3 SCHEDULING

Specific issues associated with test scheduling are listed below.

26.3.1 Building Block Test Scheduling

The design of a set of tests to demonstrate feasibility prior to the Engineering and Manufacturing Development Phase should be used. This will allow early testing of high-technical-risk items, and subsequent tests can be incorporated into the hardware as the system concept has been demonstrated as feasible.

26.3.2 Component and Subsystem Test Plans

Ensure a viable component and subsystem test plan. Studies show that almost all component failures will be the kind that cannot be easily detected or prevented in full system testing. System failure must be detected and fixed in the component/subsystem stage, as detecting and correcting failure only at the operational test level results in high cost.

26.3.3 Phasing of DT&E and IOT&E

Problems that become apparent in operational testing can often be evaluated faster with the instrumented development test and evaluation (DT&E) hardware. The integrated test plan should provide time and money to investigate test failures and eliminate causes of failures before other, similar tests take place.

26.3.4 Schedule IOT&E to Include System Interfaces with Other Systems

Whenever possible, the initial operational test and evaluation/follow-on operational test and evaluation (IOT&E)/(FOT&E) of a weapon system should be planned to include other systems that must have a technical interface with the new system. For example, the missile should be tested on most of the platforms for which they are programmed.

26.4 RESOURCES FOR TESTING

26.4.1 Identify Test Resources and Instrumentation

As early as possible, but not later than the start of the Engineering and Manufacturing Development Phase, the test facilities and instrumentation requirements to conduct operational tests should be identified and a tentative schedule of test activities prepared. This information is recorded in the Test and Evaluation Master Plan (TEMP) and Service test resource documentation.

26.4.2 Require Multi-Service OT&E

Multi-Service operational test and evaluation (OT&E) should be considered for weapon systems requiring new operational concepts involving other Services. If multi-Service testing is used, an analysis of the impact of demonstration on time and resources needed to execute the multi-Service tests should be conducted before the Milestone II decision.

26.4.3 Military Construction Program Facilities

Some programs cannot be tested without Military Construction Program facilities. To construct these facilities will require long lead times; therefore, early planning must be done to ensure that the facilities will be ready when required.

26.4.4 Test Sample Size

The primary basis for the test-sample size is usually based on one or more of the following:

- Analysis of test objectives;
- Statistical significance of test results at some specified confidence level;
- Availability of test vehicles, items, etc.;
- Support resources or facilities available;
- Time available for the test program.

26.4.5 Test Termination

One should not hesitate to terminate a test before its completion if it becomes clear that the main objective of the test is unachievable (due to hardware failure, unavailability of resources, etc.) or if additional samples will not change the outcome and conclusions of the test.

26.5 COST

26.5.1 Budget for Test

The IPS, TEMP and budgeting documents should be reviewed regularly to ensure that there are adequate identified testing funds relative to development and fabrication funds.

26.5.2 Funds for Correcting Faults Found in Testing

The IPS, TEMP and budgeting documents need careful scrutiny to ensure that there are adequate contingency funds to cover correction of difficulties at a level that matches industry/government experience on the contract. (Testing to correct deficiencies found during testing, without sufficient funding for proper correction, results in Band-Aid approaches, which require corrections at a later and more-expensive time period.)

26.6 PERFORMANCE AND OPERATIONAL ISSUES

26.6.1 Proof of Performance of Human Factors Concepts

At an appropriate time in Concept Exploration and Definition or Demonstration and Validation (DEM/VAL) Phases, T&E should authenticate the human factors concepts embodied in the proposed systems design, examining questions of safety, comfort, man-machine interfaces, as well as the number and skill of personnel required.

26.6.2 Test Planning

A summary of important test planning items that were identified by the DSB is provided below:

- Ensure that the whole system, including the system user personnel, is tested. Realistically test the complete system, including hardware, software, people and all interfaces. Get user involved from the start and understand user limitations;

- Ascertain that sufficient time and test articles are planned. When the technology is stressed, the higher risks require more test articles and time;

- In general, parts, subsystems and systems should be proven in that order before incorporating them into the next higher assembly for more complete tests. The instrumentation should be planned to permit diagnosis of trouble;

- Major tests should never be repeated without an analysis of failure and corrective action. Allow for delays of this nature.

26.7 SPECIFIC WEAPON SYSTEMS TESTING CHECKLIST

The DSB report is the result of the study of past major weapon systems acquisitions. It was hoped that this study would enhance the testing community's understanding of the role that T&E has had in identifying system problems during the acquisition process. In the foreword of the DSB study, the authors made this statement about including the obvious testing activity in their checklist:

> The T&E expert in reading this volume will find many precepts which will strike him as of this type. These items are included because examples were found where even the obvious has been neglected, not because of incompetence or lack of personal dedication by the people in charge of the program, but because of financial and temporal pressures which forced competent managers to compromise on their principles. It is hoped that the inclusion of the obvious will prevent repetition of the serious errors which have been made in the past when such political, economical and temporal pressures have forced project managers to depart from the rules of sound engineering practices....In the long run, taking short cuts during T&E to save time and money will result in significant increases in the overall costs of the programs and in a delay of delivery of the corresponding weapon systems to combatant forces.

26.7.1 Aircraft Systems

26.7.1.1 Concept Exploration and Definition Phase

- *Test Program/Total Costs.* Prior to Milestone I, all phases of the aircraft test program should be considered so the total costs and the development schedules include consideration of all likely activities in the overall program.

- *Test Facilities and Instrumentation.* Prior to Milestone I, the test facilities and instrumentation requirements to conduct tests should be generally identified along with a tentative schedule of test activities.

- *Test Resources and Failures.* Ensure that there are adequate funds, reasonable amounts of time, and acceptable numbers of aircraft planned for the various test program phases, and that provisions are made for the occurrence of failures.

- *System Interfaces.* Consider all aircraft system interfaces, their test requirements, and probable costs at the outset of the Concept Exploration and Definition Phase.

- *Major Weapon Subsystems.* If the aircraft system relies on the successful development of a specific and separately-funded major weapon (such as a gun or missile) in order to accomplish its primary mission, this major subsystem should be developed and tested concurrently with, or prior to, the aircraft.

• *Propulsion System.* If the aircraft program is paced by the propulsion system development, an early advanced-development project for the propulsion may be appropriate for a new concept.

• *Operational Scenario.* A conceptual operational scenario for operation and use of the aircraft should be developed so that general test plans can be designed. This should include purpose, roles and missions, threats, operating environments, logistics and maintenance and basing characteristics. The potential range of values on these aspects should be stated.

• *Evaluation Criteria.* Develop evaluation criteria to be used for selecting the final aircraft system design.

• *Untried Elements.* The aircraft development program should include conclusive testing to eliminate uncertainties of the untried elements.

• *Brassboard Avionics Tests.* The use of brassboard or modified existing hardware to "prove" the concept will work should be seriously scrutinized to ensure that the demonstrations and tests are applicable.

• *Nuclear Weapons Effects.* The subject of nuclear weapons effects should be addressed in the test concept for all aircraft weapons systems where operational suitability dictates that survivable exposure to nuclear weapons effects is a requirement.

26.7.1.2 Demonstration and Validation Phase

• By the end of the phase, T&E plans and test criteria should be established so there is no question on what constitutes a successful test and what performance is required.

• *Milestones and Goals.* Ensure an integrated system test plan that preestablishes milestones and goals for easy measurement of program progress at a later time.

• *Operating Concept and Environment.* The operational concept and the environments in which the aircraft will be expected to operate and be tested in OT&E should be specified.

• *Test Program Building Blocks.* In the DEM/VAL Phase, demonstrate that high-risk technology is in hand. In planning the full-scale development test program, ensure that components and subsystems are adequately qualified for incorporation into the system tests.

• *Technology Concepts.* Each concept to be used in the aircraft system (e.g., aerodynamics, structures, propulsion) should be identified and coded according to prior application, before future research. Tests for each concept should be specified with the effect of failure identified.

• *DT&E/OT&E Plan.* The aircraft DT&E/OT&E test plan should be reviewed to ensure it includes ground and flight tests necessary to safely and effectively develop the system.

• *Test Failures.* The T&E plans should be made assuming there will be failures; they are inevitable.

• *Multi-Service Testing.* When a new aircraft development program requires joint testing during OT&E and prior to Milestone II, the test plan should include the type of tests and resources required from other activities and Services.

• *Traceability.* The aircraft development and test program should be designed and scheduled so if trouble arises, its source can be traced back through the lab tests and the analytical studies.

• *Competitive Prototype Tests.* When a competitive prototype test program using test and operational crews is employed, the aircraft should be compared on the basis of the performance of critical mission.

• *Prototype Similarity to Development and Production Aircraft.* A firm determination should be made of the degree of similarity of the winning prototype (in a competitive prototype program) to the development and production aircraft. Thus, test results that are derived from the prototype in the interim period prior to availability of the engineering development aircraft can be utilized effectively.

• *Prototype Tests.* The prototype aircraft test data should be used to determine where emphasis should be placed in the engineering development program.

• *Inlet/Engine/Nozzle Match.* The aircraft test program should provide for an early and adequate inlet/engine/nozzle match through a well-planned test program, and there should be time programming for corrections.

• *Subsystem Tests.* There should be a balanced program for the aircraft subsystem tests.

• *Propulsion System.* If the aircraft is paced by the propulsion systems development, an early advanced-development project for the propulsion may be appropriate for a new concept.

• *Electromagnetic Interface (EMI) Testing.* Full-scale aircraft systems tests in an anechoic chamber are desirable for some aircraft.

• *Parts Interchange.* Early plans should provide for tests where theoretically identical parts, particularly in avionics, are interchanged to ensure that the aircraft systems can be maintained in readiness.

• *Human Factors Demonstration.* Ensure adequate demonstration of human factors is considered in the test plan.

• *Military Preliminary Evaluation.* Adequate resources should be scheduled for the aircraft Military Preliminary Evaluation (MPE) and a positive program should exist for the utilization of MPE information at the time of OT&E.

• *User Participation.* It is imperative that the operational command actively participate in the DT&E Phase to ensure that user needs are represented in the development of the system.

• *Maintenance and Training Publications.* The aircraft development program should provide for concurrent training of crews and preparation of draft technical manuals to be used by IOT&E maintenance and operating crews.

• *Research and Development (R&D) Completion Prior to IOT&E.* The testing plans should ensure that, before an aircraft system is subjected to IOT&E, the subsystems essential to the basic mission have completed R&D.

26.7.1.3 Engineering and Manufacturing Development Phase (old FSD)

• *Test Design.* Test programs should be designed to have a high probability of early identification of major deficiencies during the DT&E and IOT&E.

• *Data for Alternate Scenarios.* By careful attention to testing techniques, maximize the utility of the test data gathered; aircraft instrumentation; range instrumentation; and data collection, reduction and storage.

• *Test Milestones.* Development programs should be built around testing milestones, not calendar dates.

• *Production Engineering Influence on R&D Hardware.* Encourage that production philosophy and production techniques be brought to the maximum practicable extent into an early phase of the design process for R&D hardware.

• *Running Evaluation of Tests.* Ensure that running evaluations of test are conducted. If it becomes clear that test objectives are unattainable or additional samples will not change the test outcome, ensure that procedures are established for terminating the test.

• *Simulation.* Analysis and simulation should be conducted, where practicable, before each phase of development flight testing.

• *Avionics Mock-up.* Encourage use of a complete avionics system installed in a mock-up of the appropriate section or sections of the aircraft.

• *Escape Systems Testing.* Ensure the aircrew escape system is thoroughly tested with particular attention to redundant features, such as pyrotechnic firing channels.

• *Structural Testing.* Ensure that fatigue testing is conducted on early production airframes. Airframe production should be held to a low rate until satisfactory progress is shown in these tests.

• *Gun Firing Tests.* All forms of ordnance, especially those that create gases, must be fired from the aircraft for external effects (blast and debris), internal effects (shock) and effects on the propulsion (inlet composition or distribution).

• *Post-Stall Characteristics.* Special attention is warranted on the post-stall test plans for DT&E and OT&E.

• *Subsystem Performance History.* During DT&E and IOT&E of aircraft, ensure that a performance history of each aircraft subsystem is kept.

• *Flight Deficiency Reporting.* Composition of flight deficiencies reporting by aircrews, particularly those pertaining to avionics, should be given special attention.

• *Crew Limitations.* Ensure aircrew limitations are included in the tests.

• *Use of Operational Personnel.* Recommend experienced operational personnel help in establishing measures of effectiveness and in other operational test planning. In conducting OT&E, use typical operational aircrews and support personnel.

• *Role of the User.* Ensure that users participate in the T&E phase so their needs are represented in the development of the system concept and hardware.

• *Crew Fatigue and System Effectiveness.* In attack aircraft operational testing and particularly in attack helicopter tests where vibration is a fatiguing factor, ascertain that the tests include a measure of degradation over time.

• *Time Constraints on Crews.* Detailed operational test plans should be evaluated to determine that the test-imposed conditions on the crew do not invalidate the applicability of the collected data.

• *Complete Basic DT&E before Starting OT&E.* Before the weapon system is subjected to IOT&E, all critical subsystems should have completed basic DT&E and significant problems should be solved.

• *Realism in Testing.* Ascertain that final DT&E system tests and IOT&E flight tests are representative of operational conditions.

• *Test All Profiles and Modes.* Tests should be conducted to evaluate all planned operational flight profiles and all primary and backup, degraded operating modes.

• *Update of Operational Test Plans.* Ensure that operational test plans are reviewed and updated, as needed, to make them relevant to evolving concepts.

• *Conduct IOT&E Early.* Ensure that operational suitability tests are planned to attempt to identify operational deficiencies of new systems quickly so fixes can be developed and tested before large-scale production.

• *Missile Launch Tests.* Review the final position fix planned before launching inertial-guided air-to-surface missiles.

• *Mission Completion Success Probability.* Mission completion success probability factors should be used to measure progress in the aircraft test program.

26.7.1.4 Production and Deployment Phase

• *Operational Test Realism.* Ascertain operational testing is conducted under realistic conditions.

• *Design FOT&E for Less-Than-Optimal Condition.* Structure the FOT&E logistical support for simulated combat conditions.

• *New Threat.* Be alert to the need to extend the FOT&E if a new threat shows up.

• *Certification of Ordnance.* Ensure that ordnance to be delivered by an aircraft is certified for the aircraft.

• *Inadvertent Influence of Test.* The FOT&E plans should provide measures of ensuring that actions by observers and umpires do not unwittingly influence trial outcome.

• *Deficiencies Discovered In-Service.* Be aware that in-Service operations of an aircraft system will surface deficiencies which extensive FOT&E probably would not uncover.

• *Lead the Fleet.* Accelerated Service test of a small quantity of early production aircraft is advisable during FOT&E thereafter.

26.7.2 Missile Systems

26.7.2.1 Concept Exploration and Definition Phase

• *Weapon System Interfaces.* Consider significant weapon system interfaces, their test requirements and probable costs at the outset of the Concept Exploration and Definition Phase. Ensure that the program plan assembled before Milestone I includes an understanding of the basic test criteria and broad test plans for the whole program.

• *Number of Test Missiles.* Ensure that there is sufficient time and a sufficient number of test articles to support the program through its various phases. Compare the program requirements with past missile programs of generic similarity. If there is substantial difference, then adequate justification should be provided. The DT&E period on many programs has had to be extended as much as 50 percent.

• *Test and Evaluation Gap.* A T&E gap has been experienced in some missile programs between the time when testing with R&D hardware was completed and the time when follow-on operational suitability testing was initiated with production hardware.

• *Feasibility Tests.* Ensure experimental test evidence is available to indicate the feasibility of the concept and the availability of the technology for the system development.

• *Evaluation of Conceptual and Validation Tests.* Results of tests conducted during the Concept Exploration and Definition and the DEM/VAL Phases, which most likely have been conducted as avionics brassboard, breadboard or modified existing hardware, should be evaluated with special attention.

• *Multi-Service Testing Plans.* When a new missile development program requires multi-Service testing during OT&E, the test plan should include the type of tests and resources required from other activities and Services.

• *Test Facilities and Instrumentation Requirements.* Before Milestone I, the test facilities and instrumentation requirements to conduct tests should be generally identified along with a tentative schedule of test activities.

26.7.2.2 Demonstration and Validation Phase

• *Establish Test Criteria.* By the end of the DEM/VAL phase, test criteria should be established so there is no question on what constitutes a successful test and what performance is expected.

• *Human Factors.* Ensure that the test plan includes adequate demonstration of human factors considerations.

• *Instrumentation Diagnostic Capability and Compatibility.* Instrumentation design, with adequate diagnostic capability and compatibility in DT&E and IOT&E phases, is essential.

• *Provisions for Test Failures.* The DT&E and OT&E plans should include provisions for the occurrence of failures.

• *Integrated Test Plan.* Ensure development of an integrated system test plan that preestablishes milestones and goals for easy measurement of program progress at a later time.

• *Test and Evaluation Requirements.* Ensure that the T&E program requirements are firm before approving an R&D test program. Many missile programs have suffered severe cost impacts as a result of this deficiency. The test plan must include provisions to adequately test those portions of the operational envelope that stress the system including backup and degraded operational modes.

• *Personnel Training Plans.* Ensure that adequate training and certification plans for test personnel have been developed.

• *Test and Engineering Reporting Format.* Include a T&E reporting format in the program plan. Attention must be given to the reporting format in order to provide a consistent basis for T&E throughout the program life cycle.

• *Program-to-Program Cross Talk.* Encourage program-to-program T&E cross talk. Test and evaluation problems and their solutions, as one program, provide a valuable index of lessons learned and techniques for problem resolution on other programs.

• *Status of T&E Offices.* Ensure that T&E offices reporting to the program manager or director have the same stature as other major elements. It is important that the T&E component of the system program office has organizational status and authority equal to configuration management, program control, system engineering, etc.

• *Measurement of Actual Environments.* Thorough measurements should be made to define and understand the actual environment in which the system components must live during the captive, launch and in-flight phases.

• *Thoroughness of Laboratory Testing.* Significant time and money will be saved if each component, each subsystem, and the full system are all tested as thoroughly as possible in the laboratory.

• *Contract Form.* The contract form can be extremely important to the T&E aspects. In one program, the contract gave the contractor full authority to determine the number of test missiles; and in another, the contract incentive resulted in the contractor concentrating tests on one optimum profile to satisfy the incentive instead of developing the performance throughout important areas of the envelope.

• *Participation of Operational Command.* It is imperative that the operational command actively participate in the DT&E phase to ensure that user needs are represented in the development of the system.

26.7.2.3 Engineering and Manufacturing Development Phase (old FSD)

• *Production Philosophy and Techniques.* Encourage that production philosophy and production techniques be brought, to the maximum practicable extent, into an early phase of the design process for R&D hardware. There are many missile programs in which the components were not qualified until the missile was well into production.

• *Operational Flight Profiles.* Tests should be conducted to evaluate all planned operational flight profiles and all primary and backup degraded operating modes.

• *Failure Isolation and Responsive Action.* Does the system test plan provide for adequate instrumentation so missile failures can be isolated and fixed before the next flight?

• *Responsive Actions for Test Failures.* Encourage a closed-loop reporting and resolution process, which ensures that each test failure at every level is closed out by appropriate action; i.e., redesign, procurement, retest, etc.

• *Plan Tests of Whole System.* Plan tests of the whole system including proper phasing of the platform and supporting gear, the launcher, the missile and user participation.

• *Determination of Component Configuration.* Conditions and component configuration during development tests should be determined by the primary objectives of that test. Whenever a nonoperational configuration is dictated by early test requirements, tests should not be challenged by the fact that configuration is not operational.

• *Testing of Software.* Test and evaluation should ensure that software products are tested appropriately during each phase. Software often has been developed more as an add-on than as an integral part of the overall system. Software requirements need the same consideration as hardware requirements in the DEM/VAL Phase.

• *Range Safety Dry Runs.* Ensure the test plan includes adequate test program/ range safety dry runs. The government test ranges have to provide facilities to safely test many different projects.

• *Assemblies/Subsystems Special Requirements.*

  • Seekers and tracking devices,

  • Propulsion subsystems,

  • Connectors and their related hardware,

  • Lanyard assemblies,

  • Safeing, arming, fuzing and other ordnance devices.

• *Review of Air-to-Surface Missile (ASM) Test Position Fixes.* Review the final position fix planned before launching ASMs. There are instances in which the operational test of air-launched missiles utilized artificial position fixes just prior to missile launch.

• *Operator Limitations.* Ensure operator limitations are included in the tests. Most tactical missiles, especially those used in close support, require visual acquisition of the target by the missile operator and/ or an air/ground controller.

• *Test Simulations and Dry Runs.* Plan and use test simulations and dry runs. Dry runs should be conducted for each new phase of testing. Simulation and other laboratory or ground testing should be conducted to predict the specific test outcome. The "wet run" test should finally be run to verify the test objectives. Evaluation of the simulation vs. the actual test results will help to refine the understanding of the system.

• *Component Performance Records.* Keep performance records on components. There are many examples in missile programs that have required parts stock sweeps associated with flight failures and component aging test programs.

• *Tracking Test Data.* Ensure the test program tracks data in a readily usable manner. Reliability and performance evaluations of a missile system should break down the missile's activity into at least the following phases:

  • Prelaunch including, captive carry reliability,

  • Launch,

  • In-flight,

• Accuracy/fuzing.

• *Updating IOT&E Planning.* Periodically update MPE and IOT&E planning during the early R&D phase. Few missile system programs have had adequate user participation with the desired continuity of personnel to minimize the problems of transition from DT&E to OT&E to deployment/utilization.

• *Instrumentation Provisions in Production Missiles.* Encourage built-in instrumentation provisions in production missiles.

• *Constraints on Missile Operator.* Detailed test plans should be evaluated to determine that the test imposed constraints on the missile operator do not invalidate the applicability of the data so collected.

• *Problem Fixes Before Production.* Ensure that operational suitability tests identify operational deficiencies of new systems quickly so fixes can be developed and tested before large-scale production.

• *Flight Tests Representative of Operations.* Ascertain that final DT&E system tests and IOT&E flight tests are representative of operational flights. Some ballistic missile R&E programs have shown high success rates in R&E flight tests; however, when the early production systems were deployed, they exhibited a number of unsatisfactory characteristics such as poor alert reliability and poor operational test-flight reliability.

26.7.2.4 Production and Deployment Phase

• *System Interfaces in Operational Test.* Ensure the primary objective of an operational test is to obtain measurements on the overall performance of the weapon system when it is interfaced with those systems required to operationally use the weapons system.

• *Realistic Conditions for Operational Testing.* Ascertain operational testing is conducted under realistic combat conditions. This means that the offense/defense battle needs to be simulated in some way before the weapon system evaluation can be considered completed. Whether this exercise is conducted within a single Service (as in the test of a surface-to-surface antitank missile against tanks) or among Services (as in the test of an air-to-surface missile against tanks with antiaircraft protection), the plans for such testing should be formulated as part of the system development plan.

• *Testing All Operational Modes.* Ensure the FOT&E plan includes tests of any operational modes not previously tested in IOT&E. All launch modes including degraded, backup modes should be tested in the FOT&E because the software interface with the production hardware system should be evaluated thoroughly. Otherwise, small, easy-to-fix problems might preclude launch.

• *Extension of the FOT&E for New Threats.* Be alert to the need to extend the FOT&E if a new threat arises. Few missile programs perform any kind of testing relatable to evaluating system performance against current or new threats.

• *"Lead-the-Fleet" Production Scheduling.* Lead-the-Fleet missile scheduling and tests should be considered.

• *Test Fixes.* Test fixes result from earlier operational testing. After the IOT&E that identified problem areas in missiles, FOT&E should evaluate these areas primarily to determine the adequacy of the incorporated fixes, particularly if the

IOT&E did not run long enough to test the fixes.

• *FOT&E Feedback to Acceptance Testing.* Ensure that FOT&E results are quickly fed back to influence early production acceptance testing. Production acceptance testing is probably the final means the government normally will have to ensure the product meets specifications. Early acceptance testing could be influenced favorably by a quick feedback from FOT&E to acceptance testing. This is exemplified by a current ASM program where production has reached peak rates, and the IOT&E has not been completed.

26.7.3 Command and Control Systems

26.7.3.1 Concept Exploration and Definition Phase

• *Conceptual Test Philosophy.* The T&E planners must understand the nature of command and control (C2) systems early in the Concept Exploration and Definition Phase. In a complex command and control system, a total systems concept must be developed initially. Total systems life cycle must be analyzed so the necessary requirement for the design can be established.

• *The Importance of Software Testing.* Testers should recognize that software is a pacing item in command and control systems development.

• *Software Test Scheduling - Contractors' Facilities.* Provision should be made for including software T&E during each phase of C2 systems' acquisition. Availability of contractors' facilities should be considered.

• *Evaluation of Exploratory Development Tests.* Care should be exercised in evaluating results of tests conducted during exploratory development of command and control systems. These tests, which most likely have been conducted on brassboard, breadboard or modified existing hardware, should be evaluated with special attention.

• *Feasibility Testing for Field Compilers.* Early test planning should allow for simulating the computer system to test for field use of compilers, where applicable.

• *Evaluation of Test Plan Scheduling.* Milestones should be event-oriented, not calendar-oriented.

• *Type Personnel Needs - Effects on T&E.* A mix of personnel with different backgrounds affecting T&E is required.

• *Planning for Joint-Service OT&E Before Milestone I.* Joint-Service OT&E (multi-Service) should be considered for command and control systems.

26.7.3.2 Demonstration and Validation Phase

• *Test Prototypes.* In C2 systems, prototypes must reasonably resemble final hardware configuration from a functional-use standpoint. When high technical risk is present, development should be structured around the use of one or more test prototypes designed to prove the system concept under realistic operational conditions before proceeding to engineering development.

• *Test Objectives — Critical Issues.* In addition to addressing critical technical issues, T&E objectives during the Concept Demonstration and Validation Phase should address the functional issues of a C2 system.

• *Real-Time Software — Demonstration of "Application Patches."* Tests of real-time C2 systems should include demonstrations of interfaces whereby locally generated application patches are brought into being.

• *Independent Software Test-User Group.* An independent test-user software group is needed during early software qualification testing.

• *System Interfaces.* Critical attention should be devoted to testing interfaces with other C2 systems and to interfaces between subsystems. Particular attention should be devoted to interfaces with other C2 systems and to the interfaces between sensors (e.g., radar units), communications systems (e.g., modems) and the specific processors (e.g., CPUs). Interface with information processing C2 systems must also address data-element and code-standardization problems if data is to be processed on-line.

• *Human Factors.* In a C2 system, human factors must be considered from the earliest prototype designs and testing provided. Testing should be conducted to determine the most efficient arrangement of equipment from the human factor standpoint; e.g., displays should be arranged for viewing from an optimum angle whenever possible; adequate maneuvering room within installation constraints should be allowed considering the number of personnel normally manning the facility; and console-mounted controls should be designed and located to facilitate operation, minimize fatigue and avoid confusion.

• *Degraded Operations Testing.* When the expected operational environment of a C2 system suggests that the system may be operated under less-than-finely-tuned conditions, tests should be designed to allow for performance measurements under degraded conditions.

• *Test-Bed.* The use of a test-bed for study and experimentation with new C2 systems is needed early in the Concept Demonstration and Validation Phase.

• *Software-Hardware Interfaces.* The software-hardware interfaces, with all operational backup modes to a new C2 system, should be tested early in the program.

• *Reproducible Tests.* Test plans should contain a method for allowing full-load message input while maintaining reproducible test conditions.

• *Cost-Effectiveness.* Field-test data is needed during the DEM/VAL Phase for input to cost-effectiveness analyses of C2 systems.

26.7.3.3 Engineering and Manufacturing Development Phase (old FSD)

• *Acquisition Strategy.* The acquisition strategy for the system should:

- Allow sufficient time between the planned end of demonstration testing and major procurement (as opposed to limited procurement) decisions. This provides flexibility for modifying plans, which may be required during the test phases of the program. For instance, because insufficient time was allowed for testing one recent C2 system, the program and the contract had to be modified and renegotiated;

- Be evaluated relative to constraints imposed;

- Ensure that sufficient dollars are available, not only to conduct the planned T&E but to allow for the additional T&E that is always required due to failures, design changes, etc.

• *Problem Indications.* It is important to establish an early detection scheme so management can determine when a program is becoming "ill."

• *Impact of Software Failures.* Prior to any production release, the impact of software failures on overall system performance parameters must be considered.

• *Critical Issues.* IOT&E should provide the answers to some critical issues peculiar to C2 systems. Some critical issues that IOT&E of C2 systems should answer are:

  • Is system mission reaction time a significant improvement over present systems?

  • Is a backup mode provided for use when either airborne or ground system exhibits a failure?

  • Can the system be transported as operationally required by organic transport? (Consider ground, air and amphibious requirements.)

  • Is there a special requirement for site preparation? (For example, survey and antenna siteing.)

  • Can the system be erected and dismantled in times specified? Are these times realistic?

  • Does relocation affect system alignment?

  • Does system provide for operation during maintenance?

  • Can on-site maintenance be performed on shelterless subsystems (e.g., radar units) during adverse weather conditions?

• *Displays.* The display subsystems of a C2 system should provide an essential function to the user. Displays are key subsystems of a C2 system. They provide the link that couples the operator to the rest of the system and are, therefore, often critical to its success.

• *Pilot Test.* A pilot test should be conducted before IOT&E so sufficient time is available for necessary changes.

• *Publications and Manuals.* It is imperative that all system publications and manuals be completed, reviewed and selectively tested under operational conditions before beginning overall system suitability testing.

• *Power Sources.* Mobile, prime power sources are usually provided as government-furnished equipment (GFE) and can be a problem area in testing C2 systems.

• *IOT&E Reliability Data.* The IOT&E can provide valuable data on the operational reliability of a C2 system; this data cannot be obtained through DT&E.

• *Subsystem Tests.* Every major subsystem of a C2 system should have a successful DT&E before beginning overall system operational testing.

• *Communications.* The C2 systems must be tested in the appropriate electromagnetic environment to determine the performance of its communications system.

• *Maintenance.* In IOT&E, maintenance should include: a measurement of the adequacy of the maintenance levels and the maintenance practices; an assessment of the impact that the maintenance plan has on the operational reliability; the accessibility of the major components of the system for field maintenance (e.g., cables and connectors are installed to facilitate access); and verification that the software design for maintenance and diagnostic routines and procedures are adequate and the software can be modified to accommodate functional changes.

• *Continuity of Operations.* The IOT&E should provide for an impact assessment of the failure of any subsystem element of a C2 system on overall mission effectiveness.

• *Imitative Deception.* The IOT&E should provide for tests to assess the susceptibility of the data links of a C2 system to imitative deception.

• *Demonstration of Procedures.* Test plans should include a procedural demonstration whereby the tested C2 system works in conjunction with other systems.

• *Government-Furnished Equipment and Facilities.* Test and evaluation should be concerned about the availability of GFE equipment as specified in the proposed contract.

• *User Participation in T&E.* The varying needs of the user for a C2 system make participation in all phases of T&E mandatory.

26.7.3.4 Production and Deployment Phase

• *First Article Testing.* The preproduction, first article testing and evaluation should be designed and conducted to: (1) confirm the adequacy of the equipment to meet specified performance requirements; (2) confirm the adequacy of the software not only to meet current user needs but to accommodate changing needs; and (3) determine failure modes and rates of the total integrated system. This activity should be followed by FOT&E.

• *Test Planners and Evaluators.* Use the IOT&E personnel in the FOT&E program. The planners and evaluators for the FOT&E of the production system can do a better job if they are involved initially in planning and conducting the IOT&E.

26.7.4 Ship Systems

26.7.4.1 Concept Exploration and Definition Phase

• *Test and Evaluation Master Plan.* Prior to Milestone I, sufficient materiel should be generated to allow for evaluating the overall T&E program.

• *Test Objectives and Critical Issues.* In evaluating the initial test concept, it is important that the test objectives during the time from Milestone I to Milestone II address the major critical issues, especially technological issues.

• *OT&E Phasing.* In evaluating test plans, look favorably on phasing where the OT&E is run parallel to continued DT&E.

• *Test Facilities and Instrumentation Required.* Before Milestone I, the test facilities and instrumentation requirements to conduct developmental and operational tests and a tentative schedule of test activities should be identified.

• *Multiple Approach To Weapon System Development.* Whenever possible, the weapon system concept should not be predicated on the successful development of a single hardware or software approach in the various critical subsystems (unless it has been previously demonstrated adequately).

• *Comparison of New vs. Old System.* The procedure for examining the relative performance of new or modified systems vs. old should be indicated in the T&E plan.

• *Test Support Facilities.* The phasing of test support facilities must be planned carefully, with some schedule flexibility to cover late delivery and other unforeseen problems.

• *Fleet Operating Force Requirements.* The requirement for fleet operating forces for DT&E or OT&E should be assessed early in the program and a specific commitment made as to the types of units to be employed.

• *Mission-Related Measures of Effectiveness.* During the Concept Exploration and Definition Phase of the acquisition of a new class of ship, a study effort should be commenced jointly by the Chief of Naval Operations (CNO) and the Commander, Operational Test and Evaluation Force (COMOPTEVFOR). This effort is to establish mission-related measures of effectiveness, which may be expressed in numerical fashion and may later be made the subject of OT&E to determine how closely the new ship system meets the operational need for which it was conceived.

• *Ship T&E Management.* The management of ship T&E should ensure that test requirements are necessary and consistent relative to systems/subsystem aspects and that the necessary testing is coordinated so that test redundancy does not become a problem.

• *T&E of Large, Integrally-Constructed Systems.* Major subsystems should be proven feasible before firm commitment to a detailed hull design.

26.7.4.2 Demonstration and Validation Phase

• *Authentication of Human Factors Concepts.* Test and evaluation should authenticate the human factors concepts embodied in the proposed systems design, examining questions of safety, comfort, appropriateness of man-machine interfaces, as well as the numbers and skill levels of the personnel required.

• *Acquisition Strategy.* The acquisition strategy for a ship and its subsystems should allow sufficient time between the planned end of demonstration testing and major procurement decisions of government-furnished equipment for flexibility to modify plans (may be required during the test phases of the program).

• *Evaluation of Results of Exploratory Testing.* Results of tests conducted during exploratory development and most likely conducted on brassboards, breadboards or modified existing hardware should be evaluated carefully.

• *Software Testing.* In view of increased dependence upon computers in ship management and tactical operation, software testing must be exceptionally thorough, and integrated software testing must begin as early as possible.

• *New Hull Forms.* When a new type of ship involves a radical departure from the conventional hull form, extensive prototype testing should be required prior to further commitment to the new hull form.

• *Effects of Hull and Propulsion on Mission Capability.* The predicted effects of the proven hull and propulsion system design on the performance of the ship's critical system should be determined.

• *Advances in Propulsion.* Demonstration of the use of new propulsion systems should be conducted prior to making the decision to commit the propulsion systems to the ship in question.

• *Propulsion Systems in Other Classes.* When an engine to be used in the propulsion system of a new ship is already performing satisfactorily in another ship, this is not to be taken as an indication that shortcuts can be taken in propulsion system DT&E, or that no problems will be encountered.

• *The OT&E of Shipboard Gun Systems.* Operational tests of shipboard gun systems should simulate the stress, exposure time and other conditions of battle so that the suitability of the weapon can be evaluated in total.

• *Targets for Antiaircraft Warfare (AAW) OT&E.* Operational test of shipboard AAW weapons demands the use of targets which realistically simulate the present-day threat.

• *Waivers to T&E of Ship Systems.* Waivers to T&E of preproduction models of a system in order to speed up production and delivery should be made only after considering all costs and benefits of the waiver, including those not associated with the contract.

• *Environment Effects on Sonar Domes.* Environmental effects on sonar domes and their self-noise should be tested and evaluated before the domes are accepted as part of the sonar system.

• *Hull/Machinery Testing by Computer Simulation.* In DT&E ships, there will be cases where the best means to conduct evaluations of particular hull and machinery capabilities is through dynamic analysis using computer simulation, with later validation of the simulation by actual test.

• *Operational Reliability.* The OT&E should provide valuable data on the operational reliability of ship weapon systems that cannot be obtained through DT&E.

26.7.4.3 Engineering and Manufacturing Development Phase (old FSD)

• *Initial or Pilot Phase of IOT&E.* Before any operational tests to demonstrate operational suitability and effectiveness are conducted, an initial or pilot test should be conducted.

• *Identify Critical Subsystems.* In planning for the IOT&E of a ship system, the critical subsystems, with respect to mission performance, should be identified.

• *Reliability of Critical Systems.* Test and evaluation should determine the expected reliability at sea of systems critical to the ship's mobility and to the primary and major secondary tasks.

• *Consistency in Test Objectives.* There are various phases in testing a ship system. One should ensure the objectives of one phase are not inconsistent with the objectives of the other phases.

• *Single Screw Ships.* Test and evaluation of the propulsion systems of ships with a single screw should be especially rigorous to determine failure rates, maintenance and repair alternatives.

• *Problems Associated With New Hulls.* Whenever a new hull is incorporated into ship design, a T&E of this hull should be conducted prior to the full-rate production and incorporation of the major weapons subsystems.

26.7.4.4 Production and Deployment Phase

• *Design of Ship FOT&E.* In the testing program of a ship system, it should be recognized that, although it may be designated as a special-purpose ship, in most cases it will be used in a general-purpose role as well.

• *Operational Testing During Shakedown Periods.* The time period for FOT&E of a ship can be used more efficiently if full advantage is taken of the periods immediately after the ship is delivered to the Navy.

• *Fleet Operations in FOT&E.* A great deal of information on the operational effectiveness of a ship can be obtained from standard fleet operations through well-designed information collection, processing and analysis procedures.

• *Ship Antisubmarine Warfare (ASW) FOT&E Planning.* In planning FOT&E of shipboard systems, it is important to recognize the difficulty of achieving realism, perhaps more so than in other areas of naval warfare.

• *Variable Depth Sonar FOT&E.* The behavior of towed bodies of variable depth sonar systems and towed arrays should be tested and evaluated under all ship maneuvers and speeds likely to be encountered in combat.

• *Ship Self-Noise Tests.* The magnetic and acoustic signatures of a ship can be tested accurately only after it is completed.

• *Effect of Major Electronic Countermeasures (ECMs) on Ship Capability.* The FOT&E of a ship should include tests of the effectiveness of the ship when subjected to major ECM.

• *Ship System Survivability.* Follow-on Operational Test and Evaluation of modern ships should provide for the assessment of their ability to survive and continue to fight when subjected to battle damage.

• *Interlocks.* Shipboard electronic systems are designed with interlock switches that open electrical circuits for safety reasons when the equipment cabinets are opened. The FOT&E should be able to detect overdesign as well as minimum design adequacy of the interlock systems.

• *Intraship Communication.* In conducting lead ship trials and evaluations, particular attention should be given to the operational impact resulting from absence, by design, of intraship communications circuits and stations from important operating locations.

26.7.5 Surface Vehicle Systems

26.7.5.1 Concept Exploration and Definition Phase

• *Preparing Test Plans.* It is necessary that a detailed evaluation criteria be established that includes all items to be tested.

• *Validation Test Plans.* Prior to Milestone I, a plan should be prepared for evaluating the overall T&E program. As part of this, a detailed T&E plan for those tests to be conducted before Milestone II to validate the concept and hardware approach to the vehicle system should be developed. The objective of the validation test plan is to fully evaluate the performance characteristics of the new concept vehicle. This test plan cannot be developed, of course, until the performance characteristics are defined.

• *Performance Characteristics Range.* Stated performance characteristics derived from studies should be measured early in the program. Unrealistic performance requirements can lead to false starts and costly delays.

• *Operating Degradation.* System performance degrades under field conditions. Anticipated degradation must be considered during T&E. When a system must operate at peak performance during development test/operational test (DT/OT) to meet the specified requirements, it will then be likely to perform at a lesser level when operated in the field.

• *Test Personnel.* The test director and/or key members of the test planning group within the project office should have significant T&E experience.

• *Design Reviews.* T&E factors and experience must influence the system design. The application of knowledge derived from past experience can be a major asset in arriving at a sound system design.

• *Surrogate Vehicles.* When high technical risk is present, development should be structured around the use of one or more surrogate vehicles designed to prove the system concept under realistic operational conditions before proceeding with further development.

• *Test Facilities and Scheduling.* Before Milestone I, test range and resource requirements to conduct validation tests and a tentative schedule of test activities should be identified.

26.7.5.2 Demonstration and Validation Phase

• *Vulnerability.* The vulnerability of vehicles should be estimated on the basis of testing.

• *Gun and Ammunition Performance.* Gun and ammunition development should be considered a part of overall tank system development. When a new gun tube, or one which has not been mounted previously on a tank chassis, is being evaluated, all ammunition types (including missiles) planned for use in that system should be test fired under simulated operational conditions.

• *Increased Complexity.* The addition of new capabilities to an existing system or system type will generally increase complexity of the system and, therefore, increase the types and amount of testing required and the time to perform these tests.

• *Component Interfaces.* Prior to assembly in a prototype system, component subsystems should be assembled in a mock-up and verified for physical fit, human factors considerations, interface compatibility and for electrical and mechanical compatibility.

• *Determining Test Conditions.* During validation, test conditions should be determined by the primary objectives of that test rather than by more general considerations of realism.

• *Test Plan Development.* The test plan developed by this point should be in nearly final form and include, as a minimum:

- A description of requirements,
- The facilities needed to make evaluations,
- The schedule of evaluations and facilities,
- The reporting procedure, the objective being to communicate test results in an understandable format to all program echelons,

- The T&E guidelines, and

- A further refinement of the cost estimates which were initiated during the Conceptual Phase.

- *Demonstration Tests.* Demonstration tests should show satisfactory meeting of success criteria which are meaningful in terms of operational usage. It is essential in designing contractually required demonstration tests, upon whose outcome large incentive payments or even program continuation may depend, to specify broader success criteria than simply hit or miss in a single given scenario.

- *Reliability Testing.* Reliability testing should be performed on component and subsystem assemblies before testing of the complete vehicle system. Prior to full system testing, viable component and subsystem tests should be conducted.

- *Human Factors.* In evaluating ground vehicles, human factors should be considered at all stages starting with the design of the prototype.

- *Test Plan Scheduling.* Test plan scheduling should be tied to event milestones rather than to the calendar. In evaluating the adequacy of the scheduling as given by test plans, it is important that milestones be tied to the major events of the weapon system (meeting stated requirements) and not the calendar.

- *Test Failures.* The T&E schedule should be sufficiently flexible to accommodate failures and correction of identified problems.

26.7.5.3 Engineering and Manufacturing Development Phase (old FSD)

- *Planning the IOT&E.* The IOT&E should be cost-effective and provide meaningful results.

- *Pilot and Dry-Run Tests.* A scheduled series of tests should be preceded by a dry run, which verifies that the desired data will be obtained.

- *Comparison Testing.* The test program should include a detailed comparison of the characteristics of a new vehicle system with those of existing systems, alternate vehicle system concepts (if applicable) and those of any system(s) being replaced.

- *Simulation.* Simulation techniques and equipment should be utilized to enhance data collection. Creation of histograms for each test course provides a record of conditions experienced by the vehicle during testing. Use of a chassis dynamometer can produce additional driveling endurance testing with more complete instrumentation coverage.

- *Environmental Testing.* Ground vehicles should be tested in environmental conditions and situations comparable to those in which they will be expected to perform.

- *System Vulnerability.* For combat vehicles, some estimate of vulnerability to battle damage should be made.

- *Design Criteria Verification.* Subsystem design criteria should be compared with actual characteristics.
- *Electromagnetic Testing.* Vehicle testing should include electromagnetic testing.

- *System Strength Testing.* In evaluating ground vehicles, early testing should verify intrinsic strength. This implies operation with maximum anticipated loading, including trailed loads at maximum speeds and over worst-case grades, secondary roads and cross-country conditions for which the vehicle was developed or procured. This test is intended to identify deficient areas of design, not to break the machinery.

• *Component Compatibility.* Component compatibility should be checked through the duration of the test sequence.

• *Human Interface.* Critiques of good and bad features of the vehicle should be made early in the prototype stage while adequate time remains to make any indicated changes.

• *Serviceability Testing.* Ground vehicles should be tested and evaluated to determine the relative ease of serviceability, particularly with high-frequency operations.

• *Experienced User Critique.* Ground vehicle user opinions should be obtained early in the development program.

• *Troubleshooting During Tests.* Provisions should be made to identify subsystem failure causes. Subsystems may exhibit failures during testing. Adequate provisions should be made to permit troubleshooting and identification of defective components and inadequate design.

26.7.5.4 Production and Deployment Phase

• *Performance and Reliability Testing.* The production first-article testing should verify the performance of the vehicle system and determine the degradation, failure modes and failure rates.

• *Lead-the-Fleet Testing.* At least one production prototype or initial production model vehicle should be allocated to intensive testing to accumulate high operating time in a short period.

• *User Evaluation.* User-reported shortcomings should be followed up to determine problem areas requiring correction. Fixes should be evaluated during an FOT&E.

APPENDIX A
ABBREVIATIONS AND ACRONYMS

AAE	Army Acquisition Executive
AAH	Advanced Attack Helicopter
ACM	Advanced Cruise Missile
ADATS	Army Development and Acquisition Threat Simulators
ADM	Acquisition Decision Memorandum
AFEWES	Air Force Electronic Warfare Evaluation Simulator
AFMC	Air Force Materiel Command
AFOTEC	Air Force Operational Test and Evaluation Center
AF/TE	Air Force/Test and Evaluation Office
ALCM	Air Launch Cruise Missile
AMC	Army Materiel Command
AMARC	Army Materiel Acquisition Review Committee
APB	Acquisition Program Baseline
ASAF(A)	Asst. Secretary of Air Force (Acquisition)
ASA (RD&A)	Asst. Sec. of Army (Research, Dev. and Acquisition)
ASD(PAE)	Asst. Sec. of Def. (Program Analysis and Evaluation)
ASN (RD&A)	Asst. Sec. of Navy (Research, Dev. and Acquisition)
ATD	Advanced Technology Demonstration
ATE	Automatic Test Equipment
AWACS	Airborne Warning and Control System
BIS	Board of Inspection and Survey
BIT	Built-In Test
BLRIP	Beyond LRIP
C2	Command and Control
C3	Command, Control and Communication
C3I	Command, Control, Communications, Intelligence
CDR	Critical Design Review
CDRL	Contract Data Requirements List
CDS	Congressional Data Sheets
CE	Concept Exploration
CED	Concept Exploration/Definition Phase
CEP	Circle Error Probability
CLIN	Contract Line Item Number
CNO	Chief of Naval Operations
CNP	Candidate Nomination Proposal
COCI	Critical Operational Issues and Criteria
COEA	Cost and Operational Effectiveness Analysis
COI	Critical Operational Issue
COMOPTEVFOR	Commander, Operational Test and Evaluation Force
CSC	Computer Software Component
CSCI	Computer Software Configuration Item
CSTA	Combined Systems Test Activity
CSU	Computer Software Unit
DA	Developing Agency (Navy)

DAB	Defense Acquisition Board
DAE	Defense Acquisition Executive
DEM/VAL	Demonstration/Validation Phase
DAG	Data Automation Group
DBDD	Data Base Design Document
DCP	Decision Coordination Paper
DDDR&E	Deputy Director Defense Research and Engineering
DEMVAL	Demonstration and Validation Phase
DID	Data Item Description
DLT	Design Limit Test
DMSO	Defense Modeling and Simulation Office
DOD	Department of Defense
DODD	Department of Defense Directive
DODI	Department of Defense Instruction
DOE	Department of Energy
DOT&E	Director, Operational Test and Evaluation
DPESO	DOD Product Engineering Services Office
DPML	Deputy Program Manager, Logistics
DPRB	Defense Planning and Resources Board
DPRO	Defense Plant Representative Office
DSARC	Defense Systems Acquisition Review Council (now DAB)
DSB	Defense Science Board
DT	Development Test
DT&E	Development Test and Evaluation
DTE	Director, Test and Evaluation
DTESG	Defense Test and Evaluation Steering Group
DUSA(OR)	Deputy Undersecretary of Army (Operations Research)
DV	Demonstration/Validation Phase
DVAL	Data Link Vulnerability Analysis
EA	Evolutionary Acquisition
EC	Electronic Combat
ECM	Electronic Countermeasures
ECCM	Electronic Counter-Countermeasures
ECM	Electronic Countermeasures
ECR	Engineering Change Review
EDT	Engineering Design Test
EMD	Engineering and Manufacturing Phase
EMI	Electromagnetic Interference
EMP	Electromagnetic Pulse
EOA	Early Operational Assessment
ERAM	Extended Range Anti-armor Munitions
ESM	Electronic Support Measures
ESS	Environmental Stress Screening
EW	Electronic Warfare
FAADS	Forward Area Air Defense System
FAT	First Article Test
FCA	Functional Configuration Audit

FDT&E	Force Development Tests and Experimentation
FORSCOM	Forces Command
FOT&E	Follow-on Operational Test and Evaluation
FQR	Formal Qualification Review
FSD	Full Scale Development (now EMD)
FWE	Foreign Weapons Evaluation
FYTP	Five Year Test Program
GPMO	Government Program Management Office
HWCI	Hardware Configuration Item
HWIL	Hardware-in-the-Loop
ICBM	Intercontinental Ballistic Missile
IDD	Interface Decision Document
IEP	Independent Evaluation Plan
IFFN	Identification, Friend, Foe, Neutral
IFPP	Information for Proposal Preparation
ILS	Integrated Logistics Support
ILSMT	Integrated Logistic Support Management Team
ILSP	Integrated Logistics Support Plan
IOC	Initial Operating Capability
IOT&E	Initial Operational Test and Evaluation
IPS	Integrated Program Summary
IRA	Industrial Resource Analysis
IRS	Interface Requirements Specification
ITEA	International Test and Evaluation Association
ITP	Integrated Resource Analysis
IV&V	Independent Verification and Validation
JCG(T&E)	Joint Commanders Group (T&E)
JLF	Joint Live Fire
JRD	Joint Requirements Document
JROC	Joint Requirements Oversight Council
JT&E	Joint Test and Evaluation
JTC3A	Joint Tactical C3 Agency
LFT	Live Fire Test
LFT&E	Live Fire Test and Evaluation
LRIP	Low-Rate Initial Production
LSA	Logistics Support Analysis
LSAR	Logistics Support Analysis Report
MAA	Mission Area Analysis
MAJCOM	Major Commands
MCCR	Mission Critical Computer Resources
MCOTEA	Marine Corps Operational Test and Evaluation Agency
MIL-SPEC	Military Specification
MIL-STD	Military Standard
MMOU	Multinational Memorandum of Understanding
MNS	Mission Needs Statement
MOE	Measure of Effectiveness
MOP	Measure of Performance

MOU	Memorandum of Understanding
MPE	Military Preliminary Evaluation
MRTFB	Major Range and Test Facility Base
MS	Milestone
MSTIRC	Multi-Service Test Investment Resource Committee
NAPMA	North Atlantic Program Management Agency
NAVAIR	Naval Air Systems Command
NAVSEA	Naval Sea Systems Command
NBC	Nuclear, Biological, and Chemical
NDCP	Navy Decision Coordinating Paper
NDI	Nondevelopment Item
OJCS	Office of the Joint Chiefs of Staff
NH&S	Nuclear Hardness and Survivability
O&M	Operations and Maintenance
O&S	Operations and Support
OA	Operational Assessment
OMB	Office of Management and Budget
OPEVAL	Operational Evaluation
OPNAV	Operational Navy
OPNAVIST	Operational Navy Instruction
OPSEC	Operations Security
OPTEC	Operational Test and Evaluation Command
OPTEVFOR	Operational Test and Evaluation Force
ORD	Operational Requirement Document
ORMAS/TE	Operational Resource Mgmt Assessment Systems for T&E
OSD	Office of the Secretary of Defense
OT	Operational Test
OT&E	Operational Test and Evaluation
OTA	Operational Test Agency
OTD	Operational Test Director
OTEA	Operational Test and Evaluation Agency
OTO	Operational Test Organization
OTP	Outline Test Plan
P3I	Preplanned Product Improvements
PAT&E	Production Acceptance Test and Evaluation
PCA	Physical Configuration Audit
PCO	Primary Contracting Officer
PDR	Preliminary Design Review
PDSS	Post-Deployment Software Support
PEP	Producibility Engineering Plan
PM	Program Manager
PMO	Program Management Office
PO	Program Office, Purchase Order
POM	Program Objectives Memorandum
PPBS	Planning, Programming, and Budgeting System
PPQT	Preproduction Qualification Tests
PQT	Production Qualification Test

PRAT	Production Reliability Acceptance Test
PRESINSURV	President of the Boards of Inspection and Survey
PRR	Production Readiness Review
QOT&E	Qualification Operational Test and Evaluation
R&E	Research and Engineering
RAM	Reliability, Availability, and Maintainability
R&D	Research and Development
RAS	Requirements Allocations Sheet
RCS	Radar Cross Section
RDT	Reliability Development Testing
RDT&E	Research, Development, Test and Evaluation
RFP	Request for Proposal
RGT	Reliability Growth Test
RM	Resource Manager
RQT	Reliability Qualification Test
SAR	Selected Acquisition Report
SDD	Software Design Document
SDI	Strategic Defense Initiative
SDIO	Strategic Defense Initiative Organization
SDP	Software Development Plan
SDR	System Design Paper
SECARMY	Secretary of the Army
SECDEF	Secretary of the Defense
SECNAV	Secretary of the Navy
SEF	Stability Enhancement Function
SEMP	System Engineering Management Plan
SEMS	System Engineering Management Schedule
SIL	Software Integration Laboratory
SIS	Stall Inhibit System
SON	Statement of Operational Need
SOW	Statement of Work
SPAWAR	Space and Warfare
SPEC	Specification
SPO	System Program Office
SRR	Systems Requirements Review
SRS	Software Requirement Specification
SSD	Segment Design Document
SSR	Software Specification Review
STAR	System Threat Assessment Report
STP	Software Test Plan
SQA	Software Quality Assurance
SW	Software
T&E	Test and Evaluation
TAAF	Test, Analyze and Fix
TADS	Theater Air Defense System
TAFT	Test, Analyze, Fix and Test
TEAM	Test, Evaluation, Analysis, and Modeling

TEC	Test and Evaluation Committee
TECG	T&E Coordinating Group
TECHEVAL	Technical Evaluation (Navy Term)
TECOM	Test and Evaluation Command
TEMA	Test and Evaluation Agency
TEMP	Test and Evaluation Master Plan
TEP	Test and Evaluation Plan
TIWG	Test Integrated Working Group
TLS	Time Line Sheet
TM	Technical Manual
TMC	Test Management Council
TPO	Test Program Outline
TPM	Technical Performance Measurement
TPWG	Test Planning Working Group
TR	Test Report
TRADOC	Training and Doctrine Command
TRMS	TRADOC Resource Management System
TRR	Test Readiness Review
TSARC	Test Schedule and Review Committee
USD(A)	Under Secretary of Defense for Acquisition
WBS	Work Breakdown Structure
WSMR	White Sands Missile Range

APPENDIX B
DOD GLOSSARY OF TEST TERMINOLOGY

ACCEPTANCE TRIALS — Trials and material inspection conducted underway by the trail board for ships constructed in a private shipyard, to determine suitability for acceptance of a ship.

ACCRUED EXPENDITURES — Costs incurred during a given period representing liabilities incurred for goods and services received, other assets acquired and performance accepted, whether or not payment has been made.

ACQUISITION — The process of planning, designing, producing and distributing a weapon system/equipment. Acquisition in this sense includes the conception, validation, full-scale development, production and deployment/operational phases of the weapon systems/equipment project. For weapon systems/equipments not being procured by a project manager, it encompasses the entire process from inception of the requirement through the operational phase.

ACQUISITION CATEGORY (ACAT) — One of four acquisition categories established by the Department of Defense that governs acquisition procedures and responsibilities and assigns respective decision authority levels.

ACQUISITION RISK — The change that some elements of an acquisition program produces an unintended result with adverse effect on system effectiveness, suitability, cost or availability for deployment.

ADVANCED DEVELOPMENT (Budget Category 6.3) — Includes all projects which have moved into the development of hardware for test.

AGENCY COMPONENT — A major organizational subdivision of an agency. For example: the Army, Navy, Air Force and Defense Supply Agency are agency components of the Department of Defense. The Federal Aviation, Urban Mass Transportation and the Federal Highway Administrations are agency components of the Department of Transportation.

ALLOCATION — An authorization by a designated official of a component of the Department of Defense making funds available within a prescribed amount to an operating agency for the purpose of making allotments; i.e., the first subdivision of an apportionment.

ANALYSIS — The qualitative and/or quantified evaluation of information requiring technical knowledge and judgment.

APPORTIONMENT — A determination by the Office of Management and Budget as to the amount of obligations which may be incurred when the nature of the work involved prevents the preparation of definitive requirements, specifications or cost data. Sometimes called Letter of Intent.

AUTHORIZATION — Basic substantive legislation enacted by Congress which sets up a federal program or agency either indefinitely or for a given period of time. Such legislation sometimes sets limits on the amount that can subsequently be appropriated, but does not usually provide budget authority.

AUTOMATIC TEST EQUIPMENT (ATE) — Equipment designed to automatically conduct analysis of functional or static parameters and evaluate the degree of performance degradation and perform fault isolation of unit malfunctions.

BASELINE, APPROVED — The combination of approved program schedule, configuration, performance characteristics, acquisition, strategy, and other business aspects that constitutes the variables reflected in either the appropriate acquisition milestone approval for that acquisition category or as reflected in the latest approved program management proposal action.

BRASSBOARD CONFIGURATION — An experimental device (or group of devices) used to determine feasibility and to develop technical and operational data. Usually, it will be a model sufficiently hardened for use outside of laboratory environments to demonstrate the technical and operational principles of immediate interest. It may resemble the end-item but is not intended for use as the end-item.

BREADBOARD CONFIGURATION — An experimental device (or group of devices) used to determine feasibility and develop technical data. Usually, it will be configured only for laboratory use to demonstrate the technical principles of immediate interest. It may not resemble the end-item and is not intended for use as the projected end-item.

BUDGET — A planned program for a fiscal period in terms of: (a) estimated costs, obligations and expenditures; (b) source of funds for financing, including reimbursements anticipated and other resources to be applied; and (c) explanatory and workload data on the projected programs and activities.

CONCEPT EVALUATION PROGRAM (CEP) — A specifically-funded Army innovative testing program. The CEPs provide commanders and combat developers a quick reaction and simplified process to resolve combat development, doctrinal and training issues. In addition, CEPs solidify combat development requirements and support early milestone decisions. Also, the CEP is used to provide an experimental data base for requirements documents and to expedite the materiel acquisition process; however, CEPs are not to be used as the primary tests to support decision review production decisions. The CEP may be conducted at any time to support the concept evaluation process. Issues satisfied during the conduct of a CEP need not be examined during formal operational test to minimize testing. Data from CEPs may be used as another source for preparing the independent evaluation report.

CONTINUOUS COMPREHENSIVE EVALUATION (C2E) — A continuous process, extending from concept definition through deployment, that evaluates the operational effectiveness and suitability of a system by analysis of all available data.

COMBAT SYSTEM — The equipment, computer programs, people and documentation organic to the accomplishment of the mission of an aircraft, surface ship or submarine; it excludes the structure, material, propulsion, power and auxiliary equipment, transmissions and propulsion, fuels and control systems, and silencing inherent in the construction and operation of aircraft, surface ships and submarines.

CONFIGURATION MANAGEMENT — A discipline applying technical and administrative direction and surveillance to: (1) identify and document the functional and physical characteristics of a configuration item, (2) control changes to those characteristics, and (3) record and report change processing and implementation status.

CONTRACT — An agreement, enforceable by law, between two or more competent parties to do or not do something that is not prohibited by law for a legal consideration.

CONTRACTOR SUPPORT — An arrangement during initial development or production of end-items whereby a contractor furnished required material and maintenance of an end-item or system, pending assumption of supply by the military service.

COST AND OPERATIONAL EFFECTIVENESS ANALYSIS — An analysis of the estimated costs and operational effectiveness of alternative materiel systems to meet a mission need and the associated program for acquiring each alternative.

CRITICAL ISSUES — The aspects of a system's capability, either operational, technical or other, that must be questioned before a system's overall worth can be estimated, and are of primary importance to the decision authority in reaching a decision to allow the system to advance into the next acquisition phase.

DATA SYSTEM — Combinations of personnel efforts, forms, formats, instructions, procedures, data elements and related data codes, communications facilities and automatic data processing equipment that provide an organized and interconnected means, either automated, manual or a mixture of these for recording, collecting, processing and communicating data.

DEFENSE ACQUISITION EXECUTIVE (DAE) — The principal adviser to the Secretary of Defense on all matters pertaining to the Department of Defense Acquisition Systems. The Under Secretary of Defense for Acquisition is the DAE and the Defense Procurement Executive.

CONCEPT DEMONSTRATION APPROVAL — Milestone I decision by which the SECDEF reaffirms the mission need and approves one or more selected alternatives for competitive demonstration and validation.

DEPARTMENT OF DEFENSE ACQUISITION SYSTEM — A single, uniform system whereby all equipment, facilities and services are planned, designed, developed, acquired, maintained and disposed of within the Department of Defense. The system entails establishing policies and practices that govern acquisitions, determining and prioritizing resource requirements, directing and controlling the process, contracting and reporting to the Congress.

DESIGNATED ACQUISITION PROGRAM — Program designated by the Defense Acquisition Executive for Defense Acquisition Board milestone review.

DEVELOPER EVALUATION — The developer's evaluation addresses all aspects of the system to include technical performance, operational effectiveness and operational suitability of cost and schedule.

DEVELOPING AGENCY (DA) — The Systems Command or Chief of Naval Operations designated project manager assigned responsibility for the development, test and evaluation of a weapon system, subsystem or item of equipment.

DEVELOPMENT TEST (DT) — A technical test conducted to provide data on safety, the achievability of critical system technical characteristics, refinement and ruggedization of hardware configurations and determination of technical risks. This testing is performed on components, subsystems, materiel improvement, nondevelopment items, hardware-software integration and related software. The development test includes the testing of compatibility and interoperability with existing or planned equipment and systems and the system effects caused by natural and induced environmental conditions during the development phases of the materiel acquisition process.

EARLY OPERATIONAL ASSESSMENT — An operational assessment conducted prior to, or in support of, Milestone II.

EFFECTIVENESS — The performance or output received from an approach or a program. Ideally, it is a quantitative measure which can be used to evaluate the level of performance in relation to some standard, set of criteria or end objective.

ENGINEERING CHANGE — An alteration in the physical or functional characteristics of a system or item delivered, to be delivered or under development, after establishment of such characteristics.

ENGINEERING CHANGE PROPOSAL (ECP) — Proposal to change design or engineering features of materiel under development or production. Includes proposed engineering change and documentation by which the change is described and suggested.

ENGINEERING DEVELOPMENT — The RDTE funding category that includes development programs being engineered for Service use but not yet approved for procurement or operation. Budget Category 6.4 includes those projects in full-scale development of Service use; but they have not yet received approval for production or had production funds included in the DOD budget submission for the budget or subsequent fiscal year.

EVALUATION CRITERIA — Standards by which achievement of required technical and operational effectiveness/suitability characteristics or resolution of technical or operational issues may be evaluated. Evaluation criteria should include quantitative thresholds for the initial operating capability (IOC) system. If parameter maturity grows beyond IOC, intermediate evaluation criteria, appropriately time-lined, must also be provided.

FUTURE-YEAR DEFENSE PROGRAM (FYDP) — The official document which summarizes the Secretary of Defense approved plans and programs for the Department of Defense. It is published at least annually.

FOLLOW-ON OPERATIONAL TEST AND EVALUATION (FOT&E) — The test and evaluation that is necessary during and after the production period to refine estimates made during operational test and evaluation, to evaluate changes and to reevaluate the system to ensure it continues to meet operational needs and retains its effectiveness in a new environment or against a new threat.

FOLLOW-ON PRODUCTION TEST — A technical test conducted subsequent to a full production decision on initial production and mass production models to determine production conformance for quality assurance purposes. Program funding category — Procurement.

INDEPENDENT EVALUATION REPORT — A report that provides an assessment of item or system operational effectiveness and operational suitability vs. critical issues as well as the adequacy of testing to that point in the development of item or system.

INDEPENDENT OPERATIONAL TEST AGENCY — The Army Operational Test and Evaluation Agency, the Navy Operational Test and Evaluation Force, the Air Force Operational Test and Evaluation Center, and the Marine Corps Operational Test and Evaluation Agency.

INITIAL OPERATIONAL TEST AND EVALUATION (IOT&E) — All operational test and evaluation conducted on production or production-representative articles to support the decision to proceed beyond low-rate initial production. It is conducted to provide a valid estimate of expected system operational effectiveness and operational suitability.

IN-PROCESS REVIEW — Review of a project or program at critical points to evaluate status and make recommendations to the decision authority.

INTEGRATED LOGISTIC SUPPORT (ILS) — A disciplined, unified and iterative approach to the management and technical activities necessary to: (a) integrate support considerations into system and equipment design; (b) develop support requirements that are related consistently to readiness objectives, design and each other; (c) acquire the required support; and (d) provide the required support during the operational phase at minimum cost.

INTEROPERABILITY — The ability of systems, units or forces to provide services to, and accept from, other systems, units or forces, and to use the services so exchanged to enable them to operate together effectively.

ISSUES — Any aspect of the system's capability, either operational, technical or other, that must be questioned before the system's overall military utility can be known. Operational issues are issues that must be evaluated considering the soldier and the machine as an entity to estimate the operational effectiveness and operational suitability of the system in its complete user environment.

JOINT DEVELOPMENT TESTS (JDTs) — The JDTs provide information on intra-Service systems or equipment requirements, performance or interoperability; technical concepts, requirements or improvements; and the improvement or development of testing methodologies or resources.

JOINT OPERATIONAL TESTS (JOTs) — The JOTs use actual fielded equipment, simulators or surrogate equipment in an exercise or operational environment to obtain data pertinent to inter-Service operational doctrine, tactics and procedures.

LETHALITY — The probability that weapon effects will destroy the target or render it neutral.

LIFE-CYCLE COST — The total cost to the government for the development, acquisition, operation and logistic support of a system or set of forces over a defined life span.

LOGISTICS SUPPORTABILITY — The degree of which the planned logistics support (including test equipment, measurement and diagnostic equipment, spare and repair parts, technical data, support facilities, transportation requirements, training and manpower) allow meeting system availability and wartime usage requirements.

LONG LEAD ITEMS — The components of a system or piece of equipment that take the longest time to procure and, therefore, may require an early commitment of funds in order to meet acquisition program schedules.

LOW RATE INITIAL PRODUCTION (LRIP) — Any manufacture of a system in limited quantity to be used in initial operational test and evaluation for verifying production engineering and design maturity and to establish a production base.

MAINTAINABILITY — A characteristic of design and installation that is expressed as the probability of an item being retained in, or restored to, a specified condition within a given period of time when the maintenance is performed in accordance with prescribed procedures and resources.

MAJOR DEFENSE ACQUISITION PROGRAM — As specified in United States Code 10, sections 136a and 139a (reference 1) and DOD Directive 5000.1:

a. A DOD acquisition program that is not a highly sensitive/classified program (as determined by the Secretary of Defense) and:

(1) That is designated by the Secretary of Defense as a major defense acquisition program; or

(2) That is estimated by the Secretary of Defense to require an eventual total expenditure for research, development, test and evaluation of more than 200 million dollars (based on fiscal year 1980 constant dollars) or an eventual total expenditure for procurement of more than 1 billion dollars (based on fiscal year 1980 constant dollars).

b. A DOD acquisition program that is designated jointly by the DOT&E and DTE, as a major defense acquisition program for the purpose of carrying out the responsibilities, functions, and authorities of this Manual. Such designation for the purpose of Test and Evaluation oversight does not imply any other related review requirements.

MAJOR RANGE AND TEST FACILITY BASE (MRTFB) — The complex of major DOD ranges and test facilities.

MILESTONE — A major management decision point in the overall acquisition process of a major Department of Defense (DOD) system requiring Office of the Secretary of Defense and/or DOD Component Program review. Milestones include Joint Resource Management Board and DOD Component Equivalent Program reviews.

MILITARY REQUIREMENT — An established need justifying the timely allocation of resources to achieve a capability to accomplish approved military objectives, missions or tasks. Requirements are normally documented in a Mission Needs Statement or Operational Requirement Document.

MISSION AREA ANALYSIS (MAA) — Continuous analysis of assigned mission responsibilities in the several mission areas to identify deficiencies in the current and projected capabilities to meet essential mission needs and to identify opportunities for the enhancement of capability through more effective systems and less-costly methods.

MISSION NEED STATEMENT (MNS) — Submitted prior to Program Objectives Memorandum submission. Approval by SECDEF is Milestone 0. Documents major mission deficiencies (or opportunities for improvement) in a Service's ability to meet mission requirements when such deficiencies can be corrected by: (1) using an existing U.S. system or allied military or commercial system, (2) a major modification to an existing system, or (3) a new major acquisition. A joint MNS is prepared to document major deficiencies in two or more DOD components. The Office of the Secretary of Defense or Office of the Joint Chiefs of Staff may also prepare the MNS.

MISSION RELIABILITY — The probability that the system will perform mission essential functions for a period of time under the conditions stated in the mission profile.

MODEL — A model is a representation of an actual or conceptual system involving mathematics, logical expressions or computer simulations that can be used to predict how the system might perform or survive under various conditions or in a range of hostile environments.

MULTI-SERVICE OPERATIONAL TEST — A form of test when one or more of the Services provide support Service test or vice versa or tests that involve agreements between a Service and one or more of the other Services to evaluate a system or concept that requires testing in a multi-Service environment.

NONDEVELOPMENT ITEM (NDI) — Already developed and available hardware and/or software capable of fulfilling Service requirements, thereby minimizing or eliminating the need for costly, time-consuming, government-sponsored R&D programs. An NDI is usually off-the-shelf or a commercial-type product, but may also be equipment already developed by or for the military services or foreign military forces.

NUCLEAR HARDNESS — A quantitative description of the physical attributes of the system or component that will allow nuclear survivability in a given weapon environment. Hardness is measured by physical quantities such as overpressure, peak velocities, energy absorbed, electrical stress, etc. Hardness is achieved through design specifications and often verified by one or more test and analysis techniques. Hardness is only one of several means of attaining system-wide nuclear survivability.

OPERATIONAL ASSESSMENT — An evaluation of operational effectiveness and operational suitability made by an independent operational test activity, with user support as required, on other than production systems. The focus of an operational assessment is on significant trends noted in development efforts, programmatic voids, areas of risk, adequacy of requirements and the ability of the program to support adequate operational testing. Operational assessments may be made at any time using technology demonstrators, prototypes, mock-ups, engineering development models or simulations but will not substitute for the independent operational test and evaluation necessary to support full production decisions.

OPERATIONAL AVAILABILITY — An index of weapon system materiel readiness, including system software where applicable, in a mission environment. It is a measure of probability of an item's being in a condition, generally referred to as "up," so it can perform its intended function when called upon within acceptable limits of degradation.

OPERATIONAL EFFECTIVENESS — The overall degree of mission accomplishment of a system when used by representative personnel in the environment planned or expected (e.g. natural, electronic, threat, etc.) for operational employment of the system considering organization, doctrine, tactics, survivability, vulnerability and threat (including countermeasures; initial nuclear weapons effects; and nuclear, biological and chemical contamination threats).

OPERATIONAL EVALUATION — Addresses the effectiveness and suitability of the weapons, equipment or munitions for use in combat by typical military users and the system operational issues and criteria; provides information to estimate organizational structure, personnel requirements, doctrine, training and tactics; identifies any operational deficiencies and the need for any modifications; and assesses MANPRINT (safety, health hazards, human factors, manpower and personnel) aspects of the system in a realistic operational environment.

OPERATIONAL SUITABILITY — The degree to which a system can be placed satisfactorily in field use, with consideration being given to availability, compatibility, transportability, interoperability, reliability, wartime usage rates, maintainability, safety, human factors, manpower supportability, logistic supportability and training requirements.

OPERATIONAL TEST — Testing of materiel systems that is accomplished with representative user operators, crews, support personnel or units in as realistic an operational environment as possible to provide the evaluator data to estimate:

a. The military operational effectiveness and operational suitability (including compatibility, interoperability, reliability, availability, maintainability, supportability, operational soldier/hardware/software interface and training requirements) of new systems.

b. The system's desirability, from the use viewpoint, considering systems already available and the operational benefits and/or burdens associated with the new system.

c. The need for modifying the system.

d. The adequacy of doctrine, organization, operating techniques, tactics and training for employment of the system; the adequacy of maintenance and supply support for the system; and, when appropriate, its performance in a countermeasure environment.

OPERATIONAL TEST AND EVALUATION (OT&E) — The field test, under realistic combat conditions, of any item (or key component of), weapons, equipment or munitions for the purpose of determining the effectiveness and suitability of the weapons, equipment or munitions for use in combat by typical military users; and the evaluation of the results of such test.

OPERATIONAL TEST CRITERIA — Expressions of the operational level of performance required of the military system to demonstrate operational effectiveness for given functions during each operational test. The expression consists of the function addressed, the basis for comparison, the performance required and the confidence level.

OPERATIONAL TEST READINESS REVIEW (OTRR) — A review to identify problems that may impact the conduct of an OT&E. The OTRRs are conducted to determine changes required in planning, resources or testing necessary to proceed with the OT&E. Participants include the operational tester (chair), evaluator, material developer, user representative, logisticians, HQDA staff elements and others, as necessary.

PILOT PRODUCTION — The controlled manufacture of limited numbers of an item for Service test and evaluation purposes using manufacturing drawings and specifications, which have been developed for quantity production and with tooling that is representative of that to be used in unlimited production.

POSTPRODUCTION TESTING — Testing conducted to ensure that materiel that is reworked, repaired, renovated, rebuilt or overhauled after initial issue and deployment conforms to specified quality, reliability, safety and operational performance standards. Included in postproduction tests are surveillance tests, stockpile reliability and reconditioning tests.

PREPLANNED PRODUCT IMPROVEMENT (P^3I) — Planned future evolutionary improvement of developmental systems for which design considerations are effected during development to enhance future application of projected technology. Includes ments planned for ongoing systems that go beyond the current performance envelope to achieve a needed operational capability.

PREPRODUCTION PROTOTYPE — Items in final form employing standard parts that are representative of items to be produced on a production line with production tooling.

PRODUCT IMPROVEMENT PLAN (PIP) — Effort to incorporate configuration changes involving engineering and testing effort, on end-items and depot repairable components or changes on other-than-developmental items to increase system or combat effectiveness or extend the useful military life.

PRODUCTION QUALIFICATION TEST (PQT) — A technical test conducted post-Milestone III to ensure the effectiveness of the manufacturing process, equipment and procedures. This testing also provides data for the independent evaluation required for materiel release so the evaluator can address the adequacy of the materiel with respect to the stated requirements. These tests are conducted on a number of samples taken at random from the first production lot and are repeated if the process or design is changed significantly and when a second or alternative source is brought on line. Program funding category — Procurement.

PROGRAM MANAGER — Individual chartered by the Service Secretary reporting to the material developer or to the commander of a subordinate organization as designated by the material developer. Assigned responsibility and delegated full-line authority of the material developer for centralized management of a specified acquisition or materiel readiness program. May be superimposed over one or more product managers.

QUALIFICATIONS TESTING — Testing that verifies the design and manufacturing process and provides a baseline for subsequent acceptance tests. The completion of production qualification test and evaluation before Milestone III decisions is essential and will be a critical factor in assessing the system's readiness for production. Production qualification test and evaluation shall be conducted on production-representative items.

QUALITY ASSURANCE — A planned and systematic pattern of all actions necessary to provide adequate confidence that materiel conforms to established technical requirements and achieves satisfactory performance in service.

REALISTIC TEST ENVIRONMENT — The conditions under which a system is expected to be operated and maintained, including the natural weather and climatic conditions, terrain effects, battlefield disturbances and enemy threat conditions.

RELIABILITY — The probability that an item will perform its intended function for a specified interval under stated conditions.

REQUIRED OPERATIONAL CHARACTERISTICS — Qualitative and quantitative system parameters approved by the user that are primary indicators of a system's capability to accomplish its mission (operational effectiveness) and to be supported (operational suitability).

REQUIRED TECHNICAL CHARACTERISTICS — Quantitative system parameters approved by the Department of Defense Component that are selected as primary indicators of technical achievement of engineering thresholds. These might not be direct measures of, but should always relate to, a system's capability to perform its required mission function and to be supported

RESEARCH — Includes all effort of scientific study and experimentation directed toward increasing knowledge and understanding in the physical, engineering, environmental and life sciences fields related to long-term national security needs. It provides fundamental knowledge required for the solution of military problems. It forms a part of the base for subsequent exploratory and advanced developments in defense-related technologies, and new and improved military functional capabilities in areas such as communications, detection, tracking, surveillance, propulsion, mobility, guidance and control, navigation, energy conversion, materials and structures and personnel support. (Budget Category 6.1)

RISK — An expression of possible loss in terms of hazard severity and hazard probability.

RISK ASSESSMENT — An evaluation of a risk in terms of mission loss should a hazard result in an accident.

SAFETY — Freedom from those conditions that can cause death, injury, occupational illness or cause damage to, or loss of, equipment or property.

SAFETY/HEALTH VERIFICATION — The development of data used to evaluate the safety and health features of a system to determine its acceptability. This is done primarily during developmental test and user or operational test and evaluation and supplemented by analysis and independent evaluations.

SAFETY RELEASE — A formal document issued to a user test organization before any hands-on use or maintenance by personnel. The safety release indicates the system is safe for use and maintenance by typical user personnel and describes the system safety analyses. Operational limits and precautions are included. The test agency uses the data to integrate safety into test controls and procedures and to determine if the test objectives can be met within these limits.

SELECTED ACQUISITION REPORT — A standard, comprehensive, summary status report provided to the Congress on Department of Defense (DOD) acquisition programs for management within DOD.

SIMULATION — A simulation is a method for implementing a model. It is the process of conducting experiments with a model for the purpose of understanding the behavior of the system modeled under selected conditions or limits imposed by developmental or operational criteria. Simulation may include the use of analog or digital devices, laboratory models or "test-bed" sites. Simulations are usually programmed for solution on a computer; however, in the broadest sense, military exercises and war games are also simulations.

SPECIFICATION — A specific quantitative, contractually binding, required operational or technical characteristic.

SUBTEST — An element of a test program. A subset is a test conducted for a specific purpose (e.g., rain, dust, transportability, missile firing, fording).

SUITABILITY — A subjective determination by a decision authority that a materiel system does or does not meet minimum standards prerequisite to satisfy field Service use. The judgment may be based on the presence or absence of uncorrectable materiel deficiencies and/or the number and assessed importance of correctable and uncorrectable shortcomings. It also includes judgments on nonmateriel issues.

SURVIVABILITY — The capability of a system to avoid or withstand man-made hostile environments without suffering an abortive impairment of its ability to accomplish its designated mission.

SUSCEPTIBILITY — The degree to which a device, equipment or weapon system is open to effective attack due to one or more inherent weaknesses. (Susceptibility is a function of operational tactics, countermeasures, probability of enemy fielding a threat, etc.)

SYSTEM — A composite, at any level of complexity, of personnel, procedures, materials, tools, equipment, facilities and software. The elements of this composite entity are used together in the intended operational or support environment to perform a given task or achieve a specific production, support or mission requirement.

SYSTEM ENGINEERING, DEFENSE — That portion of the acquisition process dealing with the transformation of an operational need into an optimal set of system performance parameters and a preferred system configuration. It includes engineering/technical management, definition of system and program, design engineering, support engineering, the integration of the engineering specialties, and other such factors that affect the development, production, deployment, operation and disposal of the system.

SYSTEM ENGINEERING PROCESS — A logical sequence of activities and decisions transforming an operational need into a description of system performance parameters and a preferred system configuration.

TECHNICAL EVALUATION — Addresses the system's technical issues and criteria and the acquisition and fielding of an effective, supportable and safe system. It assists in the engineering design and development and verifies attainment of technical performance specifications, objectives, producibility, adequacy of the Technical Data Package, and supportability, determining safety, health hazards, human factors, and MANPRINT aspects. Technical evaluation encompasses the use of models, simulations and test beds as well as a prototypes or full-scale development models of the system.

TECHNICAL FEASIBILITY TEST — A technical test conducted post-Milestone 0 and pre-Milestone I or Milestone I/II (under the Army Streamlined Acquisition Process) to assist in determining safety and establishing system performance specifications and feasibility.

TECHNICAL TESTER — The command or agency that plans, conducts and reports the results of Army technical testing. Associated contractors may perform development testing on behalf of the command or agency.

TECHNICAL TESTS — A generic Army term for testing that gathers technical data during development testing, technical feasibility testing, qualification testing, joint development testing and contractor/foreign testing. Soldier operator-maintainer test and evaluation personnel are used during technical testing when appropriate.

TEST AND EVALUATION MASTER PLAN — An overall test and evaluation strategy plan that is prepared as early as possible in the acquisition process and is designed to identify and integrate objectives, responsibilities, resources and schedule for all test and evaluation to be accomplished prior to key decision milestones.

TEST BEDS — A system representation consisting partially of actual hardware and partially of computer models or prototype hardware.

TEST CRITERIA — Standards by which test results and outcome are judged.

TEST DESIGN PLAN — A statement of the conditions under which the test is to be conducted, the data required from the test, and the data handling required to relate the data results to the test conditions.

TEST INSTRUMENTATION — Test instrumentation is scientific, automated data processing equipment or technical equipment used to measure, sense, record, transmit, process or display data during tests, evaluations or examination of materiel, training concepts, or tactical doctrine. Audiovisual is included as instrumentation when used to support Army testing.

TEST RESOURCES — A collective term that encompasses all elements necessary to plan, conduct and collect/analyze data from a test event or program. Elements include test funding and support manpower (including TDY costs), test assets (or units under test), test asset support equipment, technical data, simulation models, test beds, threat simulators, surrogates and replicas, special instrumentation peculiar to a given test asset or test event, targets, tracking and data acquisition, instrumentation, equipment for data reduction, communications, meteorology, utilities, photography, calibration, security, recovery, maintenance and repair, frequency management and control, and base/facility support services.

THREAT — The sum of the potential strength, capabilities and intentions of an enemy that can limit or negate mission accomplishment or reduce force, system or equipment effectiveness.

THRESHOLDS — The minimum level of a performance parameter the system must meet (e.g., minimum flight altitude threshold of 30 feet above ground level for a missile).

TRANSPORTABILITY — The inherent capability of materiel to be moved by towing, self-propulsion or carrier via railways, highways, waterways, pipelines, ocean and airways.

USER REPRESENTATIVE — The combat developer designated to represent the user during the materiel acquisition process. The command or agency fulfilling this role represents the "mission-oriented" user and the "logistics-oriented" user by concerning itself with the operational and logistic support aspects of materiel system.

VULNERABILITY — The characteristics of a system that cause it to suffer a definite degradation (loss or reduction of capability to perform the designated mission) as a result of having been subjected to a certain (defined) level of effects in an unnatural (man-made), hostile environment.

WORK BREAKDOWN STRUCTURE (WBS) — A product-oriented family-tree division of hardware, software, services and other work tasks that organizes, defines and graphically displays the product to be produced as well as the work to be accomplished to achieve the specified product.

APPENDIX C
TEST-RELATED DATA ITEM DESCRIPTIONS

extracted from DOD 5010.12-L,
Acquisition Management System and
Data Requirement Control List (AMSDL)

ACCEPTANCE TEST PLAN	DI-QCIC-80154, -80553
AIRBORNE SOUND MEASUREMENTS TEST REPORT	DI-HFAC-80272
AIRFRAME RIGIDITY TEST REPORT	DI-T-30734
AMMUNITION TEST EXPENDITURE REPORT	DI-MISC-80060
ARMOR MATERIAL TEST REPORTS	DI-MISC-80073
BALLISTIC ACCEPTANCE TEST REPORT	DI-MISC-80246
C.P. PROPELLER TEST AGENDA	UDI-T-23737
COORDINATED TEST PLAN	DI-MGMT-80937
CORROSION TESTING REPORTS	DI-MFFP-80108
DAMAGE TOLERANCE TEST RESULTS REPORTS	DI-T-30725
DEMONSTRATION TEST:PLAN	DI-QCIC-80775
REPORT	DI-QCIC-80774
DIRECTED ENERGY SURVIVABILITY TEST PLAN	DI-R-1786
DURABILITY TEST RESULTS REPORT	DI-T-30726
ELECTROMAGNETIC COMPATIBILITY TEST PLAN	DI-T-3704B
ELECTROMAGNETIC INTERFERENCE TEST:PLAN	DI-EMCS-80201
REPORT	DI-EMCS-80200
ELECTROSTATIC DISCHARGE SENSITIVITY TEST REPORT	DI-RELI-80670
EMISSION CONTROL (EMCON) TEST REPORT	DI-R-2059
ENDURANCE TEST (EMCS) FAILURE REPORTS	DI-ATTS-80366

ENGINEER DESIGN TEST PLAN		DI-MGMT-80688
ENVIRONMENTAL DESIGN TEST PLAN		DI-ENVR-80861
ENVIRONMENTAL TEST REPORT		DI-ENVR-80863
EQUIPMENT TEST PLAN (NONSYSTEM)		DI-T-3709A
FACTORY TEST:	PLAN	DI-QCIC-80153
	EMCS PLAN	DI-ATTS-80360
	EMCS PROCEDURES	DI-ATTS-80361
	EMCS REPORTS	DI-ATTS-80362
FIRST ARTICLE QUALIFICATION TEST PLAN		DI-T-5315A
FLIGHT FLUTTER TEST REPORT		DI-T-30733
FLUTTER MODEL TEST REPORT		DI-T-30732
HARDWARE DIAGNOSTIC TEST SYSTEM DEVELOPMENT PLAN		DI-ATTS-80005
HIGH-IMPACT SHOCK TEST PROCEDURES		DI-ENVR-80709
HULL TEST RESULTS (BOATS) REPORT		UDI-T-23718
HUMAN ENGINEERING TEST:	PLAN	DI-HFAC-80743
	REPORT	DI-HFAC-80744
INSPECTION AND TEST PLAN		DI-QCIC-81110
INSTALLATION TEST:	PLAN	DI-QCIC-80155
	PROCEDURES	DI-QCIC-80511
	REPORT	DI-QCIC-80140, -80512
INTEGRATED CIRCUIT TEST DOCUMENTATION		DI-ATTS-80888
MAINTAINABILITY/TESTABILITY DEMONSTRATION TEST:	PLAN	DI-MNTY-80831
	REPORT	DI-MNTY-80832
MAINTENANCE TRAINING EQUIPMENT TEST OUTLINES		DI-H-6129A
MASTER TEST PLAN/PROGRAM TEST PLAN		DI-T-30714
NBC CONTAMINATION SURVIVABILITY TEST:	PLAN	DI-R-1779

NUCLEAR SURVIVABILITY TEST:	PLAN	DI-NUOR-80928
	REPORT	DI-NUOR-80929
PACKAGING TEST:	PLAN	DI-PACK-80456
	REPORT	DI-PACK-80457
PART, COMPONENT OR SUBSYSTEM TEST PLAN(S)		DI-MISC-80759
PARTS (NONSTANDARD) TEST DATA REPORT		DI-MISC-81058
PARTS QUALIFICATION TEST PLAN		DI-T-5477A
PERFORMANCE ORIENTED PACKAGING TEST REPORT		DI-PACK-81059
PRODUCTION TEST:	PLAN	DI-MNTY-80173
	REPORT	DI-NDTI-80492
QUALITY CONFORMANCE TEST PROCEDURES		DI-RELI-80322
RADAR SPECTRUM MANAGEMENT (RSM) TEST PLAN		DI-MISC-81113
RANDOMIZER TEST REPORT		DI-NDTI-80884
RELIABILITY TEST:	PLAN	DI-RELI-80250
	PROCEDURES	DI-RELI-80251
	REPORTS	DI-RELI-80252
RESEARCH AND DEVELOPMENT TEST AND ACCEPTANCE PLAN		DI-T-30744
ROUGH HANDLING TEST REPORT		DI-T-5144C
SHIP ACCEPTANCE TEST (SAT):	SCHEDULE	DI-T-23959B
	REPORT	DI-T-23190A
SHIPBOARD INDUSTRIAL TEST PROCEDURES		DI-QCIC-80206
SHOCK TEST:	EXTENSION REQUEST	DI-ENVR-80706
	REPORT	DI-ENVR-80708
SOFTWARE GENERAL UNIT TEST PLAN		DI-MCCR-80307
SOFTWARE TEST:	DESCRIPTION	DI-MCCR-80015A
	PLAN	DI-MCCR-80014A
	PROCEDURES	DI-MCCR-80310
	REPORT	DI-MCCR-80017A, -80311

SOFTWARE SYSTEM:	DEVEL TEST AND EVAL PLAN	DI-MCCR-80309
	INTEGRATION AND TEST PLAN	DI-MCCR-80308
SOUND TEST FAILURE NOTIF AND RECOMM		DI-HFAC-80271
SPECIAL TEST EQUIPMENT PLAN		DI-T-30702
SPECTRUM SIGNATURE TEST PLAN		DI-R-2068
STATIC TEST:	PLAN	DI-T-21463A
	REPORTS	DI-T-21464A
STRUCTUREBORNE VIBRATION ACCEL MEASUREMENT TEST		DI-HFAC-80274
SUPERIMPOSED LOAD TEST REPORT		DI-T-5463A
TEMPEST TEST:	REQUEST	DI-EMCS-80218
	PLAN	DI-T-1912A
TEST CHANGE PROPOSAL		DI-T-26391B
TEST ELEMENTS LIST		DI-QCIC-80204
TEST FACILITY REQUIREMENTS DOCUMENT (TFRD)		DI-FACR-80810
TEST PACKAGE		DI-ILSS-81085
TEST:	PLAN	DI-NDTI-80566
	PLANS/PROCEDURES	DI-NDTI-80808
	PROCEDURE	DI-NDTI-80603
	PROCEDURES	UDI-T-23732B
TEST PLAN DOCUMENTATION FOR AIS		DI-IPSC-80697
TEST PROGRAM:	DOCUMENTATION (TPD)	DI-ATTS-80284
	INTEGRATION LOGBOOK	DI-ATTS-80281
TPS AND OTPS ACCEPTANCE TEST:	PROCEDURES (ATPS)	DI-ATTS-80282A
	REPORT (ATR)	DI-ATTS-80283A
TEST REPORTS		DI-NDTI-80809A, DI-MISC-80653
TEST REQUIREMENTS DOCUMENT		DI-T-2181, DI-ATTS-80002, -80041

TEST SCHEDULING REPORT		DI-MISC-80761
TESTABILITY:	PROGRAM PLAN	DI-T-7198
	ANALYSIS REPORT	DI-T-7199
TRAINER TEST PROCEDURES AND RESULTS REPORT		DI-T-25594C
VIBRATION AND NOISE TEST REPORTS		DI-T-30735
VIBRATION TESTING:	EXTENSION	UDI-T-23752
	REPORT	UDI-T-23762
WELDING PROCEDURE QUALIFICATION TEST REPORT		DI-MISC-80876

STANDARDIZATION AREAS		LEAD SERVICE
ATTS	Automatic Test Technology Standards	10
CMAN	Configuration Management	SD
E	Engineering and Configuration Documentation	*
EMCS	Electromagnetic Compatibility	EC
ENVR	Environmental Requirements and Related Test Meth	TE
FACR	Facility Construction Design Requirements	YD
GDRQ	General Design Requirements	SD
H	Human Factors	*
HFAC	Human Factors	MI
ILSS	Integrated Logistics Support Standards	WS
IPSC	Information Processing Standards for Computers	02
MCCR	Mission Critical Computer Resources	10
MFFP	Metal Finishes and Finishing Processes and Proc	MR
MGMT	Management	var
MISC	Miscellaneous	SD
MNTY	Maintainability	17

NDTI	Nondestructive Testing and Inspection	MR
NUOR	Nuclear Ordnance	DS
PACK	Packing, Packaging, Preservation and Transport	SM
QCIC	Quality Control/Assurance and Inspection	AR
R	Related Design Requirements	*
RELI	Reliability	17
S	System/Subsystem Analysis	*
STANDARDIZATION AREAS		LEAD SERVICE
SAFT	Safety	10
TMSS	Technical Manual Specifications and Standards	TM
T	Test	*
V	Test	*

*Prior to 1 Jul 1985; being attritted out.
UDI: Indicates DID unique to originator; being attritted out.

Lead Services

02	USAF	ACS Information Systems
10	USAF	AFMC Command Standardization Office
17	USAF	AFMC Rome Air Development Center
AR	USA	AMC AMCCOM ARDEC
DS	DNA	Defense Nuclear Agency
EC	USN	SPAWAR
MI	USA	AMC MICOM
MR	USA	AMC ARL Matl Tech Lab

SD	OSD	Standardization and Data Management
SM	USA	AMC Packaging, Storage and Containerization Center
TE	USA	AMC TECOM
TM	USA	AMC Matl Readiness Spt Activity
WS	OSD	Wpn Spt Improvement and Analysis Office
YD	USN	Nav Fac Engr Cmd

APPENDIX D
BIBLIOGRAPHY

DEPARTMENT OF DEFENSE DIRECTIVES: Many of the references used in the first edition were cancelled and their intent incorporated in the 5000 series documents issued February 23, 1991, and implementing Service documents.

1. DODD 3200.11, Major Range and Test Facility Base.

2. DODD 3200.11-D, Major Range and Test Facility Base Summary of Capabilities.

3. (cancelled) DODD 4245.6, Defense Production Management.

4. (cancelled) DODD 4245.7, Transition from Development to Production.

5. DODD 5000.1, Defense Acquisition.

6. (cancelled) DODD 5000.3, Test and Evaluation.

7. (cancelled) DODD 5000.37, Acquisition and Distribution of Commercial Products.

8. (cancelled) DODD 5000.38, Production Readiness Reviews.

9. (cancelled) DODD 5000.39, Acquisition and Management of Integrated Logistic Support for Systems and Equipment.

10. DODD 5010.8, Value Engineering Program.

11. DODD 5105.31, Defense Nuclear Agency.

12. DODD 5134.1, Under Secretary of Defense (Acquisition).

13. DODD 5141.2, Director of Operational Test and Evaluation.

14. DODD 5160.5, Responsibilities for Research, Development and Acquisition of Chemical Weapons and Chemical and Biological Defense.

DEPARTMENT OF DEFENSE INSTRUCTIONS

15. (cancelled) DODI 4245.4, Acquisition of Nuclear-Survivable Systems.

16. DODI 5000.2, Defense Acquisition Management Policies and Procedures.

17. (cancelled) DODI 5000.38, Production Readiness Reviews.

18. DODI 5030.55, Joint AEC-DOD Nuclear Weapons Development Procedures.

19. (cancelled) DODI 7000.10, Contractor Cost Performance.

DEPARTMENT OF DEFENSE GUIDES/MANUALS

20. DOD 4245.7-M, Transition From Development to Production.

21.A. (cancelled) DOD 5000.3-M-1, Test and Evaluation Master Plan (TEMP) Guidelines.

21.B. DOD 5000.3-M-2, Foreign Weapons Evaluation Program and NATO Comparative Test Program Procedure Manual.

22. (cancelled) DOD 5000.3-M-3,Software Test and Evaluation Manual.

DEPARTMENT OF DEFENSE STANDARDS

23. DOD-STD-2167, Defense System Software Development.

24. DOD-STD-2168, Defense System Software Quality Program.

MILITARY STANDARDS

25. MIL-STD-480, Configuration Control — Engineering Changes, Deviations, and Waivers.

26. MIL-STD-483 (USAF), Configuration Management Practices for Systems, Equipment, Munitions, and Computer Software.

27. MIL-STD-490, Specification Practices.

28. MIL-STD-499A (USAF), Engineering Management Practices.

29. MIL-STD-1388-1A, Logistic Support Analysis.

30. MIL-STD-1512A, Technical Reviews and Audits for Systems, Equipment, and Computer Programs.

31. MIL-STD-1528, Manufacturing Master Plan.

MILITARY HANDBOOKS

32. Joint Test Director's Handbook (Draft).

OTHER MILITARY DOCUMENTS

33. Test and Evaluation in the United States Department of Defense, Office of the Under Secretary of Defense for Research and Engineering (Director Defense Test and Evaluation), August 1983.

34. Statement by Mr. James F. O'Bryon, Assistant Deputy Under Secretary of Defense (Live Fire Test) Before the Acquisition Policy Panel of the House Armed Services Committee, September 10, 1987.

35. Memorandum of Agreement of Multi-Service OT&E and Joint T&E, with changes.

36. Joint Test and Evaluation Procedures Manual, (Draft), Office of the Deputy Under Secretary for Research and Engineering (Test and Evaluation), October 1986.

37. Live Fire Test and Evaluation Planning Guide, Office of the Deputy Director, Test and Evaluation (Live Fire Test), June 1989.

38. Joint Logistics Commanders Guidance for the Use of an Evolutionary Acquisition (EA) Strategy in Acquiring Command and Control (C2) Systems, Defense Systems Management College, March 1987.

39. Concept and Approach for a Joint Test and Evaluation of U.S. Chemical Warfare Retaliatory Capabilities, JCHEM Joint Test Force, AD #B088340, February 1984.

40. FY 88 Report of the Secretary of Defense to the Congress.

41. Report of Task Force on Test and Evaluation, Defense Science Board, April 1, 1974.

42. Solving the Risk Equation in Transitioning from Development to Production, Defense Science Board Task Force Report, May 25, 1983 (later published as DOD Manual 4245.7).

43. Risk Management as a Means of Direction and Control, Program Manager's Notebook, Fact Sheet Number 4.5, DSMC, June 1992.

44. Joint Logistics Commanders Guide for the Management of Joint Service Programs, DSMC, 1987.

45. Systems Engineering Management Guide, Second Edition, DSMC, 1990.

46. Integrated Logistics Support Guide, First Edition, DSMC, May 1986.

47. Joint Logistics Commanders Guide for the Management of Multinational Programs, DSMC, 1987.

48. Defense Manufacturing Management Course, DSMC.

49. DVAL Methodology — Methodology Overview/Executive Summary, Data Link Vulnerability Analysis Joint Test Force, November 1984.

ARMY DOCUMENTS

50. AR 10-4, U.S. Army Operational Test and Evaluation Agency.

51. AR 10-11, U.S. Army Materiel Development and Readiness Command.

52. AR 10-41, U.S. Army Training and Doctrine Command.

53. AR 10-42, U.S. Armed Forces Command.

54. AR 70-24, Special Procedures Pertaining to Nuclear Weapons System Development.

55. AR 70-60, Nuclear Survivability of Army Materiel.

56. AR 70-71, Nuclear, Biological, and Chemical Contamination Survivability of Army Materiel.

57. AR 71-1, System Acquisition Policy and Procedures.

58. AR 71-3, Force Development User Testing.

59. AR 73-1, Test and Evaluation Policy.

60. AR 700-127, Integrated Logistic Support.

61. DA PAM 73-1, Test and Evaluation Guide.

62. DA PAM 700-50, Integrated Logistic Support: Developmental Supportability Test and Evaluation Guide.

63. AMC PAM 70-2, AMC-TRADOC Materiel Acquisition Handbook.

64. TECOM PAM 310-4, Index of Test Operations Procedures and International Test Operations Procedures.

65. USAOTEA Operational Test and Evaluation Guide, Special Supplement, Continuous Comprehensive Evaluation Planning for Nuclear Hardness and Survivability Test and Evaluation (Draft).

NAVY DOCUMENTS

66. SECNAVINST 5000.1B, System Acquisition.

67. SECNAVINST 5000.39, USN Acquisition and Management of ILS for Systems and Equipment.

68. SECNAVINST 5000.42C, RDT&E Acquisition Procedures.

69. SECNAVINST 5430.67A, Assignment of Responsibilities for Research, Development, Test and Evaluation.

70. OPNAVINST 3070.1A, Operations Security.

71. (cancelled) OPNAVINST 3960.10C, Test and Evaluation (Draft).

72. OPNAVINST 5000.49, ILS in the Acquisition Process.

73. OPNAVINST 5440.47F, Mission and Functions of Operational Test and Evaluation Force (OPTEVFOR).

74. NAVMATINST 3960, Test and Evaluation.

75. COMOPTEVFORINST 3960.1D, Operational Test Director's Guide.

76. NAVSO P-2457, RDT&E/Acquisition Management Guide, 10th Edition.

77. COMOPTEVFOR, Project Analysis Guide.

78. COMOPTEVFOR, AD Number AD124168, Operational Test Director Guide.

AIR FORCE DOCUMENTS

79. AFR 23-36, Air Force Operational Test and Evaluation Center (AFOTEC).

80. AFR 55-30, Operations Security.

81. AFR 55-43, Management of Operational Test and Evaluation.

82. (cancelled) AFR 80-14, Test and Evaluation.

83. AFR 80-38, Management of the Air Force Systems Survivability Program.

84. AFR 800-2, Acquisition Program Management.

85. AFR 800-8, Integrated Logistic Support (ILS) Program.

86. CAFSCR 550-15, Statistical Process Control.

87. AFSCR 550-13, Producibility.

88. AFOTEC Regulation 23-1, Organization and Functions Air Force Operational Test and Evaluation Center (AFOTEC).

89. AFOTEC, Nuclear Survivability Handbook for Air Force OT&E.

90. AFOTEC, Electronic Warfare Operational Test and Evaluation.

91. AFLC Pamphlet 800-34, ACQ Logistic Management.

92. Rome Air Development Center, Standard Procedures for Air Force Operational Test and Evaluation.

MARINE CORPS DOCUMENTS

93. MCO 3960.2, Marine Corps Operational Test and Evaluation Activity (MCOTEA); establishment of, with changes.

94. MCO 5000.11B, Testing and Evaluation of Systems and Equipment for the Marine Corps.

OTHER GOVERNMENT DOCUMENTS

95. Public Law 99-661, National Defense Authorization Act for Fiscal Year 1987.

96. Public Law 92-156.

97. Public Law 98-94.

98. GAO/PEMD-87-17, Live Fire Testing: Evaluating DOD's Programs.

99. GAO/NSIAD-87-57, Operational Test and Evaluation Can Contribute More to Decision-Making.

100. GAO/PEMD-86-12BR, Bigeye Bomb: An Evaluation of DOD's Chemical and Developmental Tests.

101. GAO/PSAD-79-86, Review of Military Services Development Test and Evaluation of Six Weapon Systems Totaling an Estimated $12 Billion in Development and Procurement Costs.

102. GAO/PSAD-78-102, Operational Testing of Air Force Systems Requires Several Improvements.

103. Memorandum of Understanding Among the Governments of the French Republic, the Federal Republic of Germany, the United Kingdom of Great Britain and Northern Ireland, and the United States of America Relating to the Mutual Acceptance of Test and Evaluation for the Reciprocal Procurement of Defense Equipment.

104. Director, Defense Test and Evaluation Memorandum of August 2, 1983; Re sponse of Admiral Isham Linder, Director, Test and Evaluation to Supplemental Questions Posed by the Senate Governmental Affairs Committee.

105. Hearing Before the Senate Committee on Governmental Affairs, June 23, 1983.

OTHER REFERENCES

106. Anderson, J.E., "Towards the Goal of Improving MIL-STD Testing," Defense Systems Management Journal, Vol. 12, Number 2, pp 30-34, April 1976.

107. The BDM Corporation, "Application of Simulations to the OT&E Process," May 1974.

108. Bolino, John V., "Software T&E in the Department of Defense."

109. The BDM Corporation, "Functional Description of the Acquisition Test and Evaluation Process," July 8, 1983.

110. Hoivik, Thomas H., "The Navy Test and Evaluation Process in Major Systems Acquisition," DTIC Accession Number AD A035935, November 1976.

111. Johnson, Larry H., "Contractor Testing and the Army Test and Evaluation Master Plan for Full Engineering Development," Defense Systems Management College, Fort Belvoir, Virginia, 1983.

112. Rhoads, D.I., "Systems Engineering Guide for Management of Software Acquisition," pp IV-1 through IV-23, Defense Systems Management College, Fort Belvoir, Virginia, 1983.

113. The MITRE Corporation, "Software Reporting Metrics," November 1985.

114. Stevens, Roger T., John Wiley & Sons, New York, "Operational Test and Evaluation: A Systems Engineering Process," 1979, pp 49-56.

115. International Test and Evaluation Association, "Test and Evaluation of Command and Control Systems Developed Evolutionary Acquisition," May 1987.

116. Mann, Greg A., LTC, Air University Press, Maxwell AFB, Alabama, "The Role of Simulation in Operational Test and Evaluation," Report AU-ARI-83-10, August 1983.

117. Georgia Institute of Technology, "The Software Test and Evaluation Project."

118. Thomas, E.F., "Reliability Testing Pitfalls," Reliability and Maintainability Symposium, Vol. 7, Number 2, pp 78-85, 1974.

119. Watt, Charles K., "The Role of Test Beds in C3I," AFCEA 35th Annual Convention, June 15-17, 1982.

120. Willoughby, Willis J., "Reliability by Design, Not Chance," Defense Systems Management Journal, Vol. 12, Number 2, pp 12-18, April 1976.

APPENDIX E
INDEX

APPENDIX F
POINTS OF CONTACT
SERVICE TEST AND EVALUATION COURSES

ARMY

Commander U.S. Army Test and Evaluation Command
ATTN: AMSTE-TE
Aberdeen Proving Ground, MD 21005-5055

Commander U.S. Army Logistics Management Group
ATTN: AMXMC-ACM-MA
Fort Lee, VA 23801-6048

Commander U.S. Army Operational Test & Evaluation Command
ATTN: CSTE-ZA
4501 Ford Avenue
Alexandria, VA 22302

NAVY

Commander Space & Naval Warfare Systems Command
ATTN: Management Operations
National Center, Bldg. 1
Washington, DC 20361

Commander Operational Test & Evaluation Force
ATTN: Dest 02B
Norfolk, VA 23511-6388

AIR FORCE

Commander Air Force Operational Test & Evaluation Center
ATTN: Asst. Director, Training
Building 20130
Kirtland Air Force Base, NM 87117-7001

Commander Air Force Institute of Technology
ATTN: Student Operations
Wright-Patterson Air Force Base, OH 45433

OSD

Defense Test & Evaluation Professional Institute
Naval Air Warfare Center, Weapons Division
Code 02PI
Point Mugu, CA 93042

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS None
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Unlimited Distribution
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Defense Systems Mgmt College 6b. OFFICE SYMBOL *(If applicable)* FD-TE	7a. NAME OF MONITORING ORGANIZATION Same as 6a
6c. ADDRESS *(City, State, and ZIP Code)* 18th Street Fort Belvoir, Virginia 22060-5426	7b. ADDRESS *(City, State, and ZIP Code)* Same as 6c
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Same as 6a 8b. OFFICE SYMBOL *(If applicable)*	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS *(City, State, and ZIP Code)* Same as 6c	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE *(Include Security Classification)*
Test and Evaluation Management Guide

12. PERSONAL AUTHOR(S)
John Claxton

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT *(Year, Month, Day)*	15. PAGE COUNT
Final	FROM ______ TO ______	August 1993	292

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Importance of T&E; T&E Process; Types of T&E; Combined & Concurrent Testing; Intro to OT&E; Specialized Testing; Logistics Infrastructure T&E; International T&E Programs

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

This guide is one of a family of educational guides written from a DOD perspective (i.e., non-service peculiar). It is intended primarily for use in the courses at the Defense Systems Management College (DSMC) and secondarily as a desk reference for program and project management personnel. It has been developed to assist the acquisition community in obtaining a better understanding of who the decision-makers are and determining how and when to plan test and evaluation events.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Sylvia Nance	22b. TELEPHONE *(Include Area Code)* 703-805-2376 22c. OFFICE SYMBOL OS-PR

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE