TEST AND EVALUATION MANAGEMENT GUIDE

JANUARY 2005

FIFTH EDITION

PUBLISHED BY
THE DEFENSE ACQUISITION UNIVERSITY PRESS
FORT BELVOIR, VA 22060-5565

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **JAN 2005**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Test and Evaluation Management Guide Fifth Edition**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defense Acquisition University Press Fort Belvoir, VA 22060-5565**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**318**	

ACKNOWLEDGMENTS

This January 2005 update to the Defense Acquisition University's *Test and Evaluation Management Guide* includes updates from the Military Services, Defense Agencies, and other organizations, as well as changes in the May 2003 DoD 5000 series. Inputs were collated and finalized by DAU Program Director for Test and Evaluation Dr. John Claxton. Freelance writer Christina Cavoli and *Defense AT&L* editor-in-chief Collie Johnson edited the final document. Freelance designer and cartoonist Jim Elmore designed the cover, and final desktop publishing and layout was accomplished by Bartlett Communications and DAU Visual Information Specialist Deborah Gonzalez.

FOREWORD

This book is one of many technical management educational guides written from a Department of Defense perspective; i.e., non-Service specific. They are intended primarily for use in courses at Defense Acquisition University and secondarily as a desk reference for program and project management personnel. These guidebooks are written for current and potential acquisition management personnel who are familiar with basic terms and definitions employed in program offices. The guidebooks are designed to assist government and industry personnel in executing their management responsibilities relative to the acquisition and support of defense systems.

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

Dr. John D. Claxton
Program Director, T&E
Defense Acquisition University

CONTENTS

Chapter 12 — OT&E to Support Decision Reviews

MODULE IV — Test and Evaluation Planning

Chapter 13 — Evaluation

Chapter 14 — Modeling and Simulation Support to T&E

Chapter 15 — Test Resources

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 (SEP), which identifies levels of performance and assists the developer in correcting deficiencies. It is a significant element in the decision-making process, providing data that support trade-off analysis, risk reduction, and requirements refinement. Program decisions on system performance maturity and readiness to advance to the next phase of development take into consideration demonstrated performance. The issue of paramount importance to the servicemember user is system performance; i.e., will it fulfill the mission. The T&E process provides data that tell the user how well the system is performing during development and if it is ready for fielding. The Program Manager (PM) 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 tradeoffs. As stated in Department of Defense Directive (DoDD) 5000.1, *The Defense Acquisition System*, "Test and evaluation shall be integrated throughout the defense acquisition process. Test and evaluation shall be structured to provide essential information to decision makers, assess attainment of technical performance parameters, and determine whether systems are operationally effective, suitable, survivable, and safe for intended use. The conduct of test and evaluation, integrated with modeling and simulation, shall facilitate learning, assess technology maturity and interoperability, facilitate integration into fielded forces, and confirm performance against documented capability needs and adversary capabilities as described in the system threat assessment."[1]

1.2 TESTING AS A RISK MANAGEMENT TOOL

Correcting defects in weapons has been estimated to add from 10 percent to 30 percent to the cost of each item.[2] Such costly redesign and modification efforts can be reduced if carefully planned and executed T&E programs are used to detect and fix system deficiencies early in the acquisition process (Figure 1-1). Fixes instituted during early work efforts (Systems Integration (SI)) in the System Development and Demonstration (SDD) Phase cost significantly less than those implemented after the Critical Design Review (CDR), when most design decisions have already been made.

T&E 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 decision maker who must examine the critical issues and weigh the facts. Only the decision maker can determine the weight and importance that is to be attributed to a system's capabilities and shortcomings and the degree of risk that can be accepted. 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 (LCC) of a system and how decisions impact program expenditures.

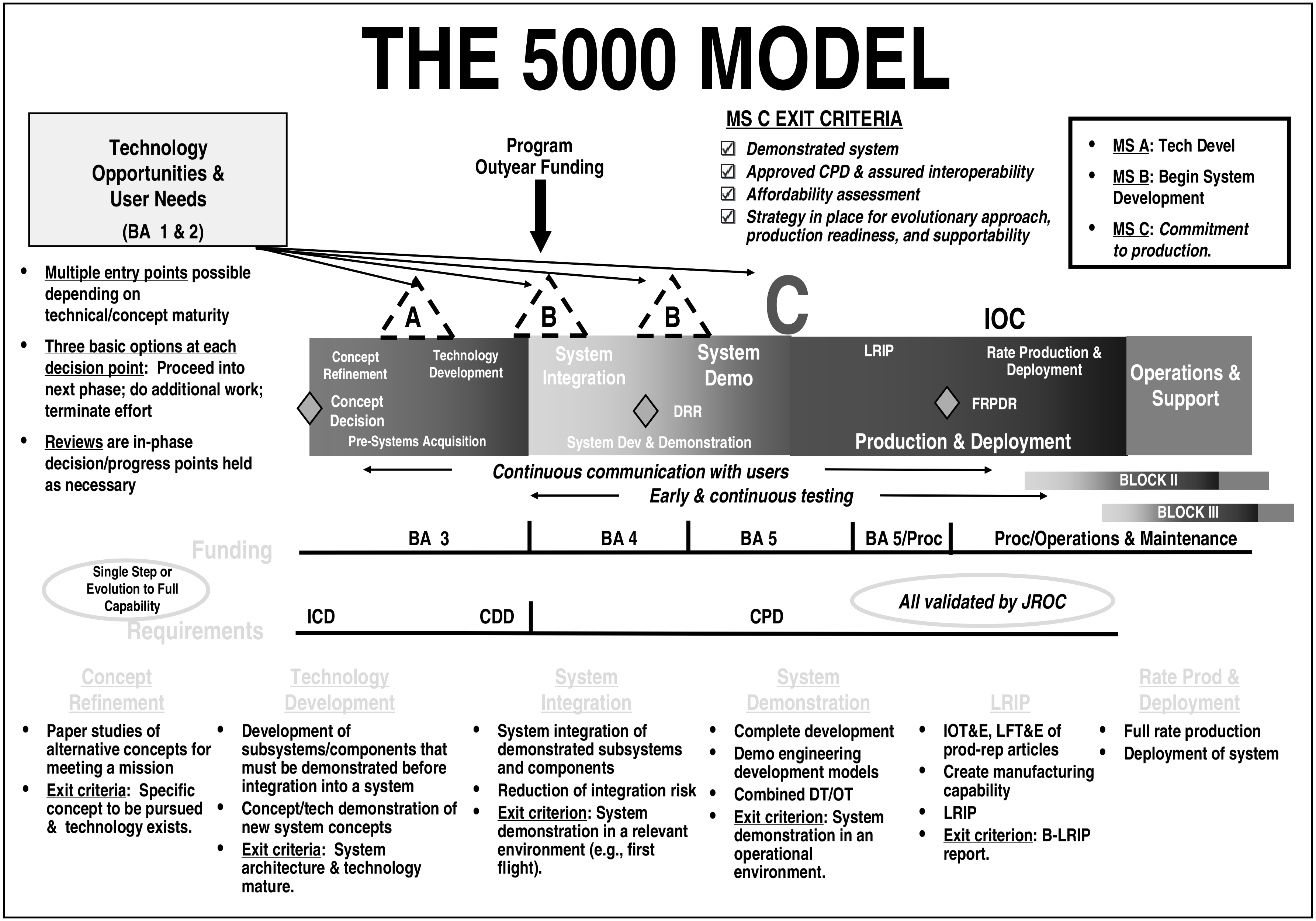


Figure 1-1. The 5000 Model

A Defense Science Board (DSB) 1999 Task Force focused on a broad overview of the state of T&E within the Department of Defense (DoD).[3] This group made the following observations about the T&E process:

- The focus of T&E should be on how to best support the acquisition process;
- T&E planning with Operational Test (OT) personnel should start early in the acquisition cycle;
- Distrust remains between the development and test communities;
- Contractor testing, developmental testing, and operational testing have some overlapping functions;
- Ensuring the test data are independently evaluated is the essential element, not the taking of the data itself;
- Responses to perceived test "failures" are often inappropriate and counterproductive.

The DSB Task Force (1983) developed a set of templates for use in establishing and maintaining low-risk programs. Each template described an area of risk and then specified technical methods for reducing that risk. PMs and test managers may wish to consult these templates for guidance in reducing the risks frequently associated with test programs. The DoD manual *Transition from Development to Production* contains sample risk management templates.[4]

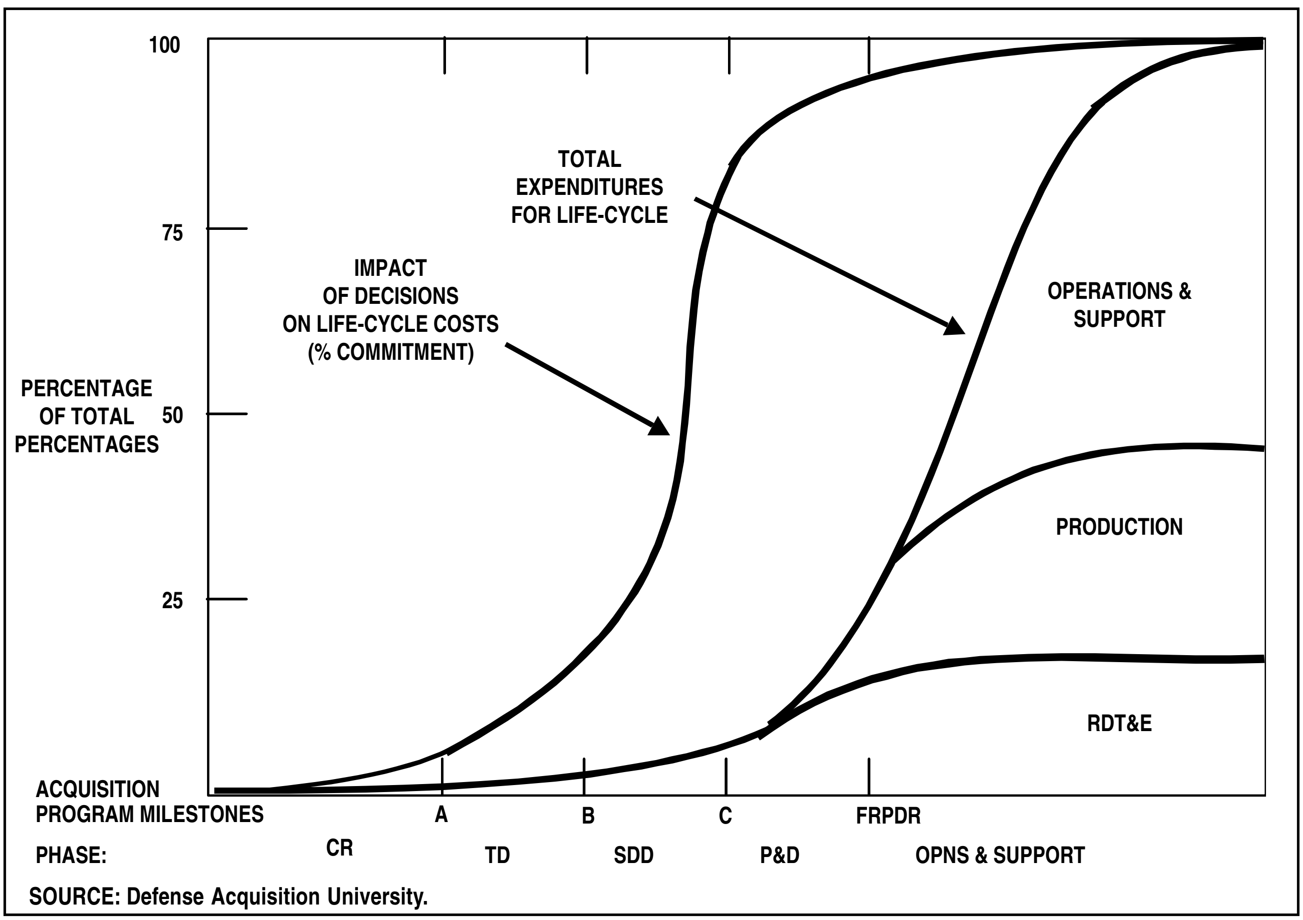


Figure 1-2. Life Cycle Cost Decision Impact and Expenditures

1.3 THE T&E CONTRIBUTION AT MAJOR MILESTONES

T&E progress is monitored by the Office of the Secretary of Defense (OSD) throughout the acquisition process. Their oversight extends to Major Defense Acquisition Programs (MDAPs) or designated acquisitions. T&E officials within OSD render independent assessments to the Defense Acquisition Board (DAB), the Defense Acquisition Executive (DAE), and the Secretary of Defense (SECDEF) at each system milestone review. 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;
- T&E, Modeling and Simulation (M&S), and data from other sources such as Service PMs, laboratories, industry developers, studies, and analyses.

At Milestone B, the OSD T&E assessments reflect an evaluation of system concepts and technology alternatives using early performance parameter objectives and thresholds found in an approved preliminary TEMP. At Milestone C, assessments include an evaluation of previously executed test plans and test results. At the Full Rate Production Decision Review (FRPDR), assessments include consideration of the operational effectiveness and suitability evaluations of 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 that improve performance to desired levels.

1.3.1 T&E Contributions Prior to Milestone B

During Concept Refinement (CR) and Technology Development (TD) activities prior to Milestone B, laboratory testing and M&S are conducted by the contractors and the development agency to demonstrate and assess the capabilities of key subsystems and components. The test and simulation designs are based on the operational needs documented in the Initial Capabilities Document (ICD). Studies, analyses, simulation, and test data are used by the development agency to explore and evaluate alternative concepts proposed to satisfy the user's needs. Also during this period, the Operational Test Agency (OTA) monitors CR and TD activities to gather information for future T&E planning and to provide effectiveness and suitability input desired by the PM. The OTA also conducts Early Operational Assessments (EOAs), as feasible, to assess the operational impact of candidate technical approaches and to assist in selecting preferred alternative system concepts.

The development agency prepares the Technology Development Strategy (TDS) with its early T&E strategy for Development Test and Evaluation (DT&E) and M&S. The TDS presents T&E plans for system design(s) engineering and performance evaluations. The OTA may provide an EOA. This information is incorporated into the PM's TEMP that documents the program's T&E strategy that supports the Milestone B decision to proceed to the next phase.

1.3.2 T&E Contributions Prior to Milestone C

During the SDD Phase, concepts approved for prototyping form the baseline that is used for detailed test planning. The design is matured into an Engineering Development Model (EDM), which is tested in its intended environment prior to Milestone C.

In SI the development agency conducts DT&E to assist with engineering design, system development, risk identification, and to evaluate the contractor's ability to attain desired technical performance in system specifications and achieve program objectives. The DT&E includes T&E of components, subsystems, and prototype development models. T&E of functional compatibility, interoperability, and integration with fielded and developing 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 logistics 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 (OT&E) agency should conduct an EOA for the Design Readiness Review (DRR) to estimate the system's potential to be operationally effective and suitable; identify needed modifications; and provide information on tactics, doctrine, organization, and personnel requirements. The early OT&E program is accomplished in an environment containing limited operational realism. Typical operational and support personnel are used to obtain early estimates of the user's capability to operate and maintain the system. Some of the most important products of user assessments of system maintainability and supportability are human factors and safety issues.

In Systems Demonstration, the objective is to design, fabricate, and test a preproduction system that closely approximates the final product. T&E activities of the EDM during this period yield much useful information. For example, data obtained during EDM T&E can be used to assist in evaluating the system's maintenance training requirements and the proposed training program. Test results generated during EDM T&E also support the user in refining and updating employment doctrine and tactics.

During Systems Demonstration, T&E 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 focus on areas for corrective actions;

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

   (d) Estimate the Reliability, Availability, and Maintainability (RAM) of the system after it is deployed;

   (e) Determine whether the system is safe and ready for Low Rate Initial Production (LRIP);

   (f) Evaluate effects on performance of any configuration changes caused by correcting deficiencies, modifications, or Product Improvements (PIs);

   (g) Assess human factors and identify limiting factors.

(2) Assess the technical risk and evaluate the tradeoffs among specifications, operational requirements, LCCs, and schedules;

(3) Assess the survivability, vulnerability, and logistics 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 weapons system;

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

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

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

The development agency evaluates the results of T&E for review by the Service headquarters and the Service acquisition review council prior to system acquisition review by the Milestone Decision Authority (MDA). The evaluation includes the results of testing and supporting information, conclusions, and recommendations for further engineering development. At the same time, the OT&E agency prepares an Independent Operational Assessment (IOA), which contains estimates of the system's potential operational effectiveness and suitability. The Operational Assessment (OA) provides a permanent record of OT&E events, an audit trail of OT&E data, test results, conclusions, and recommendations. This information is used to prepare for Milestone C and supports a recommendation of whether the design and performance of the system in development justifies proceeding into LRIP.

1.3.3 T&E Contributions Prior to Full Rate Production Decision Review

The development agency transitions the final design to LRIP while fixing and verifying any technical problems discovered during the final testing of the EDM in its intended environment. The maturity of the hardware and software configurations and logistics support system available from LRIP are assessed when the development agency considers certifying the system's readiness for Initial Operational Test and Evaluation (IOT&E).

IOT&E is conducted prior to the production decision at FRPDR to:

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

(2) Identify operational deficiencies;

(3) 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 Operations 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.

Before the certification of readiness for IOT&E, the developer should have obtained the Joint Interoperability Test Command's (JITC)'s certification of interoperability for the system components. In parallel with IOT&E, Live Fire Test and Evaluation (LFT&E) may be used to evaluate vulnerability or lethality of a weapon system as appropriate and as required by public law. The PM's briefing and the Beyond Low Rate Initial Production (BLRIP) report address the risks of proceeding into Full Rate Production (FRP).

1.3.4 T&E Contributions After the Full Rate Production Decision Review

After FRPDR, when the FRP decision is normally made, T&E activities continue to provide important insights. Tests described in the TEMP but not conducted during earlier phases are completed. The residual DT&E may include extreme weather testing and testing corrected deficiencies. System elements are integrated into the final operational configuration, and development testing is completed when all system performance requirements are met. During FRP, 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.

Post-production testing requirements may result from an acquisition strategy calling for increment changes to accommodate accumulated engineering changes or the application of Preplanned Product Improvements (P^3Is). 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; if system performance is significantly changed, some level of operational testing may be appropriate.

OT&E activities continue after the FRP 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. This verifies the operational effectiveness and suitability of the production system, determines if deficiencies identified during the IOT&E have been corrected, and evaluates areas not tested during IOT&E due to system limitations. Additional FOT&E may be conducted over the life of the system to refine doctrine, tactics, techniques, and training programs, and to evaluate future increments, modifications, and upgrades.

The OT&E agency prepares an OA report at the conclusion of each 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 after SDD. Deficiencies that are not corrected are recorded.

A final report on FOT&E may also be prepared by the using command test team, emphasizing 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 its missions 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 personnel 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 acquisition phases of the system development cycle.

1.4 SUMMARY

"Risk management," according to *Transition from Development to Production,* "is the means by which the program areas of vulnerability and concern are identified and managed."[5] T&E 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 July 2000 General Accounting Office (GAO) Report NSIAD-00-199, *Best Practices: A More Constructive Test Approach is Key to Better Weapon System Outcomes*.[6] The summary serves to underscore the importance of the T&E process as a whole:

- Problems found late in development signal weaknesses in testing and evaluation;
- Early testing to validate product knowledge is a best practice;
- Different incentives make testing a more constructive factor in commercial programs than in weapon system programs.

"To lessen the dependence on testing late in development and to foster a more constructive relationship between program managers and testers, GAO recommends that the Secretary of Defense instruct acquisition managers to structure test plans around the attainment of increasing levels of product maturity, orchestrate the right mix of tools to validate these maturity levels, and build and resource acquisition strategies around this approach. GAO also recommends that validation of lower levels of product maturity not be deferred to the third level. Finally, GAO recommends that the Secretary require that weapon systems demonstrate a specified level of product maturity before major programmatic approvals."[7]

ENDNOTES

1. DoDD 5000.1, *The Defense Acquisition System*, May 12, 2003, p. 8.

2. BDM Corporation, *Functional Description of the Acquisition Test and Evaluation Process*, July 8, 1983.

3. Defense Science Board Task Force Report, *Solving the Risk Equation in Transitioning from Development to Production*, May 25, 1983 (later published as DoD Manual 4245.7).

4. DoD 4245.7-M, *Transition from Development to Production,* September 1985.

5. *Ibid.*

6. GAO/NSIAD-00-199, *Best Practices: A more Constructive Test Approach is Key to Better Weapon System Outcomes,* July 2000, p. 11.

7. *Ibid.*

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 tradeoffs, and estimate satisfaction of operational requirements. As a system undergoes design and development, the iterative process of testing moves gradually from a concentration on Development Test and Evaluation (DT&E), which is concerned chiefly with attainment of engineering design goals, to increasingly comprehensive Operational Test and Evaluation (OT&E), which focuses on questions of operational effectiveness, suitability, and survivability. Although there are usually separate Development Test (DT) and Operational Test (OT) events, DT&E and OT&E are not necessarily serial phases in the evolution of a weapon system development. Combined or concurrent DT and OT are encouraged when appropriate, i.e., conferring possible cost or time savings.[1]

T&E 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 to testing, and new approaches are discussed in Chapter 17 of this 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.

T&E has many useful functions and provides information to many customers. 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 significantly in 2003 to change its primary objective to acquisition of quality products that satisfy user needs with measurable improvements to mission capability and operational support, in a timely manner, and at a fair and reasonable price. As it is now structured, the defense system life cycle is to replicate the preferred acquisition strategy of an Evolutionary Acquisition (EA) process that uses either incremental or spiral development processes. The three major elements—pre-system acquisition, system acquisition, and sustainment—may include the following five phases:

(1) Concept Refinement (CR)

(2) Technology Development (TD)

(3) System Development and Demonstration (SDD)

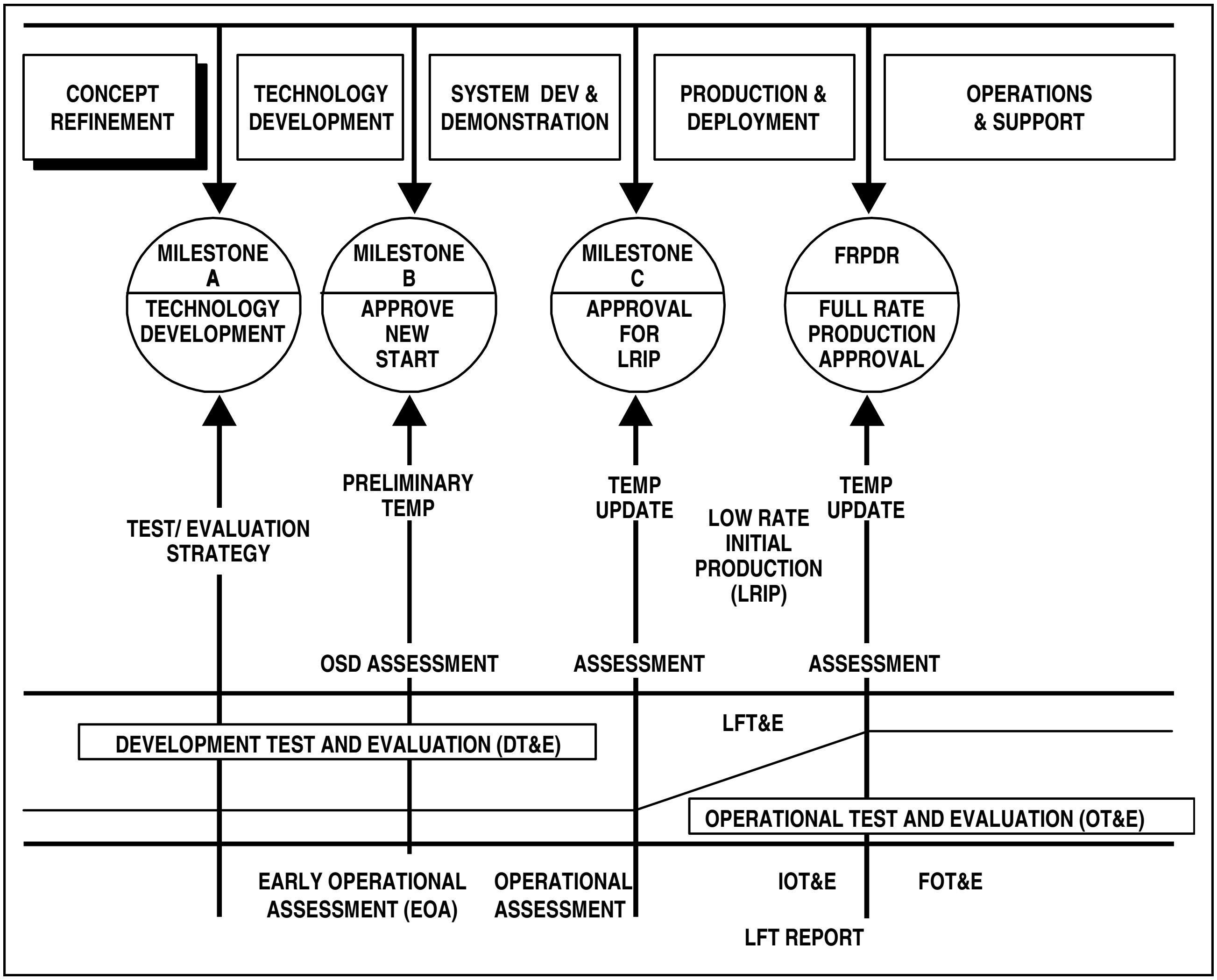


Figure 2-1. Testing and the Acquisition Process

(4) Production and Deployment (P&D)

(5) Operations and Support (O&S).

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

The following description of the defense system acquisition process, summarized from Department of Defense Instruction (DoDI) 5000.2, *Operation of the Defense Acquisition System*[2], shows how T&E fits within the context of the larger process.

2.2.1 Concept Refinement and Technology Development

A defense pre-system acquisition process begins with the requirements process identification of a material need and the development of the Initial Capabilities Document (ICD). A decision is made to start CR, and the Technology Development Strategy (TDS) evolves with the initial test planning concepts. CR ends when the MDA approves the preferred solution resulting from the Analysis of Alternatives (AoA), identifying Measures of Effectiveness (MOEs), and approves the associated TDS. The TD Phase follows a Milestone A decision during which the risks of the relevant technologies of the alternative selected for satisfying the user's needs are investigated. Shortly after the milestone decision, an integrated team begins transitioning the test planning in the TDS into the evaluation strategy for formulation of a Test and Evaluation Master Plan (TEMP). The TD effort concludes with a decision review at Milestone B when an affordable increment of militarily useful capability has been identified, the technology for that increment has been demonstrated in a relevant environment, and a system can be developed for production within a short time frame (normally less than 5 years) or the MDA decides to terminate the effort. Typical T&E-related documents at the Milestone B review are the Acquisition Decision Memorandum (ADM) (exit criteria), ICD, Capability Development Document (CDD) (performance parameters), Acquisition Strategy, System Threat Assessment (STA), an Early Operational Assessment (EOA), and the TEMP. Additional program management documents prepared before Milestone B include: the AoA, Independent Cost Estimate (ICE), and the concept baseline version of the Acquisition Program Baseline (APB), which summarizes the weapon's functional specifications, performance parameters, and cost and schedule objectives.

The program office for major programs (Office of the Secretary of Defense (OSD) oversight) must give consideration to requesting a waiver for full-up system level Live Fire Testing (LFT) and identification of Low Rate Initial Production (LRIP) quantities for Initial Operational Test and Evaluation (IOT&E).

2.2.2 System Development and Demonstration

The Milestone B decision is program initiation for systems acquisition and establishes broad objectives for program cost, schedule, and technical performance. After the Milestone B decision for a program start, the Systems Integration (SI) work effort begins during which a selected concept, typically brassboard or early prototype, is refined through systems engineering, analysis, and design. Systems engineering must manage all requirements as an integrated set of design constraints that are allocated down through the various levels of design (Figure 2-2). This work effort ends when the integration of the system components has been demonstrated through adequate developmental testing of prototypes. The Design Readiness Review (DRR) decision evaluates design maturity and readiness to either enter into System Demonstration or make a change to the acquisition strategy. The System Demonstration work effort is intended to demonstrate the ability of the system to operate in a useful way consistent with the approved Key Performance Parameters (KPPs). Work advances the design to an Engineering Development Model (EDM) that is evaluated for readiness to enter LRIP. "Successful development test and evaluation to assess technical progress against critical technical parameters, early operational assessments, and, where proven capabilities exist, the use of modeling and simulation to demonstrate system integration are critical during this effort. Deficiencies encountered in testing prior to Milestone C shall be resolved prior to proceeding beyond LRIP."[3] The EDM should have demonstrated acceptable performance in DT&E and the Operational Assessment (OA), with acceptable interoperability and operational supportability.

Preparation for the Milestone C decision establishes more refined cost, schedule, and performance objectives and thresholds, and the Capability Production Document (CPD) establishes the criteria for testing of LRIP items. Documents interesting to the T&E manager at the time of the Milestone C review include the ADM (exit criteria), updated TEMP, updated STA, AoA, Development Baseline, development testing results, and the OA.

2.2.3 Production and Deployment

During the LRIP work effort, the purpose is to achieve an operational capability that satisfies mission needs. The selected system design and

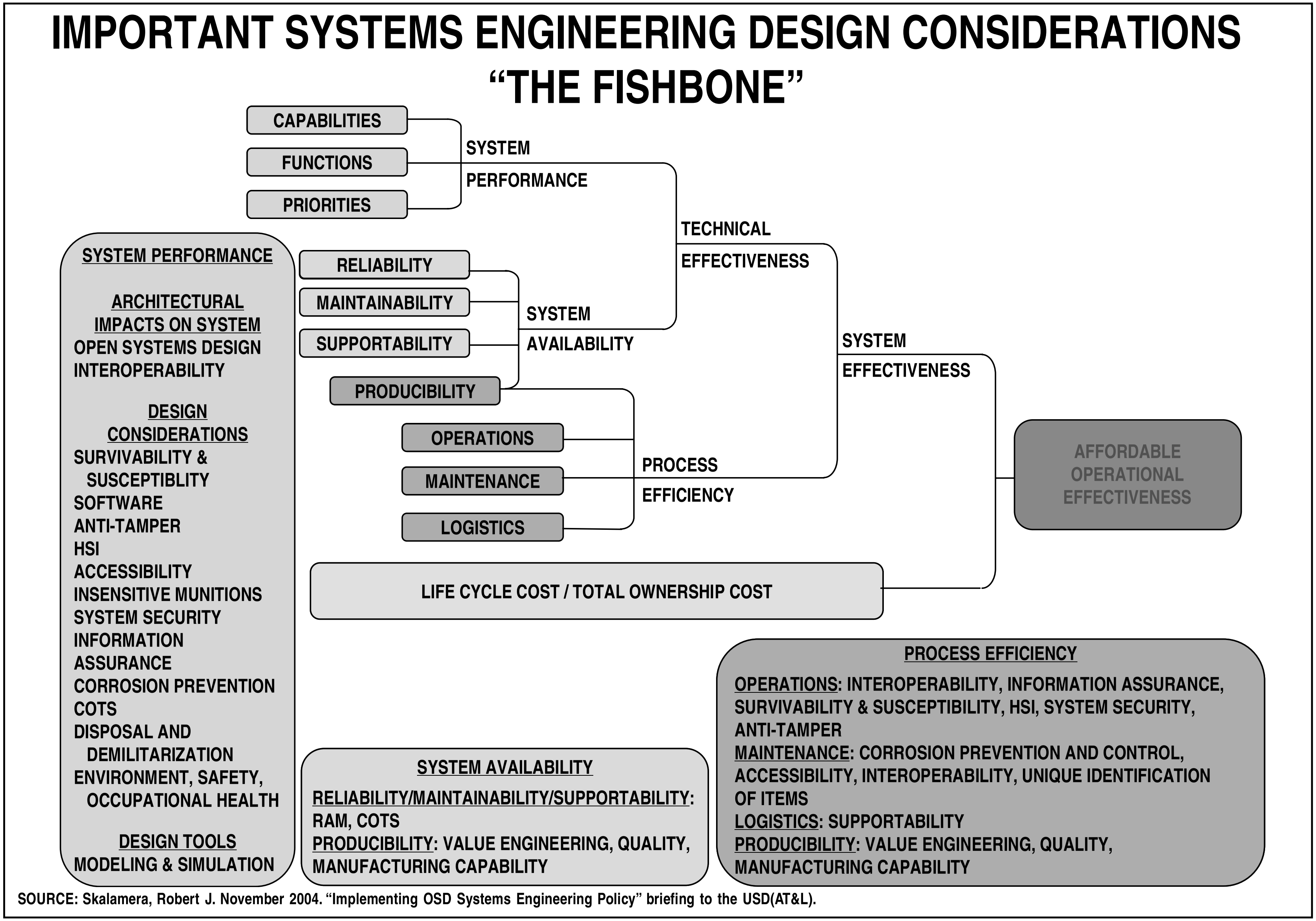


Figure 2-2. Requirements to Design

its principal items of support are fabricated as production configuration models. Test articles normally are subjected to qualification testing, full-up LFT, and IOT&E. This work effort ends with the Full Rate Production Decision Review (FRPDR) marking entry into Full Rate Production (FRP) and deployment of the system for Initial Operational Capability (IOC). Key documents for the T&E manager at the time of the FRPDR are the updated TEMP, development testing results, the Service IOT&E report, and Live Fire Test Report. For Acquisition Category (ACAT) I and designated oversight programs, the Director of Operational Test and Evaluation (DOT&E) is required by law to document the assessment of the adequacy of IOT&E and the reported operational effectiveness and suitability of the system. This is done in the Beyond LRIP (BLRIP) Report. Also mandated by law is the requirement for the DOT&E to submit the Live Fire Test Report prior to the program proceeding beyond LRIP. These DOT&E Reports may be submitted as a single document.

2.2.4 Operations and Support

The production continues at full rate allowing continued deployment of the system to operating locations and achievement of Full Operational Capability (FOC). This phase may include major modifications to the production configuration, increment upgrades, and related Follow-on Operational Test and Evaluation (FOT&E) (OA). Approval for major modifications should identify the actions and resources needed to achieve and maintain operational readiness and support objectives. The high cost of changes may require initiation of the modification as a new program. To determine whether major upgrades/modifications are necessary or deficiencies warrant consideration of replacement, the MDA may review the impact of proposed changes on system operational effectiveness, suitability, and readiness.

2.3 T&E AND THE SYSTEMS ENGINEERING PROCESS (SEP)

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.

Systems engineering, as defined in the DoD context, 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 decisionmaking (Figure 2-3).

A system's life cycle begins with the user's needs, which are expressed as constraints, and the required capabilities 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 in the past. Modern systems are too complex for a small number of generalists to manage; systems require in-depth knowledge of a broad range of areas and technical disciplines. System engineers coordinate the many specialized engineers involved in the concurrent engineering process through Integrated

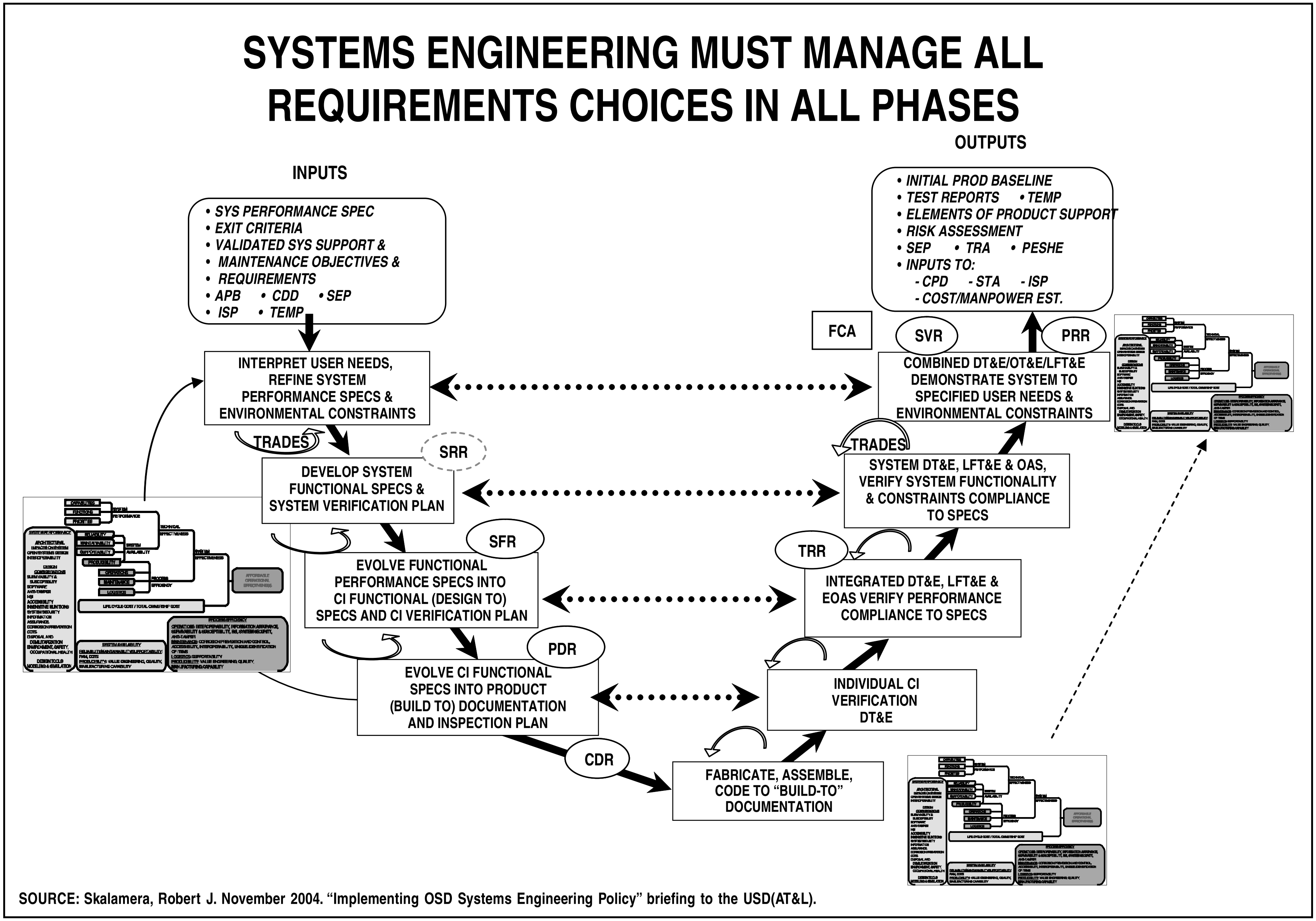


Figure 2-3. The Systems Engineering Process

Product and Process Development (IPPD). Integrated Product Teams (IPTs) are responsible for the integration of the components into a system.

Through interdisciplinary integration, a systems engineer manages the progress of product definition from system level to configuration-item level, detailed level, deficiency correction, and modifications/Product Improvements (PIs). 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 (TPM), trade-off analysis, and specifications.

What happens during systems engineering? The process determines what specialists are required, what segments and Non-Developmental Items (NDIs) are used, design performance limits, trade-off criteria, how to test, when to test, how to document (specifications), and what management controls to apply (TPM and design reviews).

Development and operational testing support the technical reviews by providing feedback to the SEP. More information on the reviews is contained in Chapter 8.

2.3.1 The Systems Engineering Process

The SEP 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 requirements analysis, functional analysis and allocation, synthesis, and verification of performance (T&E), which support decisions on tradeoffs and formalize the description of system elements. The system engineering plan is described in Chapter 16 of the Defense Acquisition University guide to *Systems Engineering Fundamentals*, January 2001.

The requirements analysis activity is a process used by the program office, in concert with the user, to establish and refine operational and design requirements that result in the proper balance between performance and cost within affordability constraints. Requirements analysis shall be conducted iteratively with Functional Analysis/Allocation (FA/A) to develop and refine system-level functional and performance requirements, external interfaces, and provide traceability among user requirements and design requirements.

The FA/A activity identifies what the system, component, or part must do. It normally works from the top downward ensuring requirements traceability and examining alternative concepts. This is done without assuming how functions will be accomplished. The product is a series of alternative Functional Flow Block Diagrams (FFBDs). 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 CR and TD phases, developmental testers assist the functional analysis activity to help determine what each component's role will be as part of the system being developed. Performance requirements are allocated to system components.

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 (RASs), 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 prototype testing to determine how the components will perform assigned functions to assist this synthesis activity.

The verification loop and decision activity allows tradeoff 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 (LCCs); effectiveness; Reliability, Availability, and Maintainability (RAM); risk limits; and schedule. It is repeated at each level of development. The verification and decision activity is assisted by developmental testers during the later SDD 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 Audit (PCA) and Functional Configuration Audit (FCA). Operational testers may 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 2-4, Systems Engineering Process and Test and Evaluation, depicts the activities and their interactions.

2.3.2 Technical Management Planning

The technical management planning incorporates top-level management planning for the integration of all system design activities. Its purpose is to develop the organizational mechanisms for direction and control, and identify personnel for the attainment of cost, performance, and schedule objectives. Planning defines and describes the type and degree of system engineering management, the SEP, and the integration of related engineering programs. The design evolution process forms the basis for comprehensive T&E planning.

The TEMP must be consistent with technical management planning. The testing program outlined in the TEMP must provide the TPMs data required for all design decision points, audits,

TESTING	DEFINED IN	PERFORMED AGAINST	PARTICIPANTS
	SYSTEMS ENGINEERING PROCESS		
QUALIFICATION ACCEPTANCE	SOW CONTRACT SPECS WBS	SPECIFICATION MISSION PROFILE/ ENVIRONMENT	PROJ ENGR DPML FUNC ENGR Kr/GOVT TEAM
	TEST AND EVALUATION		
SYSTEM DT & OT	TEMP DT/OT PLANS	STA/CDD/CPD OPS CONCEPT MAINT CONCEPT	DT&E/DOT&E DT/OT PERSONNEL USER/SUPT PERS TEST IPT

Figure 2-4. Systems Engineering Process and Test and Evaluation

and reviews that are a part of the system's engineering process. The configuration management process controls the baseline for the test programs and incorporates design modifications to the baseline determined to be necessary by T&E.

The TEMP and technical management planning 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. Technical thresholds and objectives 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 TPM Program. For example, failure criteria for reliability thresholds during OT&E testing must be delineated and agreed upon by the Program Manager (PM) and the Operational Test Director (OTD), and reflected in the TEMP.

2.3.3 Technical Performance Measurement

TPM identifies critical technical parameters that are at a higher level of risk during design. It tracks T&E 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 the systems engineering progress. For more information, see Chapter 8. Figure 2-5 depicts the technical reviews that usually take place during the SEP and the related specification documents.

2.3.4 System Baselining and T&E

The SEP establishes phase baselines throughout the acquisition cycle. These baselines (functional, allocated, product) can be modified with the results of engineering and testing. The testing used to prove the technical baselines is rarely the same as the operational testing of requirements.

Related to the baseline is the process of configuration management. Configuration management benefits the T&E community in two ways. Through configuration management, the baseline 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

T&E is the deliberate and rational generation of performance data, which describes the nature of the emerging system and the transformation of data into information useful for 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 of Defense Acquisition Acronyms and Terms*, produced by the Defense Acquisition University defines "Test" and "Test and Evaluation" as follows:[4]

> A "test" is any program or procedure that is designed to obtain, verify, or provide data for the evaluation of any of the following: 1) progress in accomplishing developmental objectives; 2) the

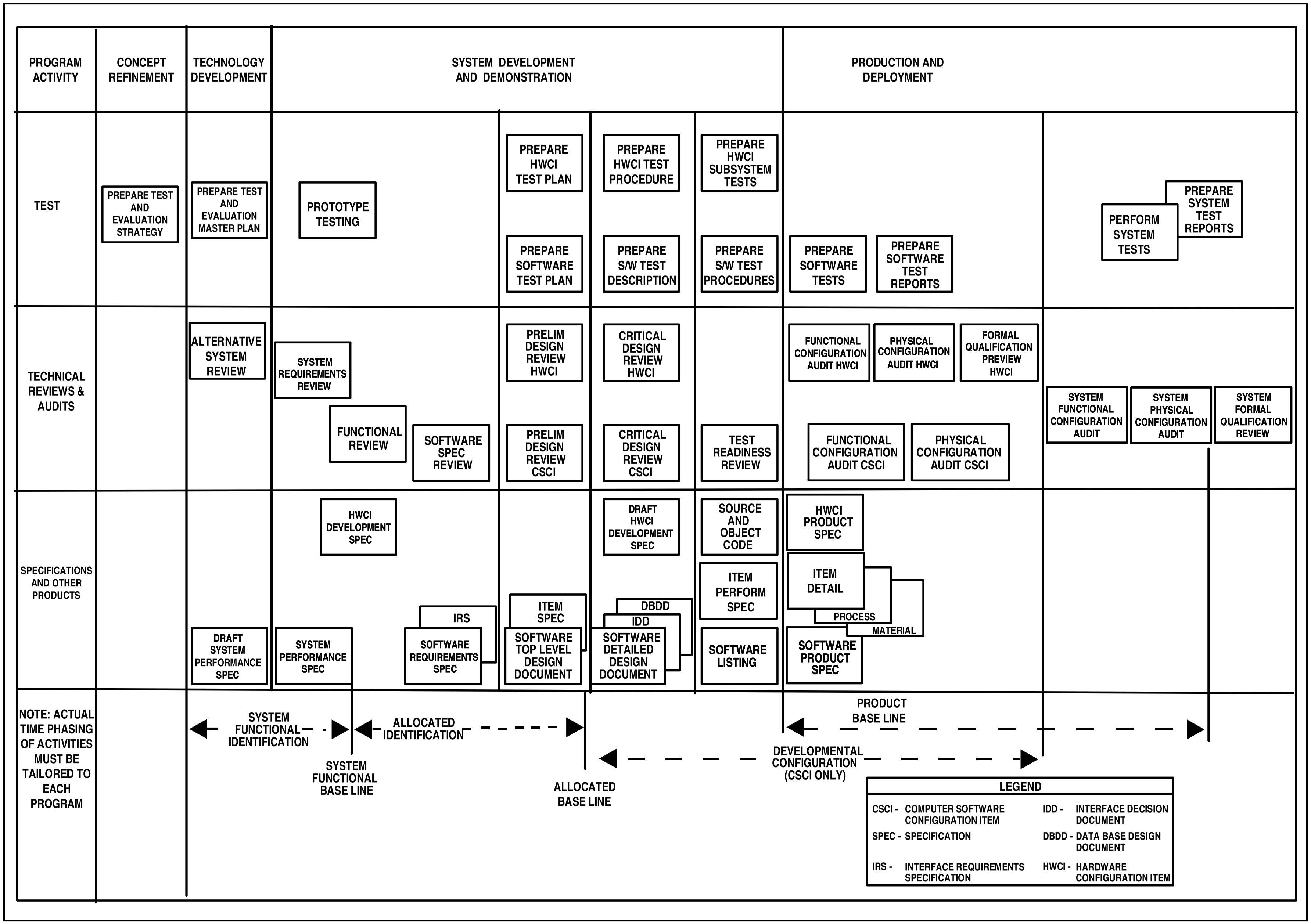


Figure 2-5. Design Reviews

performance, operational capability, and suitability of systems, subsystems, components, and equipment items; and 3) the vulnerability and lethality of systems, subsystems, components, and equipment items.

"Test and Evaluation" is the process by which a system or components are exercised and results analyzed to provide performance-related information. The information has many uses including risk identification and risk mitigation and empirical data to validate models and simulations. T&E enables an assessment of the attainment of technical performance, specifications, and system maturity to determine whether systems are operationally effective, suitable, and survivable for intended use, and/or lethality.

2.5 THE DOD T&E PROCESS

The DoD Test and Evaluation Process (Figure 2-6) is an iterative five-step process that provides answers to critical T&E questions for decision makers at various times during a system acquisition. The T&E process begins during the formative stages of the program with the T&E Coordination Function, in which the information needs of the various decision makers are formulated in conjunction with the development of the program requirements, acquisition strategy, and AoAs.

Given certain foundation documentation, Step 1 is the identification of T&E information required by the decision maker. The required information usually centers on the current system under test, which may be in the form of concepts, prototypes, EDMs, or production representative/production systems, depending on the acquisition

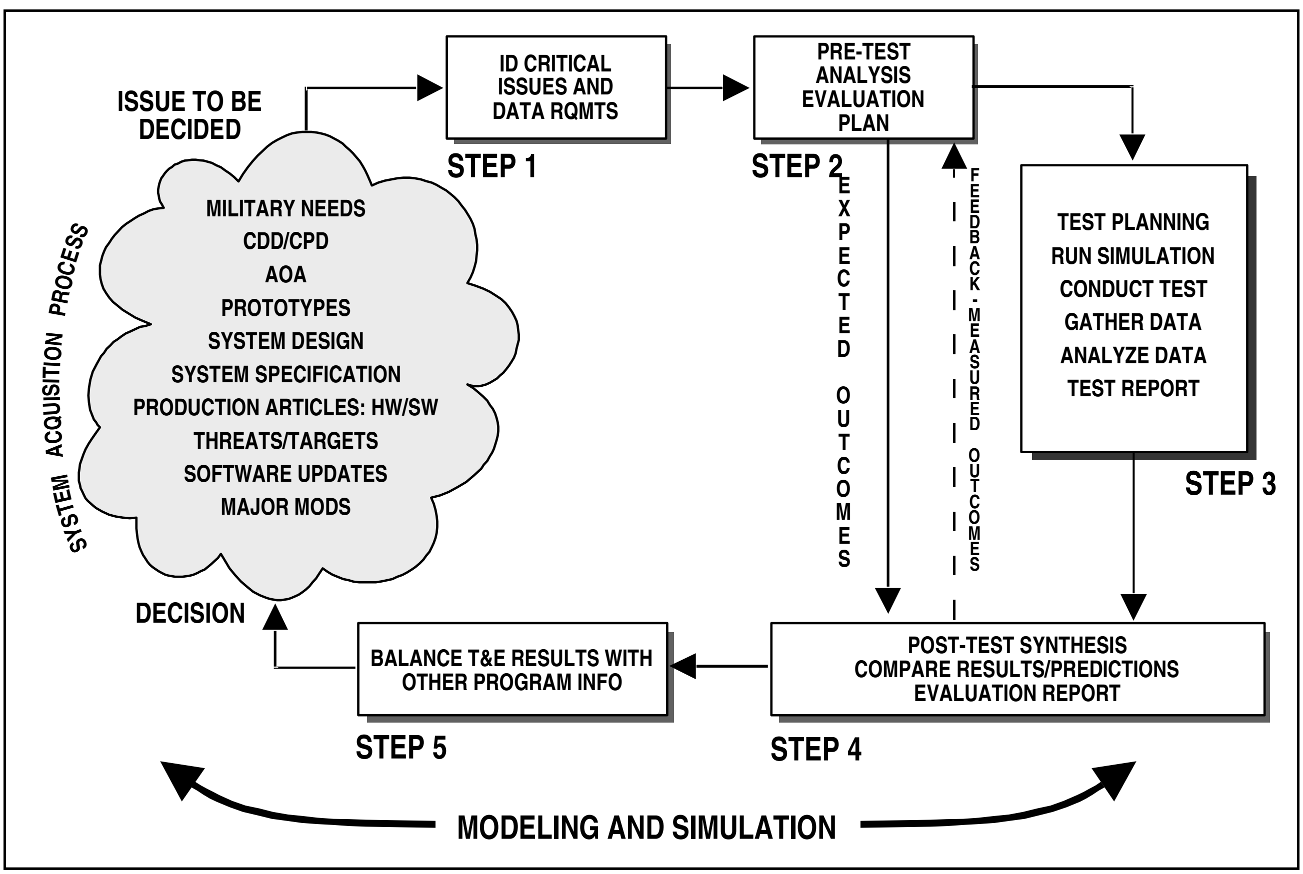


Figure 2-6. DoD Test and Evaluation Process

phase. The required information consists of performance evaluations of effectiveness and suitability, providing insights into how well the system meets the user's needs at a point in time.

Step 2 is the pre-test analysis of the evaluation objectives from Step 1 to determine the types and quantities of data needed, the results expected or anticipated from the tests, and the analytical tools needed to conduct the tests and evaluations. The use of validated models and simulation systems during pre-test analysis can aid in determining: how to design test scenarios; how to set up the test environment; how to properly instrument the test; how to staff and control test resources; how best to sequence the test trials; and how to estimate outcomes.

Step 3, test activity and data management, is the actual test activity planning. Tests are conducted and data management for data requirements is identified in Step 2. T&E managers determine what valid data exist in historical files that can be applied and what new data must be developed through testing. The necessary tests are planned and executed to accumulate sufficient data to support analysis. Data are screened for completeness, accuracy, and validity before being used for Step 4.

Step 4, post-test synthesis and evaluation, is the comparison of the measured outcomes (test data) from Step 3 with the expected outcomes from Step 2, tempered with technical and operational judgment. This is where data are synthesized into information. When the measured outcomes differ from the expected outcomes, the test conditions and procedures must be reexamined to determine if the performance deviations are real or were the result of test conditions, such as lack of fidelity in computer simulation, insufficient or incorrect test support assets, instrumentation error, or faulty test processes. The assumptions of tactics, operational environment, systems performance parameters, and logistics support must have been carefully chosen, fully described, and documented prior to test. Modeling and Simulation (M&S) may normally be used during the data analysis to extend the evaluation of performance effectiveness and suitability.

Step 5 is when the decision maker weighs the T&E information against other programmatic information to decide a proper course of action. This process may identify additional requirements for test data and iterate the DoD T&E process again.

2.6 SUMMARY

T&E is an engineering tool used to identify technical risk throughout the defense system acquisition cycle and a process for verifying performance. This iterative cycle consists of acquisition phases separated by discrete milestones. The DoD T&E process consists of developmental and operational testing that is used to support engineering design and programmatic reviews. This T&E process forms an important part of the SEP used by system developers and aids in the decision process used by senior decision authorities in DoD.

ENDNOTES

1. DoDI 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003.

2. *Ibid.*

3. *Ibid.*

4. Defense Acquisition University Press. *The Glossary of Defense Acquisition Acronyms and Terms,* Vol. 11, September 2003.

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 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

Congress has shown a longstanding 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, 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 (OTAs). The Navy Operational Test and Evaluation Force (OPTEVFOR) was established in the late 1960s. In 1983, similar concerns led Congress to direct the establishment of the independent Office of the Director, Operational Test and Evaluation (DOT&E) within OSD. In 1985, a report released by another President's Blue Ribbon Commission on Defense Management, this time chaired by David Packard, made significant recommendations on the management and oversight of DoD's acquisition process, specifically, T&E. All the Commission's recommendations have not been implemented, and the full impact of these recommendations is not yet realized. In Fiscal Year (FY) 87 the Defense Authorization Act required Live Fire Testing (LFT) of weapon systems before the Production Phase begins. The earmarking of authorizations and appropriations for DoD funding, specific testing requirements, and acquisition reform legislation continues to indicate the will of Congress for DoD implementation.

Congress requires DoD to provide the following reports that include information on T&E:

- Selected Acquisition Report (SAR). Within the cost, schedule, and performance data in the report, SAR describes Acquisition Category (ACAT) I system characteristics required and outlines significant progress and problems encountered. It lists tests completed and issues identified during testing. The Program Manager (PM) uses the Consolidated Acquisition Reporting System (CARS) software to prepare the SAR.

- Annual System Operational Test Report. This report is provided by the DOT&E to the SECDEF and the committees on Armed Services, National Security, and Appropriations. The report provides a narrative and resource summary of all Operational Test and Evaluation (OT&E) and related issues, activities, and assessments. When oversight of LFT was moved to DOT&E, this issue was added to the report.

- Beyond Low Rate Initial Production (BLRIP) Report. Before proceeding to BLRIP for each Major Defense Acquisition Program (MDAP), DOT&E 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. When oversight of LFT was moved to the DOT&E, the Live Fire Test Report was added to the BLRIP Report content.

- Foreign Comparative Test (FCT) Report. The Under Secretary of Defense for Acquisition, Technology, and Logistics (USD(AT&L)) should notify Congress a minimum of 30 days prior to the commitment of funds for initiation of new FCT evaluations of equipment produced by selected allied and friendly foreign countries.

3.3 OSD OVERSIGHT STRUCTURE

The DoD organization for the oversight of T&E is illustrated in Figure 3-1. In USD(AT&L), Development Test and Evaluation (DT&E) oversight is performed by the Deputy Director, Developmental Test and Evaluation, Defense Systems (DT&E/DS), and the DOT&E provides OT&E oversight for the SECDEF. The management of MDAPs in OSD is performed by the Defense Acquisition Executive (DAE), who uses the Defense Acquisition Board (DAB) and Overarching Integrated Product Teams (OIPTs) to process information for decisions.

3.3.1 Defense Acquisition Executive (DAE)

The DAE position, established in September 1986, is held by the USD(AT&L). As the DAE, the USD(AT&L) uses the DAB and its OIPTs to provide the senior-level decision process for the acquisition of weapon systems. The DAE responsibilities include establishing policies for acquisition (including procurement, Research and Development (R&D), logistics, development testing, and contracts administration) for all elements of DoD. The DAE's charter includes the authority over the Services and defense agencies on policy, procedure, and execution of the weapon systems 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 co-chair is the Vice Chairman of the Joint Chiefs of Staff (VCJCS). The DAB is composed of the department's senior acquisition officials, including the DOT&E. The DAB conducts business through OIPTs and provides decisions on ACAT ID programs.[1]

3.3.3 Deputy Director, Developmental Test and Evaluation, Defense Systems (DT&E/DS)

The DT&E/DS serves as the principal staff assistant and advisor to the USD(AT&L) for T&E matters. The Director, Defense Systems works for the Deputy Under Secretary of Defense for Acquisition and Technology (DUSD(A&T)) and has authority and responsibility for all DT&E conducted on designated MDAPs and for FCT. The DT&E/DS is illustrated in Figure 3-1.

3.3.3.1 Duties of the DT&E/DS

Within the acquisition community, the DT&E/DS:

- Serves as the focal point for coordination of all MDAP Test and Evaluation Master Plans (TEMPs). Recommends approval by the OIPT lead of the DT&E portion of TEMPs;

- Reviews MDAP documentation for DT&E implications and resource requirements to provide comments to the OIPT, USD(AT&L) DAE or DAB;

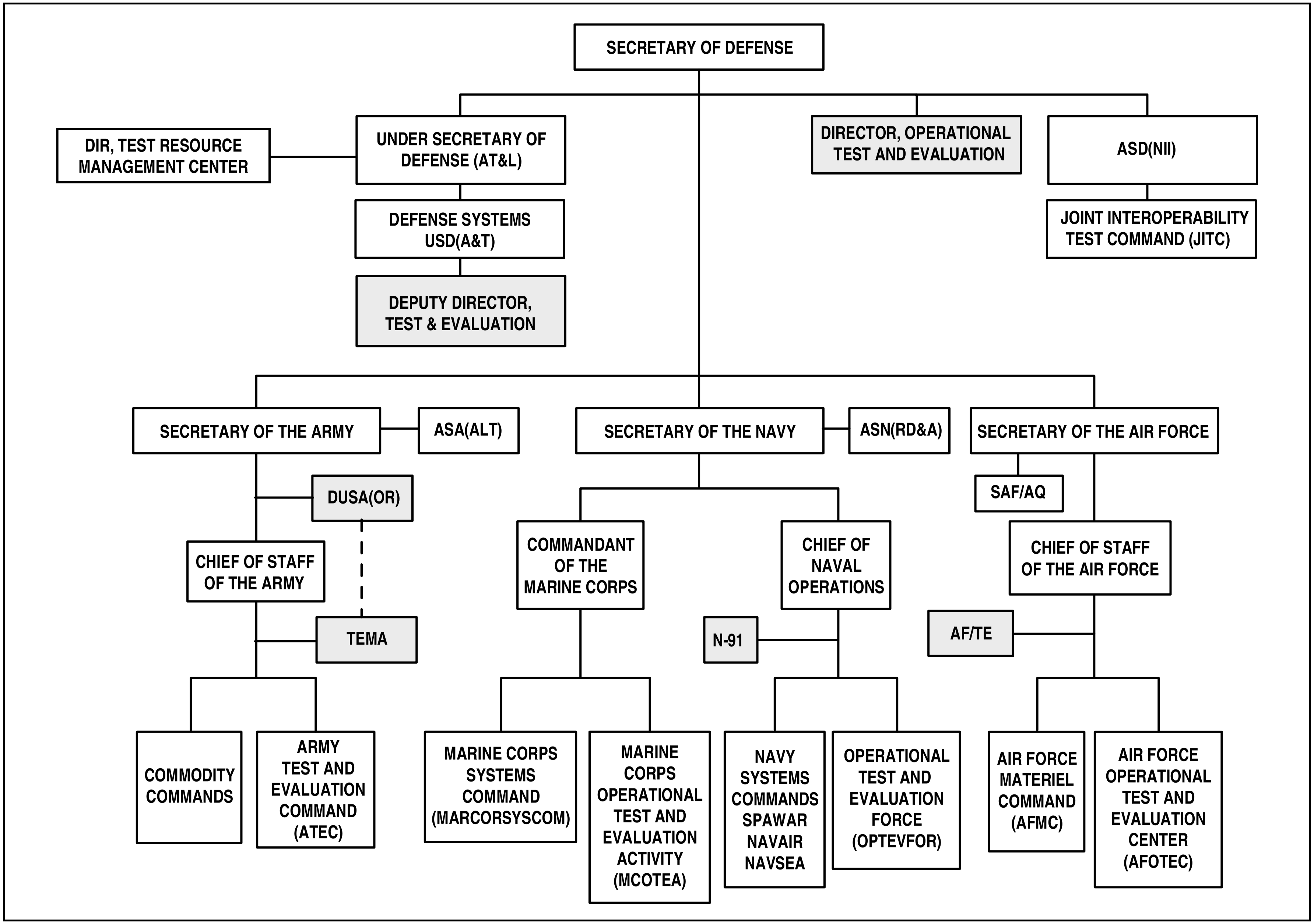


Figure 3-1. DoD Test and Evaluation Organization

- Monitors DT&E to ensure the adequacy of testing and to assess test results;

- Provides assessments of the Defense Acquisition Executive Summary (DAES) and technical assessments of DT&E and Systems Engineering Processes (SEPs) conducted on a weapon system;

- Provides advice and makes recommendations to OSD senior leadership, and issues guidance to the Component Acquisition Executives (CAEs) with respect to DT&E;

- Leads the defense systems efforts for joint testing and is a member of the Joint T&E Program's Technical Advisory Board for Joint Development Test and Evaluation programs.

3.3.3.2 DT&E/DS and Service Reports

During the testing of ACAT I and designated weapon systems, the DT&E/DS and Services interaction includes the following reporting requirements:

- A TEMP (either preliminary or updated, as appropriate) must be provided for consideration and approval before each milestone review, starting with Milestone B.

- A technical assessment of Developmental T&E is provided to the DS and DOT&E listing the T&E results, conclusions, and recommendations prior to a milestone decision or the final decision to proceed to BLRIP.

3.3.4 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 Congress. The DOT&E's responsibility to Congress is to provide an unbiased window of insight into the operational effectiveness, suitability, and survivability of new weapon systems.

3.3.4.1 Duties and Functions of the DOT&E

The specific duties of DOT&E, outlined in DoD Directive (DoDD) 5141.2, *Director of Operational Test and Evaluation,* include:[2]

- 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 ACAT I and oversight TEMPs for approval of OT&E and approving the OT&E funding for major systems acquisition;

- Approving operational test plans on all major defense acquisition systems and designated oversight programs prior to system starting initial operational testing (approval in writing required before operational testing may begin); oversight extends into Follow-on Operational Test and Evaluation (FOT&E);

- 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 OT&E performed;

- Presenting the BLRIP report to the SECDEF and to congressional Committees on Armed Services and Appropriations on the adequacy of Live Fire Test and Evaluation (LFT&E) and OT&E, and analyzing if results confirm the system's operational effectiveness and suitability;

- Providing oversight and approval of major program LFT.

3.3.4.2 DOT&E and Service Interactions

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

- A draft copy of the Operational Test Plan concept for review;
- Significant test plan changes;
- The final Service IOT&E report, which must be submitted to DOT&E before the Full Rate Production Decision Review (FRPDR);
- The LFT&E plan for approval, and the Service Live Fire Test Report for 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 Assistant Secretary of the Army for Acquisition, Logistics, and Technology (ASA(ALT)) has the principal responsibility for all Department of the Army (DA) matters and policy related to acquisition, logistics, technology, procurement, the industrial base, and security cooperation. Additionally, the ASA(ALT) serves as the Army Acquisition Executive (AAE). The AAE administers acquisition programs by developing/

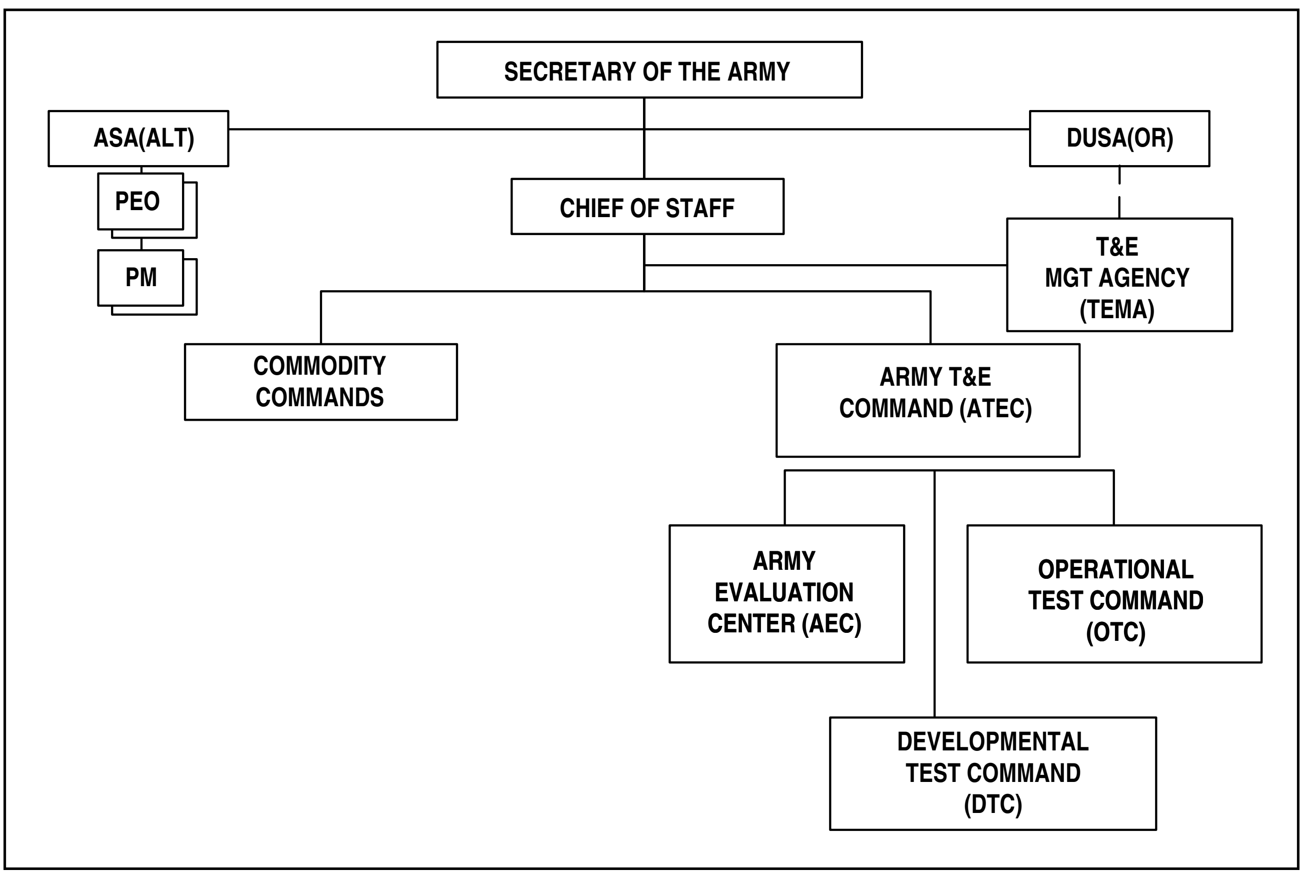


Figure 3-2. Army Test and Evaluation Organization

promulgating acquisition policies and procedures as well as appointing, supervising, and evaluating assigned Program Executive Officers (PEOs) and direct reporting PMs.

3.4.1.2 Army T&E Executive

The Deputy Under Secretary of the Army for Operations Research (DUSA(OR)), establishes, reviews, enforces, and supervises Army T&E policy and procedures including overseeing all Army T&E associated with the system research, development, and acquisition of all materiel and Command, Control, Communications, Computers and Intelligence (C^4I)/Information Technology (IT) systems. The Test and Evaluation Management Agency (TEMA) within the Office of the Chief of Staff, Army, provides the DUSA(OR) oversight. As delegated by the AAE, the DUSA(OR) is the sole Headquarters, Department of the Army (HQDA) approval authority for TEMPs.

3.4.1.3 Army Test and Evaluation Command

The U.S. Army Test and Evaluation Command (ATEC) supports the systems acquisition, force development, and experimentation processes through overall management of the Army's T&E programs. In this role, ATEC manages the Army's developmental and operational testing, all system evaluation, and management of joint T&E. ATEC is the Army's independent Operational Test Activity (OTA) reporting directly to the Vice Chief of Staff, Army (VCSA) through the Director of the Army Staff (DAS).

3.4.1.3.1 Army Developmental Testers

The U.S. Army Developmental Test Command (DTC) within ATEC has the primary responsibility for conducting Developmental Tests (DTs) for the Army. DTC responsibilities include:

- Perform the duties of governmental developmental tester for all Army systems except for those systems assigned to the Communications-Electronics Research, Development, and Engineering Center (CERDEC) of the Army Materiel Command (AMC) Research, Development and Engineering Command (RDECOM) by HQDA (Deputy Chief of Staff, G-6), Medical Command (MEDCOM), Intelligence and Security Command (INSCOM), Space and Missile Defense Command (SMDC), and Army Corps of Engineers (ACE);

- Provide test facilities and testing expertise in support of the acquisition of Army and other defense systems;

- Operate and maintain the Army's portion of the Major Range and Test Facility Base (MRTFB) (except for the United States Army Kwajalein Atoll (USAKA) and the High Energy Laser System Test Facility (HELSTF)) in accordance with DoDD 3200.11, *Major Range and Test Facility Base (MRTFB)*, May 1, 2002;

- Provide testers with a safety release for systems before the start of pretest training for tests that use soldiers as test participants;

- Provide safety confirmations for milestone acquisition decision reviews and the materiel release decision;

- Manage the Army Test Incident Reporting System (ATIRS).

3.4.1.3.2 Army Operational Testers

The U.S. Army Operational Test Command (OTC) within ATEC has the primary responsibility for conducting Operational Tests (OTs) for the Army and supporting Army participation in Joint T&E. OTC responsibilities include:

- Perform the duties of operational tester for all Army systems except for those systems assigned to MEDCOM, INSCOM, SMDC, and ACE;

- Perform the duties of operational tester for assigned multi-Service OTs and (on a customer service basis) for Training and Doctrine Command (TRADOC) Force Development Tests and Experimentation (FDTE);

- Provide test facilities and testing expertise in support of the acquisition of Army and other defense systems, and for other customers on a cost-reimbursable and as-available basis;

- Program and budget the funds to support OT except out of cycle tests (which are usually paid for by the proponent);

- Develop and submit OT and FDTE Outline Test Plans (OTPs) to the Army's Test Schedule and Review Committee (TSARC).

3.4.1.3.3 Army Evaluation Center

The U.S. Army Evaluation Center (AEC) is an independent subordinate element within ATEC that has the primary responsibility for conducting Army system evaluations and system assessments in support of the systems acquisition process. Decision makers use AEC's independent report addressing an Army system's operational effectiveness, suitability, and survivability. AEC responsibilities include:

- Perform the duties of system evaluator for all Army systems except for those systems assigned to MEDCOM, INSCOM, and the commercial items assigned to ACE; conduct Continuous Evaluation (CE) on all assigned systems;

- Develop and promulgate evaluation capabilities and methodologies;

- Coordinate system evaluation resources through the TSARC;

- Preview programmed system evaluation requirements for possible use of Modeling and Simulation (M&S) to enhance evaluation and reduce costs; perform manpower and personnel integration (MANPRINT) assessments in coordination with Deputy Chief of Staff (DCS), G-1 Army Research Laboratory-Human Research and Engineering Division/Directorate (ARL-HRED);

- Perform the Integrated Logistics Support (ILS) program surveillance for Army systems. Perform independent logistics supportability assessments and report them to the Army Logistician and other interested members of the acquisition community.

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 (SECNAV) has assigned general and specific Research, Development, Test, and Evaluation (RDT&E) responsibilities to the Assistant Secretary of the Navy (Research, Development, and Acquisition) (ASN(RDA)) and to the Chief of Naval Operations (CNO). The CNO has responsibility for ensuring the adequacy of the Navy's overall T&E program. The T&E policy and guidance are exercised through the Director 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.

3.4.2.1 Navy DT&E Organizations

The Navy's senior systems development authority is divided among the commanders of the system commands with Naval Air Systems Command (NAVAIR) developing and performing DT&E on aircraft and their essential weapon systems; Naval Sea Systems Command (NAVSEA) developing and performing DT&E on ships, submarines, and their associated weapon systems; and Space and Naval Warfare Systems Command (SPAWAR) developing and performing DT&E on all other systems. System acquisition is controlled by a

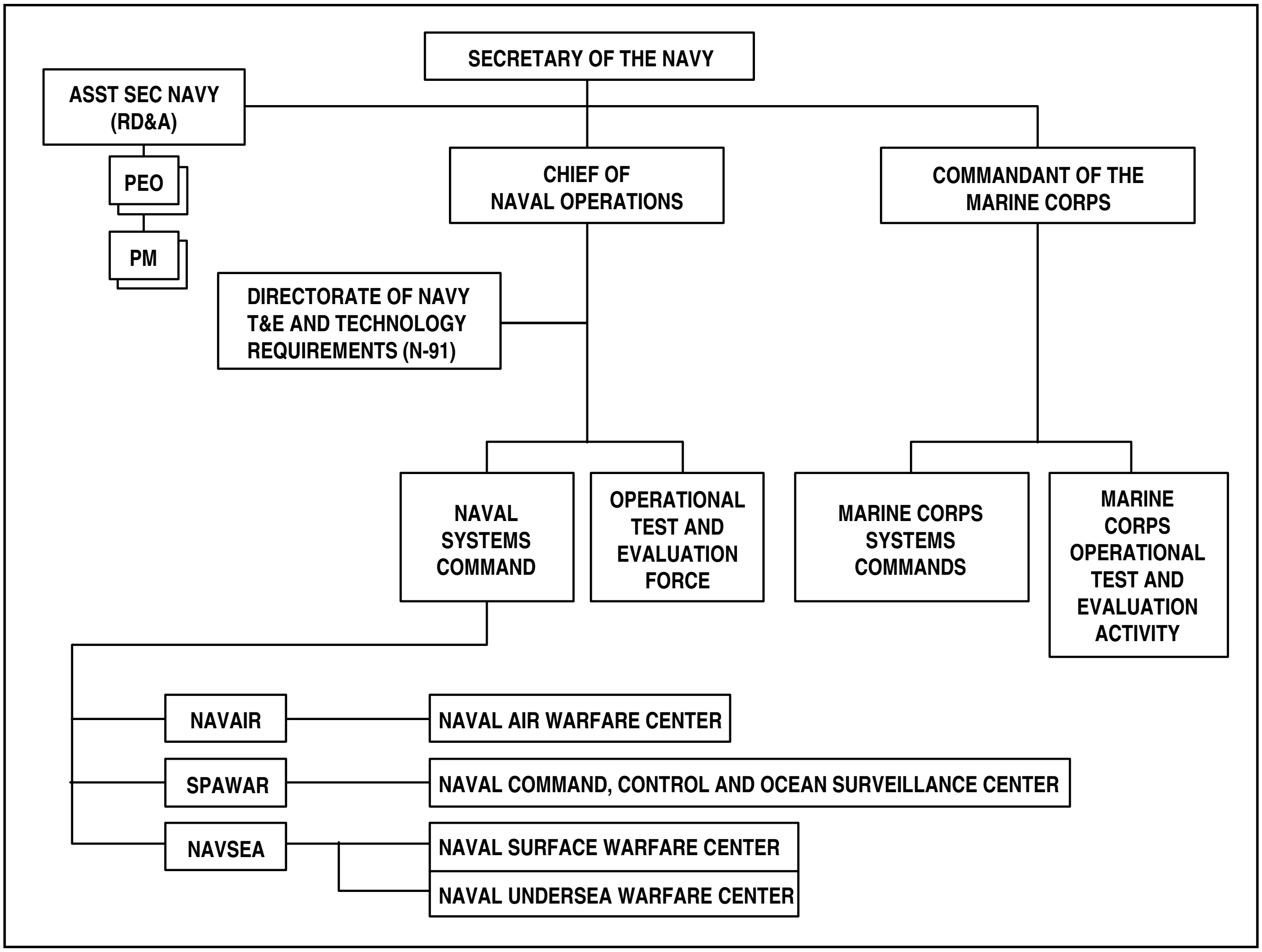


Figure 3-3. Navy Test and Evaluation Organization

chartered PM 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 T&E planning in the TEMP. Developing agencies are responsible for:

- Developing test issues based on the thresholds established by the user in the requirements documents;
- Identifying the testing facilities and resources required to conduct the DT&E;
- Developing the DT&E test reports and quick-look reports.

3.4.2.2 Navy Operational Test and Evaluation Force

The Commander, Operational Test and Evaluation Force (COMOPTEVFOR) commands the Navy's independent OT&E activity and reports directly to the CNO. The functions of the COMOPTEVFOR include:

- Establishing early liaison with the developing agency to ensure an understanding of the requirements and plans;
- Reviewing acquisition program documentation to ensure that documents are adequate to support a meaningful T&E program;
- 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;

- Reporting directly to the CNO, the President of the Board of Inspection and Survey (PRES-INSURV) is responsible for conducting acceptance trials of new ships and aircraft acquisitions and is the primary Navy authority for Production Acceptance Test and Evaluation (PAT&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 (SAF/AQ) is the senior-level authority for Research, Development, and Acquisition (RD&A) within the Air Force. The Director of Space Acquisition (SAF/USA) obtains space assets for all Services. As illustrated in Figure 3-4, the SAF/AQ is an advisor to the Secretary of the Air Force and interfaces directly with the DT&E and DOT&E. The SAF/AQ receives DT&E and OT&E results as a part of the acquisition decision process. Within the SAF/AQ structure, there is a military deputy (acquisition) who is the Air Force primary staff officer with responsibility for RD&A. This staff officer 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

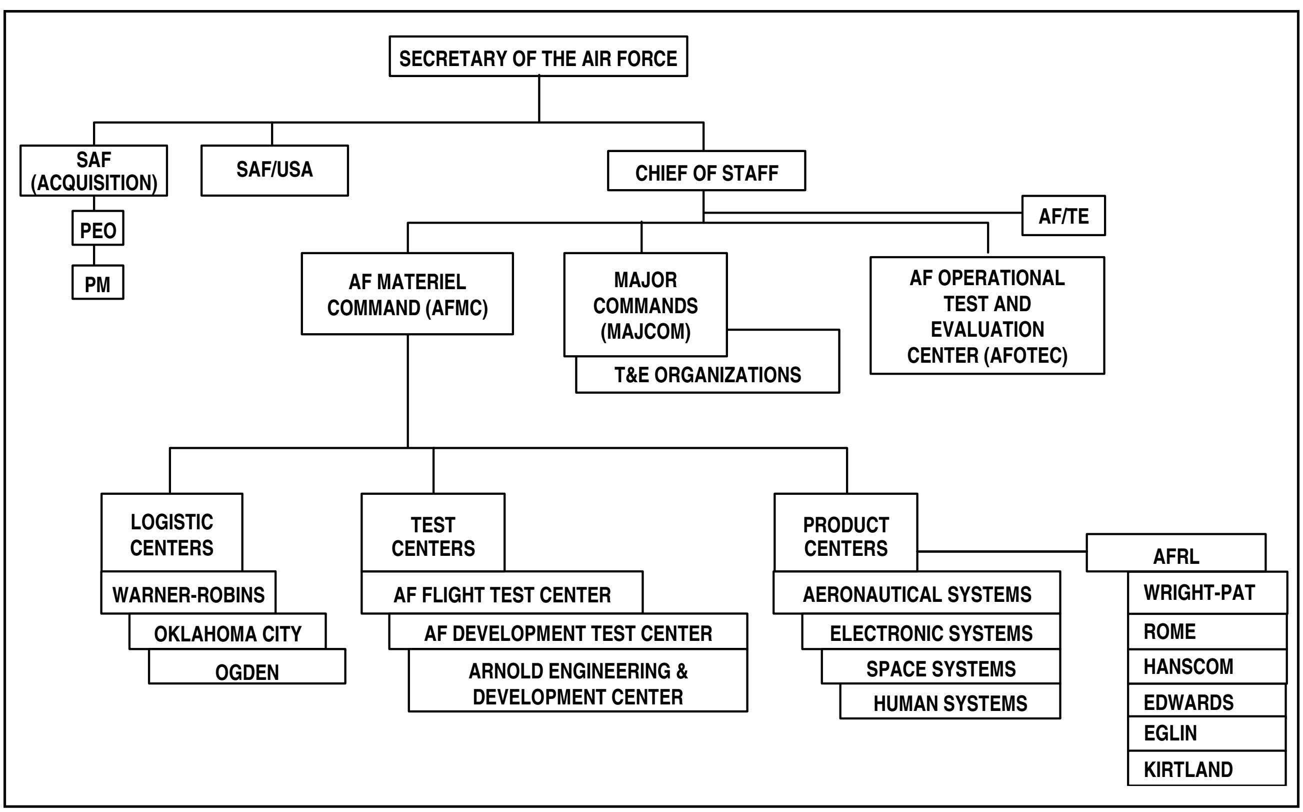


Figure 3-4. Air Force Test and Evaluation Organization

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 weapon systems, support systems, and equipment; and conducts all DT&E. The acquisition PMs 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, AFMC retains management responsibility for developing and testing system improvements, enhancements, or upgrades. The Air Force Space Command (AFSC) is responsible for DT&E of space and missile systems.

3.4.3.3 Air Force OT&E Organization

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, AFOTEC, is responsible to the Secretary of the Air Force and the Chief of Staff for the independent T&E of all major and non-major systems acquisitions. The Commander is supported by the operational commands and others in planning and conducting OT&E.

The AFOTEC reviews operational requirements, employment concepts, tactics, maintenance concepts, and training requirements before conducting 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

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. The DCS/R&D's 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 and 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 Activity

The Marine Corps Operational Test and Evaluation Activity (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, conducting, and reporting of OT&E. Marine Corps aviation systems are tested by OPTEVFOR. The Fleet Marine Force performs troop T&E of materiel development in an operational environment.

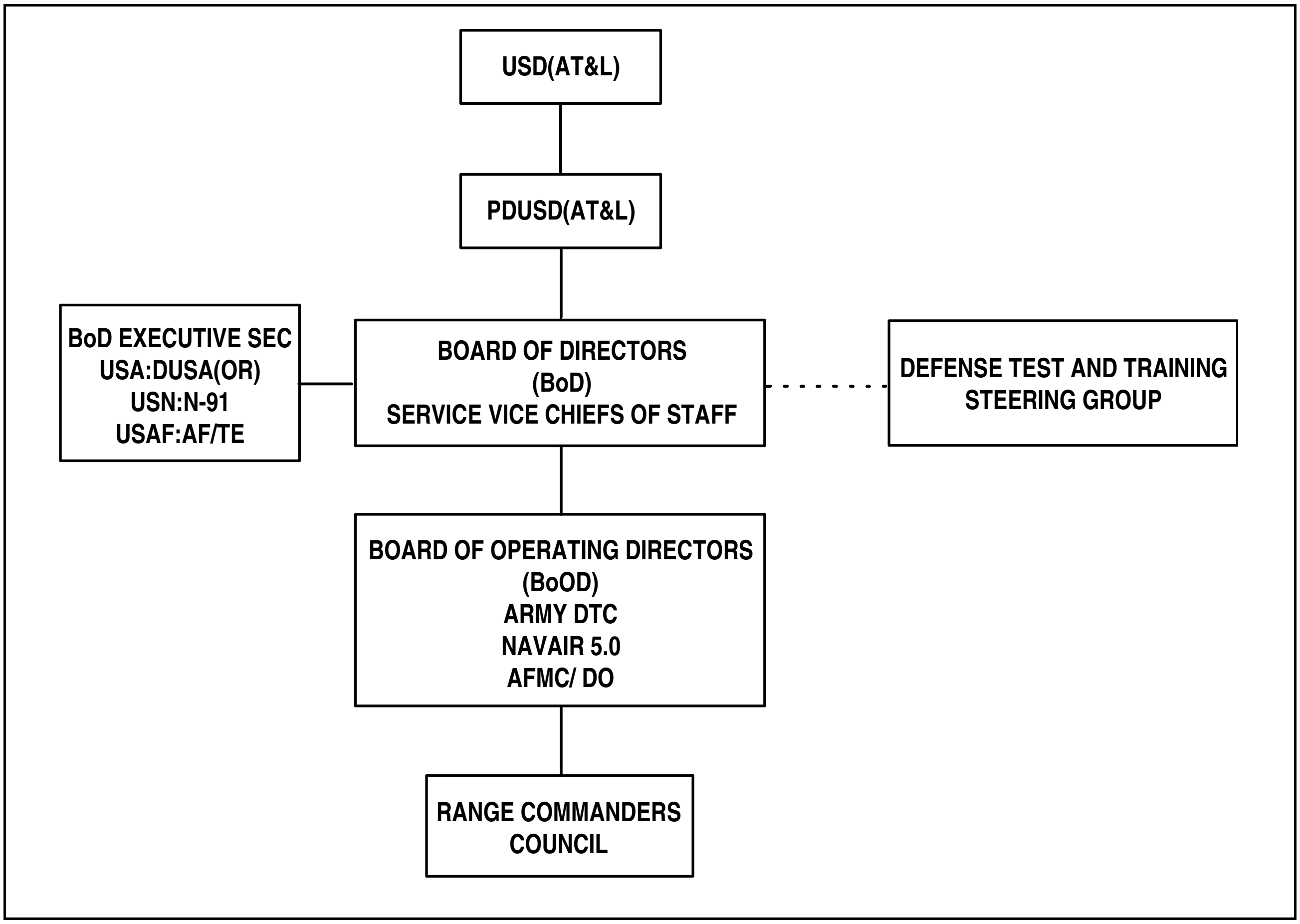


Figure 3-5. Test and Evaluation Executive Agent Structure

3.5 THE T&E EXECUTIVE AGENT STRUCTURE

In 1993, the USD(AT&L) approved a T&E Executive Agent structure to provide the Services with more corporate responsibility for the management and policies that influence the availability of test resources for the evaluation of DoD systems in acquisition (Figure 3-5). The DT&E/DS has functional responsibility for the execution of the processes necessary to assure the T&E Executive Agent structure functions effectively. The DT&E/DS also participates in the Operational Test and Evaluation Coordinating Committee, chaired by the DOT&E. This committee manages the OT&E Resources Enhancement Project, and the DT&E/DS draws input to the T&E Executive Agent structure for coordination of all T&E resource requirements.

The Board of Directors (BoD) (Service Vice Chiefs) is assisted by an Executive Secretariat consisting of the Army DUSA(OR), the Navy N-91, and the Air Force AF/TE. The BoD provides guidance and decisions on policy and resource allocation to their subordinate element, the Board of Operating Directors (BoOD) (Army DTC, NAVAIR 5.0, and AFMC Director of Operations). The BoD also provides program review and advocacy support of the T&E infrastructure to OSD and Congress.

The BoOD is supported by a Secretariat and the Defense Test and Training Steering Group (DTTSG). The DTTSG manages the T&E Resources Committee (TERC), the Training Instrumentation Resource Investment Committee (TIRIC), and the CROSSBOW Committee. The DTTSG is instrumental in achieving efficient acquisition and integration of all training and associated test range instrumentation and the development of acquisition policy for embedded weapon system training and testing capabilities. The TERC supports the DTTSG in overseeing infrastructure requirements development from a T&E community perspective, both development testing and operational testing, and manages OSD funding in executing the Central T&E Investment Program (CTEIP). The TIRIC is chartered to ensure the efficient acquisition of common and interoperable range instrumentation systems. The CROSSBOW Committee provides technical and management oversight of the Services' development and acquisition programs for threat and threat-related hardware simulators, emitters, software simulations, hybrid representations, and surrogates.

3.6 SUMMARY

An increased emphasis on T&E has placed greater demands on the OSD and DoD components to carefully structure organizations and resources to ensure maximum effectiveness. Renewed interest by Congress in testing as a way of assessing systems utility and effectiveness, the report by the President's Blue Ribbon Panel on Acquisition Management, and acquisition reform initiatives have resulted in major reorganizations within the Services. These policy changes and reorganizations will be ongoing for several years to improve the management of T&E resources in support of acquisition programs.

ENDNOTES

1. DoD Instruction (DoDI) 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003.
2. DoDD 5141.2, *Director of Operational Test and Evaluation (DOT&E)*, May 25, 2000.

4

PROGRAM OFFICE RESPONSIBILITIES FOR TEST AND EVALUATION

4.1 INTRODUCTION

In government acquisition programs, there should be an element dedicated to the management of Test and Evaluation (T&E). This element has the overall test program responsibility for all phases of the acquisition process. T&E expertise may be available through matrix support or reside in the Program Management Office (PMO) engineering department during the program's early phases. By System Demonstration, the PMO should have a dedicated T&E manager. In the PMO, the Deputy for 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 test resources, facilities, and their support required for each phase of testing. In addition, the Deputy for T&E or a staff member, will be responsible for managing the Test and Evaluation Master Plan (TEMP), and planning and managing any 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, the Deputy for T&E will be responsible for supporting production acceptance testing and the test portions of later increments, Preplanned Product Improvements (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 T&E. 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 T&E function is often handled by matrix support from the materiel command. Matrix T&E support or the Deputy for T&E should be responsible for development of the T&E 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: 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), portions of the Acquisition Plan, Information for Proposal Preparation (IFPP), and, if a joint acquisition program, feedback on the Capability Development Document (CDD) integrating multiple Services' requirements.

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 (MOAs/

MOUs) between Services, North Atlantic Treaty Organization (NATO) countries, test organizations, etc., which outline the responsibilities of each organization. The RFP and SOW outline contractor/government obligations and arrangements on the access and use of test facilities (contractor or 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. The Deputy for T&E must have access to all contractor testing and test results, and the Deputy for T&E is responsible for disseminating the results to government agencies that need this data. Additionally, the Deputy for T&E 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*[1] 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. The Deputy for T&E ensures that the 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 test performance to ensure that the office of the Deputy for T&E has time to approve the procedures or implement modifications.

The Deputy for T&E must receive the test results and reports on time to enable the office of the Deputy for T&E, the PM, and higher authorities to make program decisions. The data received should be tailored to provide the minimum information needed. The Deputy for T&E must be aware that data requirements in excess of the minimum needed may lead to an unnecessary increase in overall program cost. For data that are 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).

The contract must specify the data the contractor will supply the Operational Test Agency (OTA). Unlike Development Test and Evaluation (DT&E), the contractor will not be making the Operational Test and Evaluation (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 its own test plans and procedures for OT&E. In fact, OTA representatives should attend the 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 stay informed of all OTA activities, understand the test procedures, and plan and receive the test reports. Unlike DT&E and contractor documentation, the PMO Deputy for T&E will not have report or document approval authority for OT&E items. The Deputy for T&E is responsible for keeping the PM informed of OT&E results.

4.3.3 Test Schedule Formulation

An 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, mock-ups, full-scale models, etc., once the contract is awarded. The Deputy for T&E uses this information and the early test planning in the Technology Development Strategy (TDS) document 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 draft integrated schedule (Part II, TEMP) 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 ensure that the test schedule is not so success-oriented that retesting of failures will cause serious program delays for either the government test agencies or the contractor.

Another important early activity to be performed by the Deputy for T&E 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 straw-man schedule (based on previous experience) and present this schedule to the operational test representative at the initial T&E Integrated Product Team (IPT) meeting for review, or the Deputy for T&E 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 PM. The test schedule must allow time for DT&E testing and OT&E testing when testing is not combined or test assets are not limited. Before setup 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 overrun so that the IOT&E test schedule is adversely impacted, reducing program schedule time with inadequate operational testing or rushing the reporting of IOT&E results. For example, if the DT&E schedule slips 6 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.

4.3.4 Programmatic Environmental 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 (CFR), Parts 1500-1508; National Environmental Policy Act (NEPA) Regulations; Executive Order 12114, *Environmental Effects Abroad of Major Federal Actions*; DoD 5000 Series; etc.) require analysis of hazardous materials and appropriate mitigation measures during each acquisition phase. As stated in DoD 5000.2-R, "Planning shall consider the potential testing impacts on the environment."

Litigation resulting in personal fines and imprisonment against government employees have raised the environmental awareness at test ranges and facilities. Environmental Impact Statements (EISs) (supported by long, thorough studies and public testimony) or Environmental Analysis and Assessments are generally required before any system testing can be initiated.

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, OTAs, 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 process. The Deputy for T&E also ensures the RFP includes provisions for government attendance at contractors' 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, the Deputy for T&E 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 T&E input.

4.5 INTEGRATED PRODUCT TEAMS FOR TEST AND EVALUATION

Before the final version of the RFP is created, the Deputy for T&E should form an IPT, a test planning/integration working group. This group includes the OTA, 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. 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. 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. The Deputy for T&E 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 the 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 the test program requirements; and the test organization sends the program office an estimate of the test program costs. The Deputy for T&E then obtains cost estimates from all test sources that the Deputy for T&E 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 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, the Deputy for T&E plans, directs, or monitors government testing; however, in the reviews and audits, the Deputy examines the contractor's approach to the test problem and evaluates the validity of the process and the accuracy of the contractor's results. The Deputy for T&E uses personal experience and background in T&E to assess 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, the Deputy must inform the contractor, the PM, and the PCO to validate the 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 all observations for future reference.

4.8 CONTRACTOR TESTING

The Deputy for T&E is responsible for ensuring that contractor-conducted tests are monitored by the government. The Deputy for T&E must also be given access to all contractor internal data, test results, and test reports related to the 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 (DT&E/OT&E) 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 system performance 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 (Section 3) 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. The Deputy for T&E 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 (Section 4) shape the DT&E testing scenario, test ranges, test support, targets, etc., and are very important to the Deputy for T&E.

4.10 INDEPENDENT TEST AND 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, the Deputy for T&E must depend on those DT or OT 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 (DT&E) 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 (DT/OT) 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. The Army is the only Service to have a designated independent evaluation agency, which provides feedback to the program office.

4.11 PMO RESPONSIBILITIES FOR OPERATIONAL TEST AND EVALUATION

In the government PMO, there should be a section responsible for T&E. Besides being responsible

- 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: Matt Reynolds, NAVSEA.

Figure 4-1. Lessons Learned from OT&E for the PM

for DT&E support to the PM, this section should be responsible for program coordination with the OT&E agency (Figure 4-1). The offices of the systems engineer or the Deputy for T&E may be designated to provide this support to the PM. In some Services, responsibilities of the Deputy for T&E include coordination of test resources for all phases of OT&E.

4.11.1 Contract Responsibilities

The Deputy for T&E or a T&E representative ensures that certain sections of the 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 (FRP)), 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 OTAs, 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.

4.11.2 Data Requirements

The contract must specify the data the contractor will supply the 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 its own test plans and procedures for OT&E. In fact, OTA representatives should attend the 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 stay informed of all OTA activities, understand the test procedures and plans, and receive the test reports. Unlike DT&E, the PMO Deputy for T&E will not have report or document approval authority as does the Deputy for T&E over contractor documentation. The Deputy for T&E is always responsible for keeping the PM informed of OT&E results.

4.11.3 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 draft schedule (based on previous experience) and present this schedule to the operational test representative at the initial Test Planning Working Group (TPWG) for review; or the Deputy for T&E 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 PM. 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 setup of IOT&E, the Certification of Readiness for IOT&E process 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.

4.11.4 Contractor Support

The Deputy for T&E provides all T&E input to the RFP/SOW. The Deputy for T&E must determine, before the beginning of the program acquisition phase, whether the contractor will be involved in supporting IOT&E and, if so, to what extent. According to Title 10, United States Code (U.S.C.), 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 IOT&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).

4.11.5 Statement of Work

One of the most important documents receiving input from the Deputy for T&E is the SOW. The Deputy for T&E 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.

4.11.6 OT&E Funding

The Deputy for T&E provides the PM estimates of PMO test program costs to conduct OT&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 PM.

4.11.7 Test and Evaluation Master Plan

The Deputy for T&E is responsible for managing the TEMP throughout the test program. The OTA usually is tasked to complete the OT section of the TEMP and outline its proposed test program through all phases of OT&E. The resources required for OT&E should also be entered in Part V. It is important to keep the TEMP updated regularly so that test organizations involved in OT&E understand the scope of their test support. The TEMP should be updated regularly by the OTA. 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 DT and OT requirements.

4.11.8 Program Management Office Support for OT&E

Even though operational testing is performed by an independent organization, the PM plays an important role in its planning, reporting, and funding. The PM must coordinate program activities with the test community in the TE Working-level Integrated Product Team (WIPT), especially the OTAs. The PM 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 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 the Office of the Secretary of Defense 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. The PMO must assure that contractor and government DT&E supports the decision to certify the system's readiness for IOT&E.

4.11.9 Support for IOT&E

The Certification of Readiness for the IOT&E process provides a forum for a final review of test results and support required by the IOT&E. The Deputy for T&E must ensure the contract portions adequately cover the scope of testing as outlined by the OTA. 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; the Deputy for T&E will keep the PM informed of progress and problems that arise during testing and will monitor required PMO support to the test organization. Sufficient LRIP items must be manufactured to run a complete and adequate OT&E program. The DOT&E determines the number of LRIP items for IOT&E on oversight programs, and the OTA makes this determination for all others.[2] 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 FRP decision can be made. As in the operational assessments, the Deputy for T&E will be monitoring test procedures and results and keeping the PM 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.

4.11.10 FOT&E and Modifications, Upgrades, Enhancements, or Additions

During FOT&E, the Deputy for T&E monitors the testing, and the level of contractor involvement is usually negotiated. 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 or in future increments. If the PM 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.

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 Non-Developmental 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 re-coordinated with the test organizations. The design of the DT&E and FOT&E program usually requires coordination with the engineering, contracting, and program management IPTs of the program office.

4.11.11 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 PM programs funding for support of OT. Funding for OT&E is identified in the TEMP and funded in the PMO's budget. The Services' policy specifies how the OTAs develop and manage their own budgets for operational testing.

4.12 SUMMARY

Staffing requirements in the PMO vary with the program phase and the T&E workload. T&E 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 prototype testing. 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.

The PMO should be proactive in its relations with the Service OTA. 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 surprises downstream. Operational testing is an essential component of the system 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.

ENDNOTES

1. DoD 5010.12-L, *Acquisition Management Systems and Data Requirements Control List (AMSDL)*, April 2001.
2. DoD Instruction (DoDI) 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003.

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 Capability Development Document (CDD) and Capability Production Document (CPD); program decision documentation such as the Acquisition Decision Memorandum (ADM) with exit criteria; and program management documentation such as the Acquisition Strategy, Baseline documentation, the Systems Engineering Plan (SEP), the logistics support planning, and the Test and Evaluation Master Plan (TEMP). Of importance to Program Managers (PMs) 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 (OTPs/TPOs), evaluation plans, and test reports. This chapter concludes with a description of the End-of-Test Phase and Beyond Low Rate Initial Production (BLRIP) Reports, and 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 Capability Analyses

The Joint Chiefs, via the Joint Requirements Oversight Council (JROC), are required to conduct continuing mission analyses of their assigned areas of responsibility.[1] These Functional Analyses (Area, Needs, Solutions) may result in recommendations to initiate new acquisition programs to reduce or eliminate operational deficiencies. If a need cannot be met through changes in Doctrine, Organization, Training, Leadership and Education, Personnel and Facilities (DOTLPF), and a Materiel (M) solution is required, the needed capability is described first in an Initial Capabilities Document (ICD) and then in the CDD. When the cost of a proposed acquisition program is estimated to exceed limits specified in *Operation of the Defense Acquisition System*, the proposed program considered a Major Defense Acquisition Program (MDAP) and requires JROC approval (JROC Interest).[2] Lesser programs for Service action are designated Joint Integration or Independent programs. The ICD is completed at the beginning of the analyses for a mission area; the CDD and CPD are program-/increment-specific and reviewed to evaluate necessary system modifications periodically.

5.2.2 Initial Capabilities Document

The ICD makes the case to establish the need for a material approach to resolve a specific capability gap. The ICD supports the Analysis of Alternatives (AoA), development of the Technology Development Strategy (TDS), and various milestone decisions. The ICD defines the capability gap in terms of the functional area(s), relevant range of military operations, time, obstacles to overcome, and key attributes with appropriate Measures of Effectiveness (MOEs). Once approved, the ICD is not normally updated. Format for the ICD is found in CJCSM [Chairman of the Joint Chiefs of Staff Manual] 3170.01A.[3]

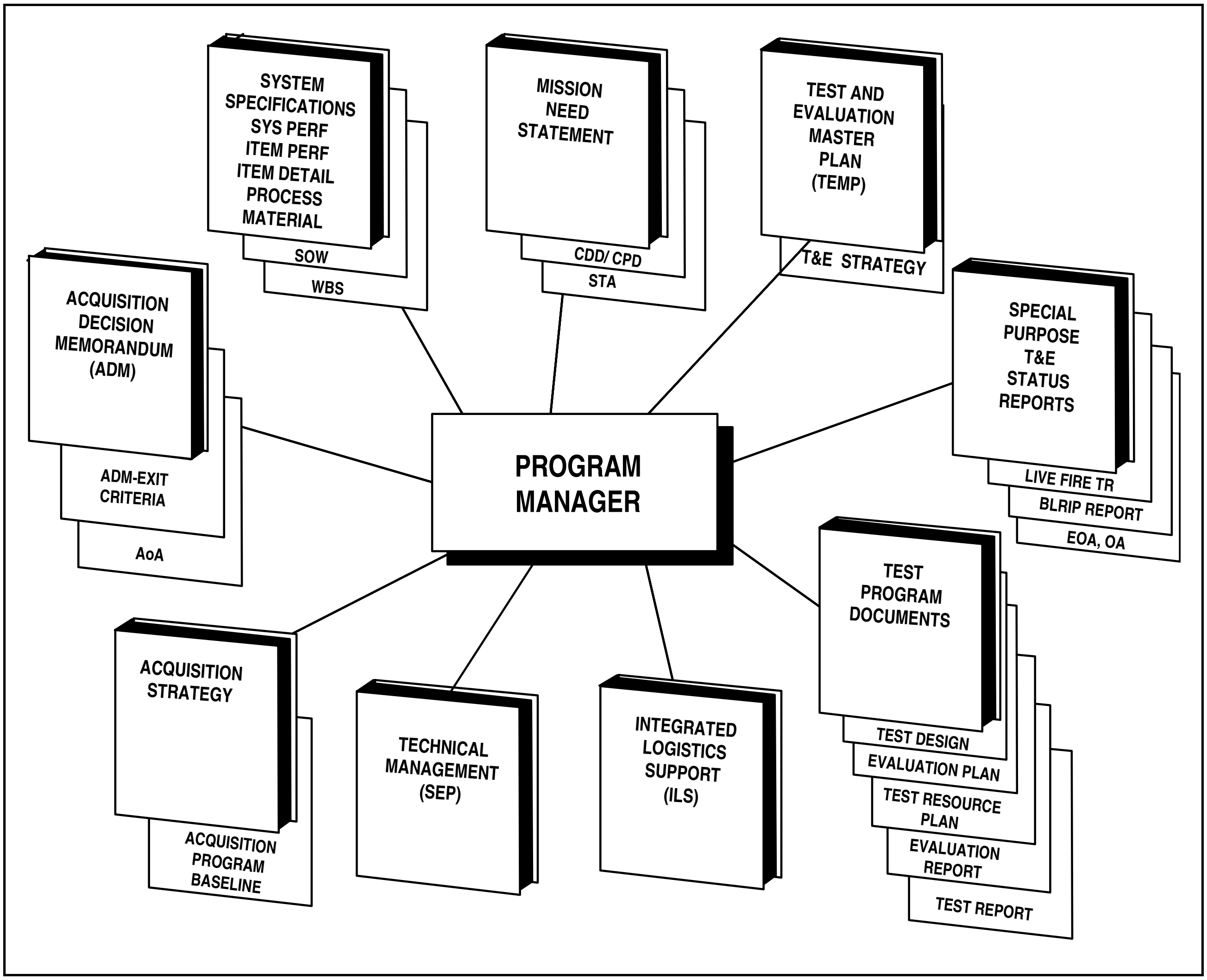


Figure 5-1. Test-Related Documentation

5.2.3 Capability Development Document

The CDD outlines an affordable increment of militarily useful and supportable operational capability that can be effectively developed, produced or acquired, deployed, and sustained. Each increment of capability will have its own set of attributes and associated performance values with thresholds and objectives established by the sponsor (Service) with input from the user. The CDD performance attributes apply to only one increment and include Key Performance Parameters (KPPs) and other parameters that will guide the development, demonstration, and testing of the current increment. The CDD will outline the overall strategy to develop the full or complete capability by describing all increments (Figure 5-2). The CDD will be approved at Milestone B and not normally updated. The format is found in CJCSM 3170.01A.[4]

5.2.4 Capability Production Document

The CPD addresses the production attributes and quantities specific to a single increment and is finalized by the sponsor after the Design Readiness Review (DRR), when projected capabilities of the increment in development have been specified with sufficient accuracy to begin production. Design trades during System Development and Demonstration (SDD) should have led to the final production design needed for Milestone C.

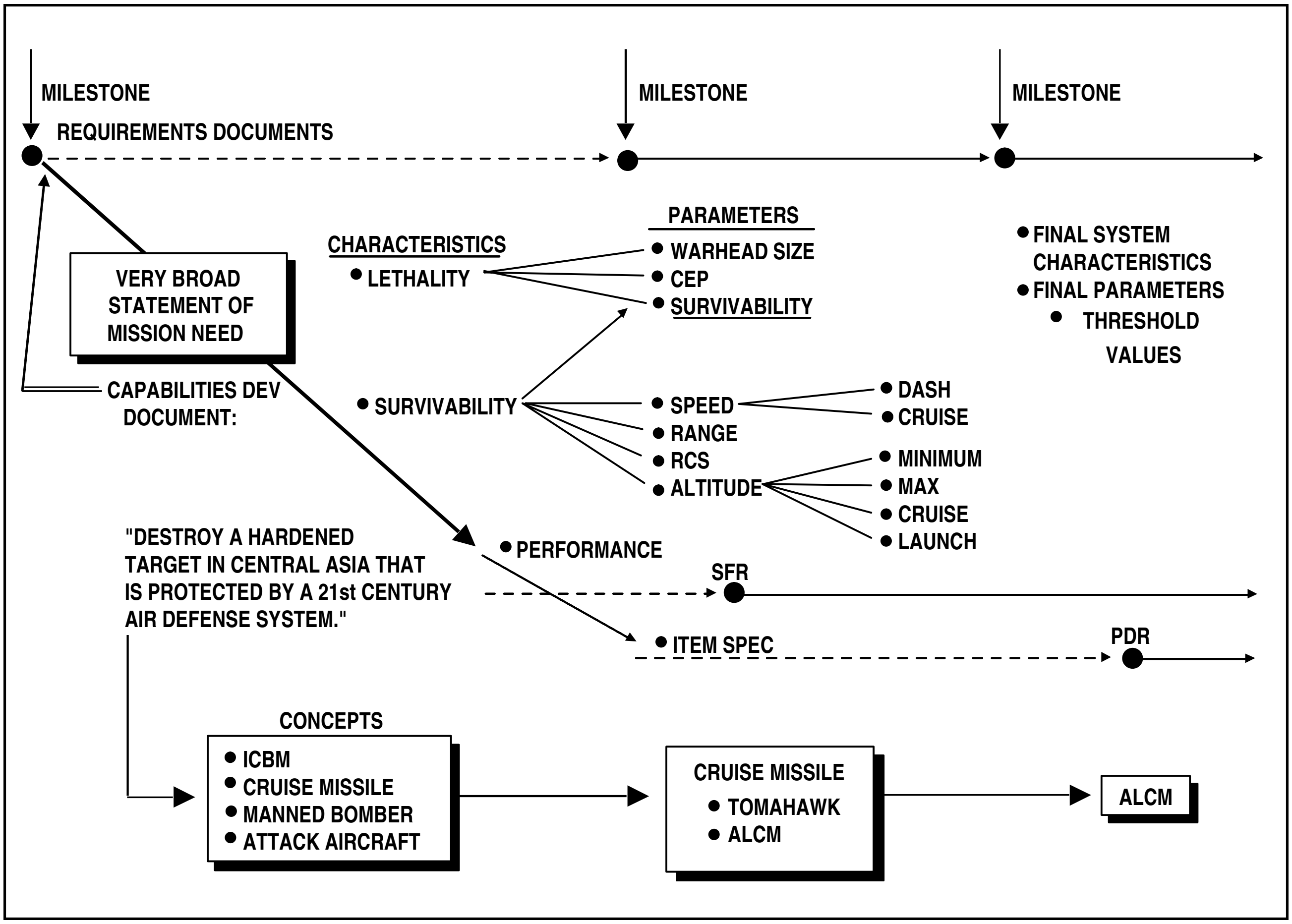


Figure 5-2. Requirements Definition Process

Initial Operational Test and Evaluation (IOT&E) will normally test to the values in the CPD. The threshold and objective values from the CDD are, therefore, superseded by the specific production values detailed in the CPD. Reduction in any KPP threshold value will require reassessment of the military utility of the reduced capability. The CPD will always reference the originating ICD. The format is found in CJCSM 3170.01A.[5]

5.2.5 System Threat Assessment (STA)

An STA is prepared, starting at Milestone B, by the Department of Defense (DoD) Component Intelligence Command or Agency, and for Acquisition Category (ACAT) ID programs, and is validated by the Defense Intelligence Agency (DIA).[6] The STA, for Defense Acquisition Board (DAB) programs, 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 PM when evaluating the program against the threat. Threat projections start at the Initial Operational Capability (IOC) and extend over the following 10 years. The STA provides the basis for the test design of threat scenarios and the acquisition of appropriate threat targets, 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 (LFT) of ACAT I and II systems are contingent on valid threat descriptions. A summary of the STA is included in Part 1 of the TEMP.

5.3 PROGRAM DECISION DOCUMENTATION

5.3.1 Acquisition Decision Memorandum

Under Secretary of Defense for Acquisition, Technology, and Logistics (USD(AT&L)) decisions at major defense ACAT ID milestones are recorded in a document known as an ADM. The ADM documents a USD(AT&L) decision on program progress and on the Acquisition Program Baseline (APB) at milestones and decision reviews. In conjunction with an ADM, and its included exit criteria for the next phase, the APB is a primary program guidance document providing KPP thresholds/objectives for systems cost, schedule, and performance.

5.3.2 Analysis of Alternatives

An AoA is normally prepared by a DoD Component agency (or Principal Staff Assistant (PSA) for ACAT IA programs) other than the Program Management Office (PMO) for each milestone review beginning at Milestone A. The AoA aids decision makers by examining the relative advantages and disadvantages of program alternatives, shows the sensitivity of each alternative to possible changes in key assumptions, and provides the rationale for each option. The guidance in the *Defense Acquisition Guidebook*, Chapter 4, requires a clear linkage between the AoA, system requirements, and system evaluation MOE.

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 validate AoA 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. 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 milestone and decision review using a Working-level Integrated Product Team (WIPT) structure 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 tailored conceptual basis for formulating other program functional plans such as the TEMP.

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 APB will initially be developed before Milestone B or program initiation and revised for each subsequent milestone. 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. MOEs or Measures of Performance (MOPs) shall be used in describing needed capabilities early in a program. Guidance on the formulation of baselines is found in the *Defense Acquisition Guidebook*.[7] 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) for KPPs beyond which the PM may not trade off cost, schedule, or performance without authorization by the Milestone Decision Authority (MDA). Baseline and test documentation must reflect the same expectations for system performance. The total number of performance parameters shall be the minimum number needed to characterize the major drivers of operational effectiveness and suitability, schedule, technical progress, and cost for a production system intended for deployment. The performance parameters may not completely define operational effectiveness or suitability.

5.4.3 Acquisition Logistics Planning

Supportability analyses are 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. Support concepts describe the overall logistics support program and include logistics requirements, tasks, and milestones for the current and succeeding phases of the program. The analyses serve as the source document for logistics support testing requirements.

Guidelines for logistics support analyses are documented in *Logistics Support Analysis*.[8] This standard identifies how T&E programs should 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 systems engineering disciplines, especially those involved in logistics support analyses. Areas of particular interest include Reliability, Availability, and Maintainability (RAM), Human Systems Integration (HSI), Environmental Safety and Occupational Health (ESOH), and post-deployment evaluations. The support analyses should be drafted shortly before program start to provide a skeletal framework for Logistics Support Analysis (LSA), to identify initial logistics testing requirements that can be used as input to the TEMP, and to provide test feedback to support Integrated Logistics Support (ILS) development. Test resources will be limited early in the program.

5.4.4 Specification

The system specification is a document used in development and procurement that describes the technical performance requirements for items, materials, and services including the procedures by which it will be determined that the requirements have been met. Specification evolves over the developmental phases of the program with increasing levels of detail: system, item performance, item detail, process, and material. Section 4 of the specification identifies what procedures (inspection, demonstration, analysis, and test) will be used to verify the performance parameters listed in Section 3. Further details may be found in Military Standard (MIL-STD)-961D, *DoD Standard Practice for Defense Specifications*.[9]

5.4.5 Work Breakdown Structure (WBS)

A program WBS shall be established that provides a framework for program and technical planning, cost estimating, resource allocations, performance measurements, and status reporting. Program offices shall tailor a program WBS for each program using the guidance in Military Handbook *Work Breakdown Structure*.[10] Level 2 of the WBS hierarchical structure addresses system-level T&E with sublevels for DT&E and OT&E. Additionally, each configuration item structure includes details of the integration and test requirements.

5.5 TEST PROGRAM DOCUMENTATION

5.5.1 Test and Evaluation Master Plan

An evaluation strategy is developed by the early concept team that describes how the capabilities in the CDD will be evaluated once the system is developed. The evaluation strategy, part of the TDS, provides the foundation for development of the program TEMP at the milestone supporting program start. The TEMP is the basic planning document for T&E related to a DoD system acquisition (Figure 5-3). It is prepared by the PMO with the OT information provided by the Service Operational Test Agency (OTA). 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 Parameters (CTPs), performance characteristics (MOE), and Critical Operational Issues (COIs); and describes the objectives, responsibilities, resources, and schedules for all completed and planned T&E. The TEMP, in the *Defense Acquisition Guidebook* specified format, is required by DoD Instruction (DoDI) 5000.2 for ACAT I, IA, and designated oversight programs (see enclosure 5 to DoDI 5000.2 for more information regarding the TEMP). Format is at Service discretion for ACAT II and III programs. For example, the Army requires that each TEMP contain the full set of approved Critical Operational

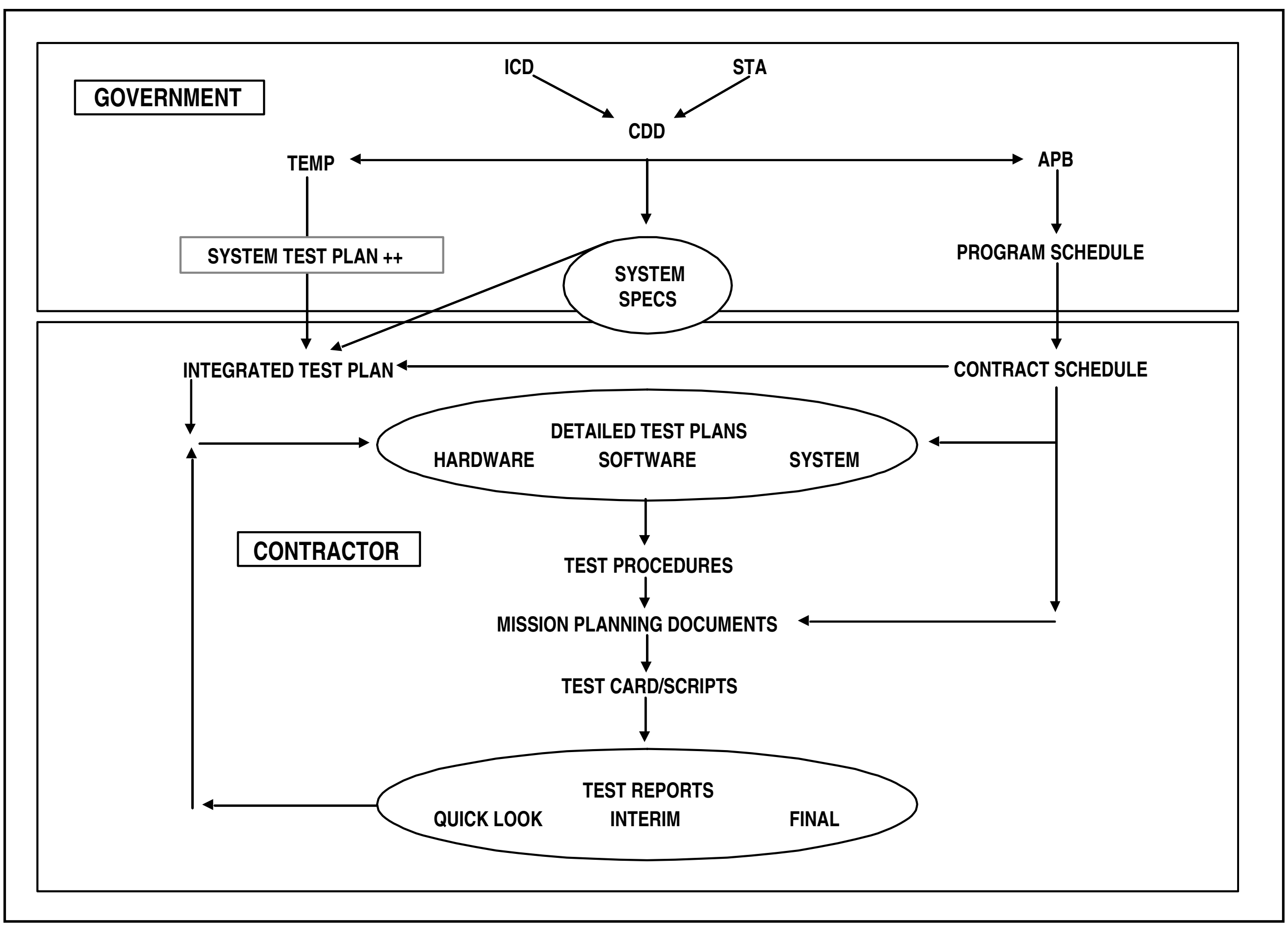


Figure 5-3. Test Program Documentation

Issues and Criteria (COIC), with associated scope and rationale, forming the basis for planning the system's evaluations.[11]

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 have 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 a System Evaluation Plan (SEP) describing the evaluation approach that addresses the technical and operational aspects necessary to address the system's operational effectiveness, suitability, and survivability.

The objective of the Army's emphasis on evaluation is to address the issues; describe the evaluation of issues that require data from sources other than test; state the technical or operational issues and criteria; identify data sources (data source matrix); state the approach to the independent evaluation; and specify the analytical plan and identify program constraints.[12]

Evaluation plans are prepared for all systems in development by the independent evaluators during concept exploration and in coordination with the system developer. The Army SEP complements the TEMP and is developed in support of the initial TEMP. 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 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? 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 MOEs and 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 (M&S) 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 multi-volume 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.

PRELIMINARY PAGES

- i. TITLE PAGE
- ii. APPROVAL PAGE
- iii. CONDITIONS FOR RELEASE OF TECHNICAL DATA
- iv. PREFACE (ABSTRACT)
- v. EXECUTIVE SUMMARY
- vi. TABLE OF CONTENTS*

*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. BACKGROUND
3. TEST ITEM DESCRIPTION
4. OVERALL TEST OBJECTIVES
5. CONSTRAINTS AND LIMITATIONS
6. TEST RESOURCES
7. SAFETY REQUIREMENTS
8. SECURITY REQUIREMENTS
9. TEST PROJECT MANAGEMENT
10. TEST PROCEDURES
11. TEST REPORTING
12. LOGISTICS
13. ENVIRONMENTAL PROTECTION

APPENDICES

- A. TEST CONDITION MATRIX
- B. REQUIREMENTS TRACEABILITY
- C. TEST INFORMATION SHEETS
- D. PARAMETERS LIST
- E. DATA ANALYSIS PLAN
- F. INSTRUMENTATION PLAN
- G. LOGISTICS SUPPORT PLAN
- H-Z AS REQUIRED

LIST OF ABBREVIATIONS

DISTRIBUTION

SOURCE: *AFFTC Test Plan Preparation Guide*, May 1994.

Table 5-1. Sample Test Plan Contents

5.5.5 Outline Test Plan/Resources Plan

The Army's Outline Test Plan (OTP) and Air Force's Test Resources Plan (TRP) are essential test planning documents. They are formal resource documents specifying the resources required to support the test. Since the OTP or TRP provide 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 database. 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 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) M&S

(9) RAM 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 M&S to generalize the results and extrapolate to other conditions. In the case of the Army, a separate System Evaluation Report is prepared by independent evaluators within the Army Evaluation Center (AEC).

5.6 OTHER TEST-RELATED STATUS REPORTS

5.6.1 End of Test Reports

The Services are required by DoDI 5000.2, *Operation of the Defense Acquisition System* to submit to OSD T&E offices copies of their formal DT&E, OT&E, and Live Fire Test and Evaluation (LFT&E) reports that are prepared at the end of each period of testing for ACAT I, IA, and oversight programs.[13] These reports will generally be submitted in a timely manner to permit OSD review.

5.6.2 Beyond Low-Rate Initial Production (BLRIP) Report

Before an ACAT I or Director, Operational Test and Evaluation (DOT&E)-designated program can proceed beyond Low Rate Initial Production (LRIP), the DOT&E must submit a BLRIP report to the Secretary of Defense and the Senate and House of Representatives Committees on Armed Services, National Security, and Appropriations. This report addresses whether the OT&E performed was adequate and whether the IOT&E results confirm that the items or components tested are effective and suitable for use in combat by typical military users. The report may include information on the results of LFT&E for applicable major systems.

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 PM must work to ensure that T&E requirements are considered at the outset when the acquisition strategy is formulated. The PM must also require early, close coordination and a continuing dialogue among those responsible for integration of functional area planning and the TEMP.

ENDNOTES

1. CJCS Instruction (CJCSI) 3170.013, *Joint Capabilities Integration and Development System* (JCIDS), March 12, 2004; and CJCSM 3170.01A, *Operation of the Joint Capabilities Integration and Development System*, March 12, 2004.

2. DoDI 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003.

3. CJCSM 3170.01A, Appendix A, enclosure D.

4. CJCSM 3170.01A, Appendix A, enclosure E.

5. CJCSM 3170.01A, Appendix A, enclosure F.

6. DoD Directive (DoDD) 5105.21, *Defense Intelligence Agency*, May 19, 1977. Cancelled.

7. *Defense Acquisition Handbook*, http://akss.dau.mil/DAG, Chapter 9.

8. MIL-STD-1388-1A, *Logistics Support Analysis*, April 11, 1983.

9. MIL-STD-961D, *DoD Standard Practice for Defense Specifications*, March 22, 1995. Military Defense Specification Standard Practices incorporated portions of MIL-STD-490, *Specification Practices*, which is fully exempt from the MIL-STD-491D waiver process because it is a "Standard Practice."

10. Military Handbook (MIL-HDBK)-881, *Work Breakdown Structure*, January 2, 1998.

11. Army Regulation (AR) 73-1, *Test and Evaluation Policy*, January 7, 2002.

12. *Ibid.*

13. DoDI 5000.2, enclosure 5.

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 sub-element of development testing. Other important types of T&E are introduced. They include: multi-Service testing; joint T&E; Live Fire Testing (LFT); Nuclear, Biological, and Chemical (NBC) testing; and Nuclear Hardness and Survivability (NH&S) 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 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, test beds, and prototypes or full-scale engineering development models of the system. DT&E 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.

The DT&E may support the system design process through an iterative Modeling and Simulation (M&S), such as Simulation, Test, and Evaluation Process (STEP) that involves both contractor and government personnel. Because contractor testing plays a pivotal role in the total test program, it is important that the contractor establish an integrated test plan early to ensure that the scope of the contractor's test program satisfies government and contractor test objectives.

Anti-tamper component-level verification testing shall take place prior to production as a function of Development Test/Operational Test (DT/OT). Component-level testing shall not assess the strength of the anti-tamper mechanism provided, but instead verify that anti-tamper performs as specified by the source contractor or government agency.[1]

The Program Manager (PM) remains responsible for the ultimate success of the overall program. The PM and the test specialists on the PM's 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 Test (PQT)

Qualification testing is a form of development testing that verifies the design and manufacturing process. PQTs are formal contractual tests that

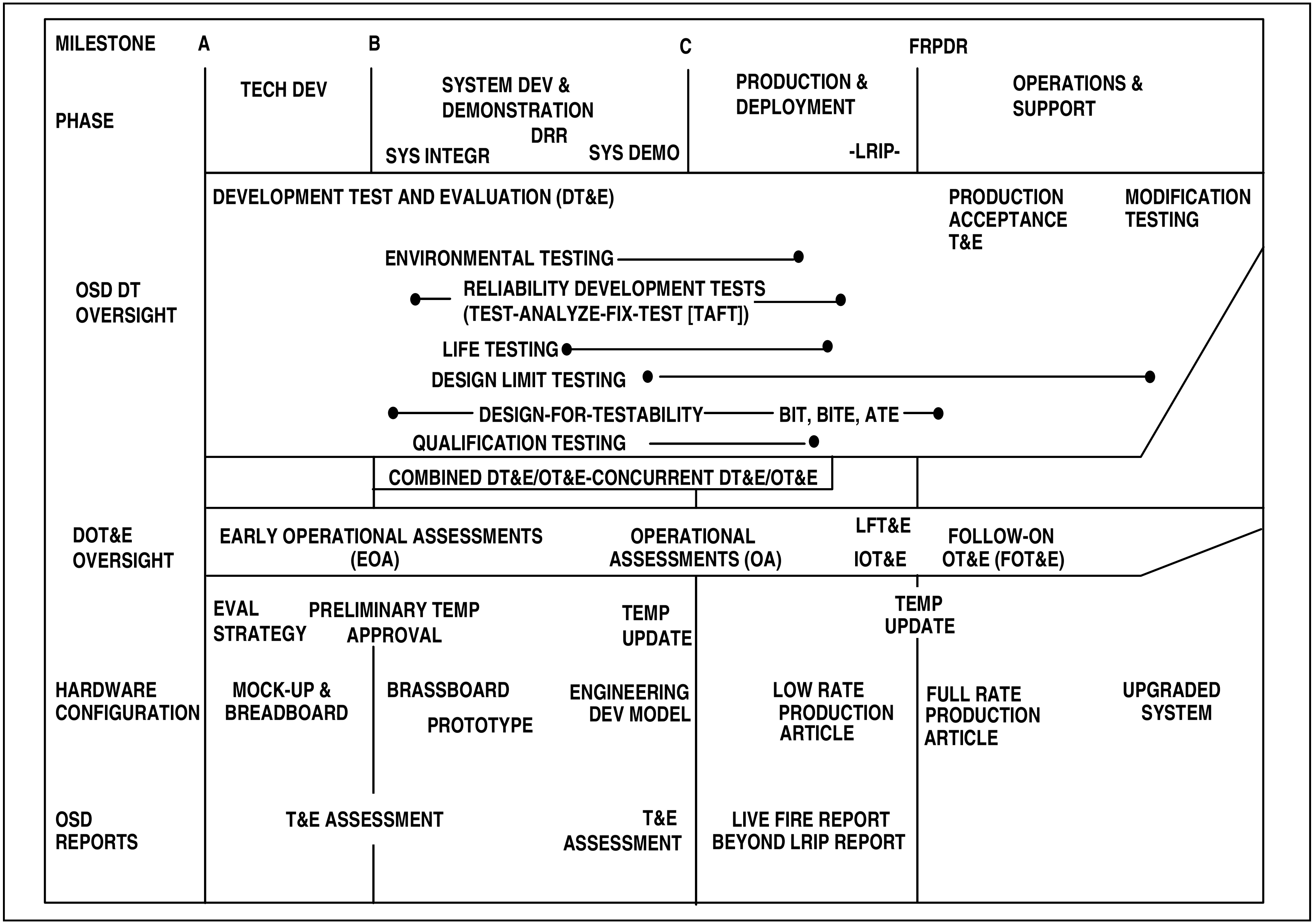


Figure 6-1. Testing During Acquisition

confirm the integrity of the system design over the operational and environmental range in the specification. These tests usually use production hardware fabricated to the proposed production design specifications and drawings. Such tests include contractual Reliability and Maintainability (R&M) demonstration tests required before production release. PQT must be completed before Full Rate Production (FRP) in accordance with *Operation of the Defense Acquisition System*.[2]

PQTs may be conducted on Low Rate Initial Production (LRIP) items to ensure the maturity 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 online. These tests are also conducted against contractual design and performance requirements.

6.3 OPERATIONAL TEST AND EVALUATION

6.3.1 The Difference Between Development and Operational Testing

The 1979 *Management of Operational Test and Evaluation*, 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 that 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.[3]

Captain Foulois' experience may seem humorous now, but it illustrates the need for operational testing. It also 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 OT focuses on the actual functioning of the equipment in a realistic combat environment in which the equipment must interact with humans and peripheral equipment. While DT&E and OT&E are separate activities and are

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

DT&E	IOT&E
• CONTROLLED BY PM • ONE-ON-ONE TESTS • CONTROLLED ENVIRONMENT • CONTRACTOR INVOLVEMENT • TRAINED, EXPERIENCED OPERATORS • PRECISE PERFORMANCE OBJECTIVES AND THRESHOLD MEASUREMENT • TEST TO SPECIFICATION • DEVELOPMENT TEST ARTICLE	• CONTROLLED BY INDEPENDENT AGENCY • MANY-ON-MANY TESTS • REALISTIC/TACTICAL ENVIRONMENT WITH OPERATIONAL SCENARIO • RESTRICTED SYSTEM CONTRACTOR INVOLVEMENT • USER TROOPS RECENTLY TRAINED ON EQUIPMENT • PERFORMANCE MEASUREMENT OF OPERATIONAL EFFECTIVENESS AND SUITABILITY • TEST TO REQUIREMENTS • PRODUCTION REPRESENTATIVE TEST ARTICLE

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 OT&E

Operational Test and Evaluation is defined in Title 10, United States Code (U.S.C.), Sections 139 and 2399:

> The field test, under realistic combat conditions, of any item of (or key component of) weapons, equipment, or munitions for the purposes 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. This term does not include an operational assessment based exclusively on computer modeling, simulation, or an analysis of system requirements, engineering proposals, design specifications, or any other information contained in program documents.[4]

Definitions of operational effectiveness and operational suitability are listed below:

Operational Effectiveness: Measure of the overall ability of a system to accomplish a mission when used by representative personnel in the environment planned or expected for operational employment of the system considering organization, doctrine, tactics, supportability, survivability, vulnerability, and threat.[5]

Operational Suitability: The degree to which a system can be placed and sustained satisfactorily in field use with consideration being given to availability, compatibility, transportability, interoperability, reliability, wartime usage rates, maintainability, safety, human factors, habitability, manpower, logistics supportability, natural environmental effects and impacts, documentation, and training requirements.[6]

In each of the Services, operational testing is conducted under the auspices of an organization that is independent of the development agency, in environments as operationally realistic 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. The major questions addressed in OT&E are shown in Figure 6-2.

Operational Assessment (EOA, OA): An evaluation of operational effectiveness and operational suitability made by an independent OT activity, with user support as required, on other than production systems. The focus of an OA is on significant trends noted in development efforts, programmatic voids, risk areas, adequacy of requirements, and the ability of the program to support adequate operational testing. An OA may be conducted at any time using technology demonstrators, prototypes, mock-ups, Engineering Development Models (EDMs), or simulators, but will not substitute for the Initial Operational T&E (IOT&E) necessary to support FRP decisions. Normally conducted prior to Milestone C.[7]

The OA normally takes place during each phase of the acquisition process prior to IOT&E. It is used to provide an early assessment of potential operational effectiveness and suitability for decision makers at decision points. These assessments attempt to project the system's potential to meet the user's requirements. Assessments conducted early in the program development process may be called Early Operational Assessments (EOAs). When using an evolutionary acquisition strategy, an OA is required for all increments following the one experiencing an IOT&E before the increment is released to the user.[8]

6.3.3 Initial Operational Test and Evaluation

The OT&E performed in support of the FRP decision is generally known as IOT&E. The Navy calls this event OPEVAL (Operational Evaluation). The IOT&E occurs during LRIP and must

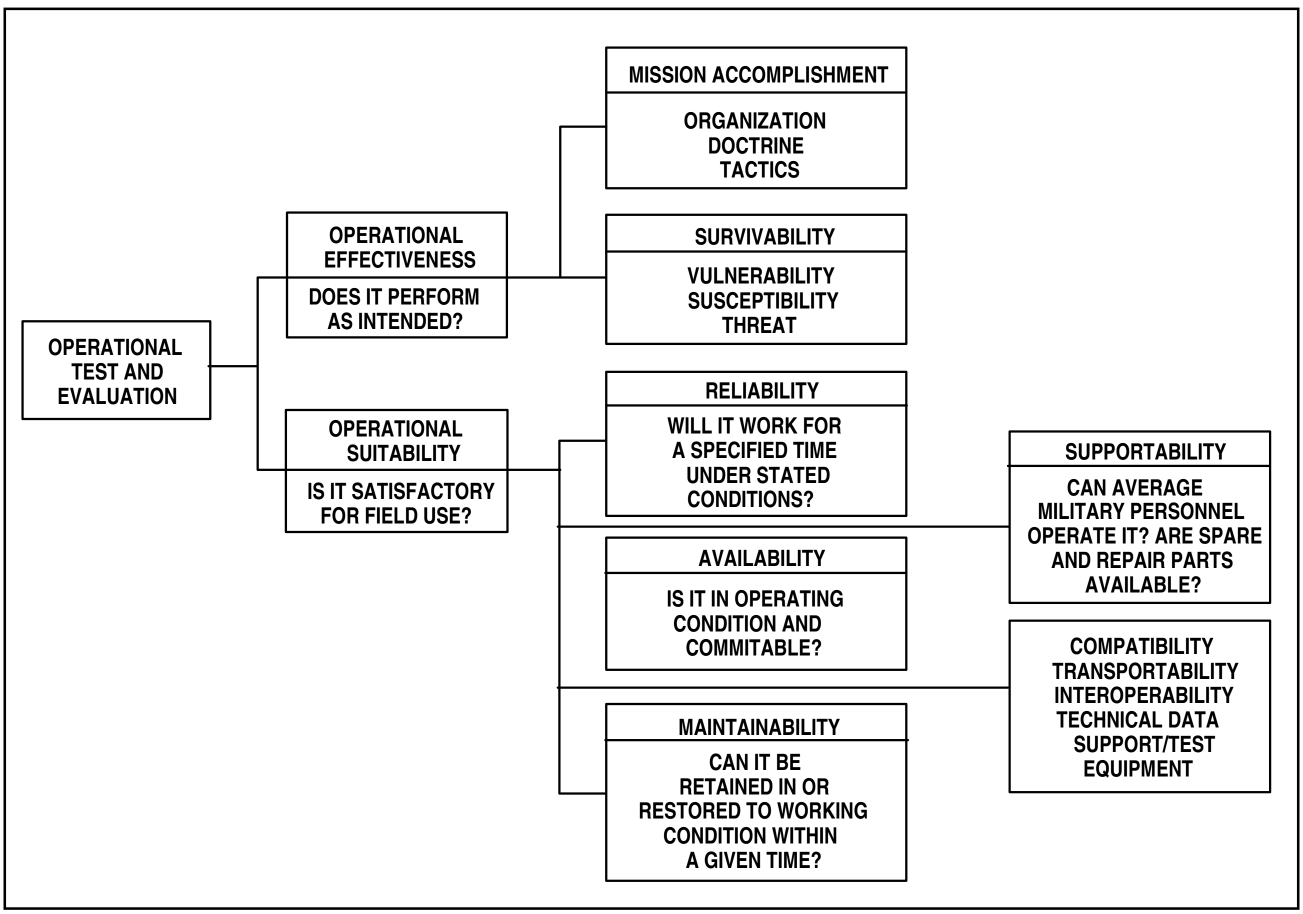


Figure 6-2. Sample Hierarchy of Questions Addressed in OT&E

be completed before the Full Rate Production Decision Review (FRPDR). More than one IOT&E may be conducted on the system if there are system performance problems requiring retest, the system is decertified, or a need exists to test in different environments. The IOT&E is conducted on a production or production representative system using typical operational personnel in a realistic combat scenario.

6.3.4 Follow-on Operational Test and Evaluation (FOT&E)

The OT&E that may be necessary after the FRPDR to refine the estimates made during IOT&E, to evaluate changes, and to reevaluate the system to ensure that it continues to meet operational needs and retains its effectiveness in a new environment or against a new threat.[9] FOT&E is conducted during fielding/deployment and operational support, and sometimes may be divided into two separate activities. Preliminary FOT&E is normally conducted after the Initial Operational Capability (IOC) is attained to assess full system capability. It is conducted by the OTA or designated organization to verify the correction of deficiencies, if required, and to assess system training and logistics status not evaluated during IOT&E. 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 later FOT&E often is conducted by the operating command.

6.4 MULTI-SERVICE TEST AND EVALUATION

Multi-Service Test and Evaluation is T&E conducted by two or more Department of Defense (DoD) Components for systems to be acquired by more than one DoD component, or for a DoD component's systems that have interfaces with equipment of another DoD component.[10] All affected Services and their respective Operational Test Agencies (OTAs) 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 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 joint acquisition program conducting multi-Service T&E 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 may prepare an Independent Evaluation Report (IER) in its own format and submits that report through its normal Service channels. The lead Service OT&E agency may prepare the documentation that goes forward to the Milestone Decision Authority (MDA). This documentation is coordinated with all participating OT&E agencies.

6.5 JOINT TEST AND EVALUATION

Joint Test and Evaluation (JT&E) is not the same as multi-Service T&E. JT&E is a specific program activity sponsored and largely funded by an Office of the Secretary of Defense (OSD) (Director, Operational Test and Evaluation (DOT&E)). JT&E programs are not acquisition-oriented but 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 Commands, and the

Services. The overarching policies and details of formulating a JT&E program are set forth in DoD Directive (DoDD) *Joint Test and Evaluation (JT&E) Program*.[11]

Responsibility for management of the JT&E program was transferred from the Under Secretary of Defense for Acquisition, Technology, and Logistics (USD(AT&L)) DT&E to the Office of the DOT&E in December 2002. Each year nominations are submitted by Combatant Commands, DoD agencies, and Services to the JT&E Program Office within DOT&E for review and processing. The JT&E Program Office identifies those to be funded for a feasibility study, the DOT&E assigns lead Service responsibility for a Quick Reaction Test (6 to 12 months) or a full JT&E (up to three years), and a Joint Test Force Office is formed to conduct JT&E activities consistent with the chartered objectives from the participating Services. The lead and participating Services provide personnel, infrastructure support, and test resources for JT&E activities.

The JT&E program's primary objectives include: assess Service system interoperability in joint operations; evaluate joint technical and operational concepts and recommend improvements; validate testing methodologies that have joint applications; improve M&S validity with field exercise data; increase joint mission capability, using quantitative data for analysis; provide feedback to the acquisition and joint operations communities; improve joint Tactics, Techniques, and Procedures (TTP).

Each JT&E typically produces one or more test products that provide continuing benefits to DoD, such as improvements in: joint warfighting capabilities; joint tactics, techniques, and procedures; joint and individual Service training programs; new testing methodologies; joint application of M&Ss that can support acquisition, TTP, and Concept of Operations (COOs) development; universal task list update and input.

6.6 LIVE FIRE TESTING

The Live Fire Test (LFT) Program was mandated by the Congress in the National Defense Authorization Act (NDAA) for Fiscal 1987 (Public Law 99-661) passed in November 1986. Specifically, this law stipulated that a major [Acquisition Category (ACAT) I and II] program development may not proceed Beyond Low Rate Initial Production (BLRIP) 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, Joint LFT, was initially divided into two distinct parts: Armor/Anti-armor 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 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 FRP 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 (LFT&E) Guidelines were first issued by the Deputy Director, T&E (Live Fire Testing) in May 1987 to supplement DoD Test and Evaluation Master Plan (TEMP) guidelines in areas pertaining to live fire testing.[12] These guidelines encompass all Major Defense Acquisition Programs (MDAPs) and define LFT requirements. In 1994 Public Law 103-355 directed that oversight of Live Fire Testing be moved within DoD to the DOT&E. Guidelines for this program are now found in DoD Instruction (DoDI) 5000.2 and the *Defense Acquisition Guidebook*.

6.7 NUCLEAR, BIOLOGICAL, AND CHEMICAL (NBC) WEAPONS TESTING

The testing of NBC weapons is highly specialized and regulated. PMs 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 (LTBT) that generally restricts nuclear detonations to the underground environment. Nuclear weapons testing requires extensive coordination between Service and DOE test personnel.[13]

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 1970s.[14]

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.[15]

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.[16]

In addition to the customary development tests (conducted to determine if a weapon meets technical specifications) and OTs (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.[17]

6.8 NUCLEAR HARDNESS AND SURVIVABILITY (NHS) 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. DoD policy requires the incorporation of NH&S 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.[18]

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

T&E is a spectrum of techniques used to address questions about critical performance parameters during system development. These questions may involve several issues including: technical and survivability (development testing); effectiveness and suitability (operational testing); those affecting more than one Service (multi-Service and joint testing); vulnerability and lethality (LFT), nuclear survivability; or the use of other than conventional weapons (i.e., NBC).

ENDNOTES

1. *Defense Acquisition Guidebook*, Chapter 3, Test and Evaluation. http://akss.dau.mil/DAG.

2. DoDI 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003.

3. Air Force Manual (AFM) 55-43, *Management of Operational Test and Evaluation,* June 1979.

4. Title 10, U.S.C., Section 139, *Director of Operational Test and Evaluation*, and Section 2399, *Operational Test and Evaluation of Defense Acquisition Programs*, March 18, 2004.

5. Chairman of the Joint Chiefs of Staff Instruction (CJCSI) 3170.01D, *Joint Capabilities Integration and Development System (JCIDS)*, March 12, 2004.

6. *Ibid.*

7. Defense Acquisition University Press, *Glossary of Defense Acquisition Acronyms & Terms,* September 2003.

8. DoDI 5000.2

9. Defense Acquisition University Press.

10. *Ibid.*

11. DoDD 5010.41, *Joint Test and Evaluation (JT&E) Program*, February 23, 1998.

12. 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.

13. DoDI 5030.55, *Joint AEC-DoD Nuclear Weapons Development Procedures,* January 25, 2001.

14. DoDD 5160.5, *Responsibilities for Research, Development, and Acquisition of Chemical Weapons and Chemical and Biological Defense,* May 1, 1985.

15. *Ibid.*

16. *Ibid.*

17. *Ibid.*

18. DoDI 4245.4, *Acquisition of Nuclear-Survivable Systems,* July 25, 1988. Cancelled.

II

MODULE

DEVELOPMENT 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. DT&E is performed in the factory, in the laboratory, and on the proving ground. It is conducted by subcontractors as they are developing the components and subassembly; the prime contractor, as the components are assembled and integration of the system is ensured; 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 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. DT&E 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 AND THE SYSTEM ACQUISITION CYCLE

As illustrated in Figure 7-1, DT&E is conducted throughout the system life cycle. DT&E may begin before program initiation with the evaluation of evolving technology, and it continues after the system is in the field.

7.2.1 DT&E Prior to Program Initiation

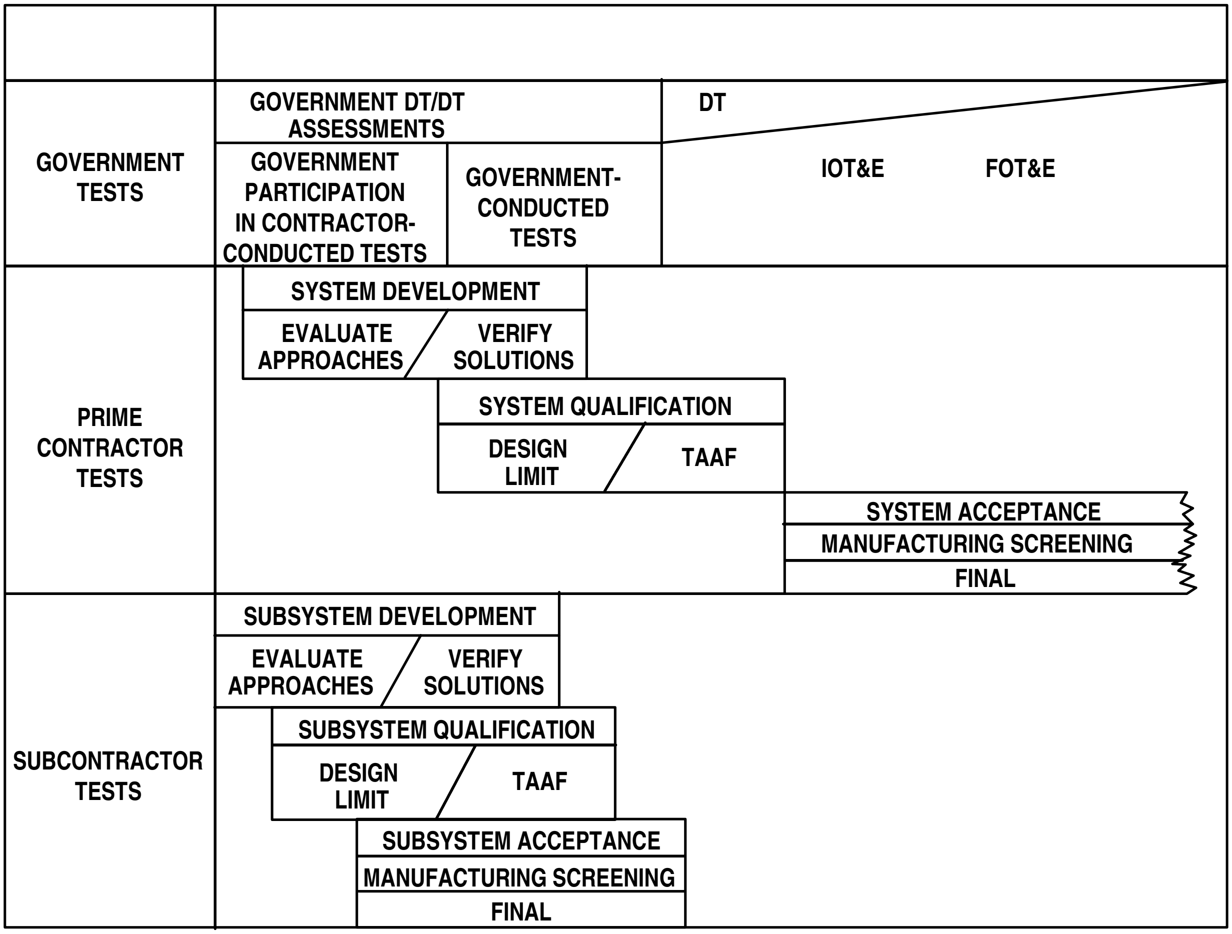
Prior to program initiation, modeling, simulations, and technology, feasibility testing is conducted to confirm that the technology considered for the proposed weapon development is the most advanced available and that it is technically feasible.

7.2.2 DT&E During Concept Refinement (CR) and Technology Development (TD)

Development testing that takes place 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 technical maturity of the test article's design. Government test evaluators participate in this testing because information obtained can be used to support a Systems Requirements Review (SRR), early test planning in the Technology Development Strategy (TDS), and development of the Capability Development Document (CDD) and the Request for Proposal (RFP) for Milestone B. 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).

7.2.3 DT&E During System Development and Demonstration (SDD)

Development testing is 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 decision



SOURCE: Adapted and modified from "Solving the Risk Equation in Transitioning from Development to Production." January 19, 1984.

Figure 7-1. Contractor's Integrated Test Requirements

review that considers transition from prototype design into advanced engineering and construction of the Engineering Development Model (EDM). This DT&E includes contractor/government integrated testing, engineering design testing, and advanced development verification testing.

Development testing during systems integration 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 system demonstration. Component interface problems and equipment performance capabilities are evaluated. The use of properly validated analysis, Modeling and Simulation (M&S) is encouraged, especially during the early phases to assess those areas that, for safety or testing capability limitations, cannot be observed directly through testing. M&Ss can provide early projections of systems performance, effectiveness, and suitability, and can reduce testing costs. This Test and Evaluation (T&E) 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 testing/technical testing. The early DT&E of the AAH was conducted by the Army Engineering Flight Activity (EFA). The test was conducted in conjunction with an Early Operational Assessment (EOA), 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 database. After evaluating the results, the Army selected a contractor to proceed with the next development phase of the AAH.

DT&E conducted during system demonstration provides the final technical data for determining a system's readiness to transition into Low Rate Initial Production (LRIP). It is conducted using advanced EDMs 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, RAM, compatibility, interoperability, safety, and supportability. They include tests of human engineering and technical aspects of the system. Demonstrations indicating that engineering is reasonably complete and that solutions to all significant design problems are in hand are also included.

7.2.4 DT&E During Production and Deployment (P&D)

7.2.4.1 Low Rate Initial Production

DT&E may be conducted on EDMs or LRIP articles as a prelude to certifying the system ready for Initial Operational Test and Evaluation (IOT&E). Each Service has different and specific processes incorporated in the certification for IOT&E documentation. The Navy conducts additional DT&E for certification called TECHEVAL (Technical Evaluation). This is a DT&E event controlled by the program office that is conducted in a more operationally realistic test environment. The Air Force has developed a guide with a structured process using templates to assist the Program Manager (PM) in assessing the program's readiness for IOT&E.

As an example of testing done during this 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, Arizona. 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 (OT) personnel were completing OT design, bringing together test resources, and observing the DT&E tests. Additionally, 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 (CDR) members to ensure that each configuration item design was essentially completed. The Army conducted a Physical Configuration Audit (PCA) to provide a technical examination to verify that each item "as built" conformed to the technical documentation defining that item.

7.2.4.2 Post Full Rate Production Decision Review (FRPDR)

Development testing may be necessary after the Full-Rate Production (FRP) decision is made. 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, limited quantity production sampling processes are used. This testing determines whether the system has successfully transitioned from engineering development prototype to production, and whether it meets design specifications.

7.2.5 Operations and Support (O&S)

The DT, which occurs soon after the Initial Operating Capability (IOC) 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 (PIs) and block upgrades, and ensures that Integrated Logistics Support (ILS) 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. For mature systems, DT&E is performed to assist in modifying the system to help meet new threats, add new technologies, or aid in extending service life.

Once a system approaches the end of its usefulness, the DT 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.

7.3 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; the Operational Test and Evaluation (OT&E) agency acts as observer. 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 AAH illustrates the type of DT performed by contractors and the relationship of this type of testing to government DT&E activities. During the contractor competitive testing of the Army AAH, prime contractor and subcontractor testing included design support tests, testing of individual components, establishing fatigue limits, 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 Design Reviews (PDRs), and Critical Design Reviews (CDRs) were conducted.

Following successful completion of the 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 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.

7.3.1 Contractor Testing

Materiel DT&E is an iterative process in which a contractor designs hardware and software, evaluates performance, makes necessary changes, 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 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, an RFP requires that the winning contractor submit an Integrated Engineering Design Test Plan within a short period 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, 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 (ITP).

If the contractor has misinterpreted the RFP requirements and the Integrated 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.

7.3.2 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 decisionmaking; 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 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/T&E 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. They must be responsive to the needs of the PM by providing test facilities, personnel, and data acquisition services, as required.

7.4 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 timed, 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 ITP must also be developed to describe these tests. The PM is responsible for coordinating the total T&E program. The PM performs this task with the assistance of the T&E IPT 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. The PM 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.

7.4.1 Integrated Test Plan

The 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.

7.4.2 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 data from all sources for system evaluation.

7.5 AREAS OF DT&E FOCUS

7.5.1 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.

7.5.2 Design Evaluation/Verification Testing

Design evaluation and verification testing are conducted by the contractor and/or the development agency with the primary objective of influencing system design. Design evaluation is fully integrated into the DT cycle. 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 OTs.

7.5.3 Design Limit Testing

Design Limit Tests (DLTs) 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 extremes 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.

7.5.4 Reliability Development Testing (RDT)

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. RDT 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 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.

7.5.5 Reliability, Availability, and Maintainability

The RAM requirements are assessed during all contractor and government testing. The data are collected from each test event and placed in a RAM database, which is managed by the development agency. Contractor and government DTs 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. Early projections of RAM are important to logistics support planners. The test data facilitate determination of spares quantities, maintenance procedures, and support equipment.

7.6 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/BIT capability. Evaluation of these BIT/BITE/ATE systems must be included in the test program.

7.7 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 fixed-price production contracts. Incentives are the main thrust of warranties, and 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

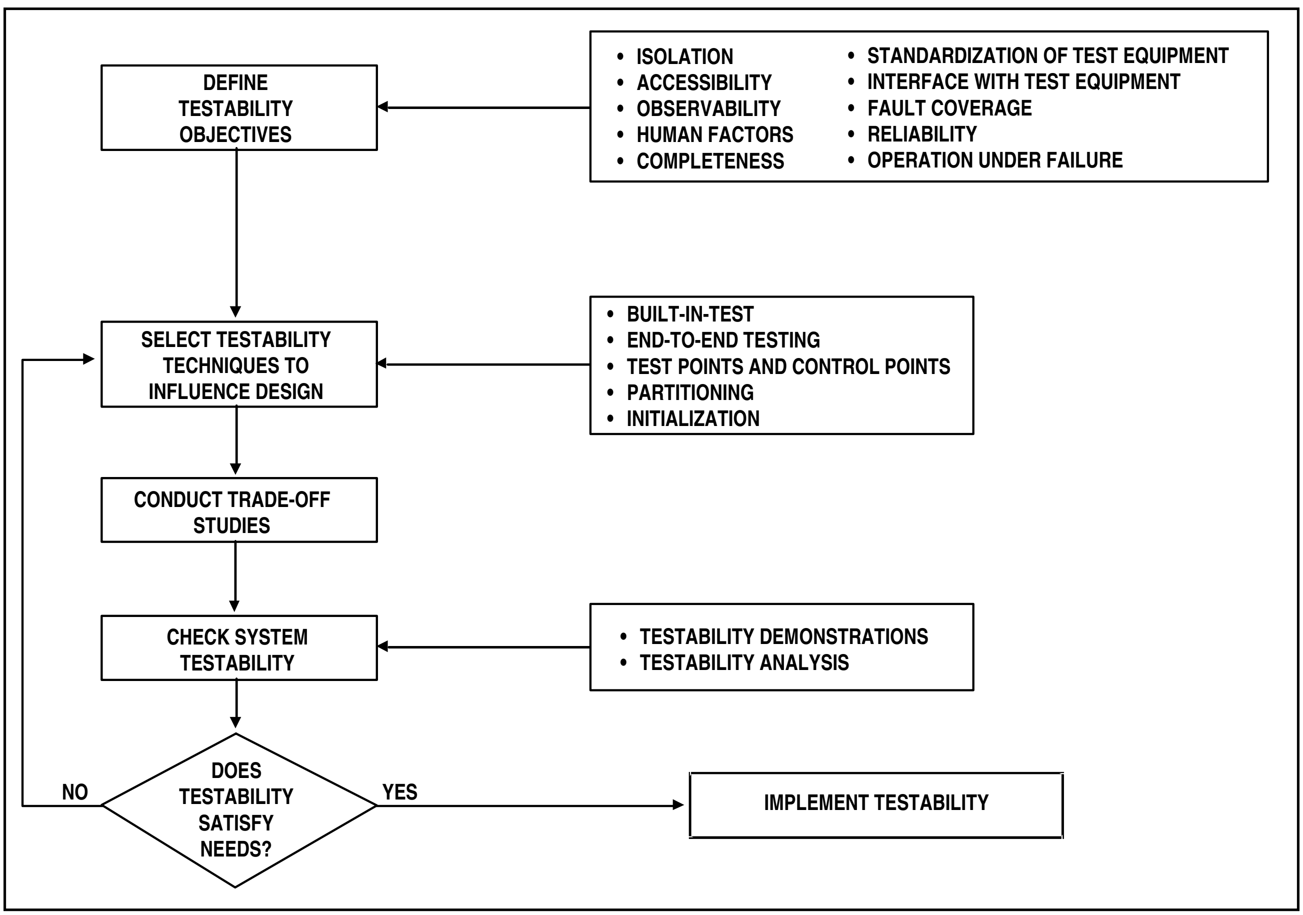


Figure 7-2. Design for Testing Procedures

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.

7.8 DT&E OF LIMITED PROCUREMENT QUANTITY PROGRAMS

Programs that involve the procurement of relatively few items, such as satellites, some large missiles, and unique intelligence equipment, typically over an extended period, are normally subjected to modified DT&E. Occasionally, a unique test approach that deviates from the standard timing and reporting schedule will be used. The DT&E principle of iterative testing starting with components, subsystems, prototypes, and first-production models of the system is normally applied to limited procurements. It is important that DT&E and OT&E organizations work together to ensure that integrated T&E plans are adapted/tailored to the overall acquisition strategy.

7.9 SUMMARY

DT&E 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 performance requirements 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. DT&E 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 (Figure 8-1). The DT&E results are used throughout the Systems Engineering Process (SEP) to provide valuable data in support of formal design reviews. This chapter describes the test's relationship to the Formal Design Reviews (FDRs) essential to the SEP.

8.2 DT&E AND THE REVIEW PROCESS

8.2.1 The Technical Review Process

Technical reviews and audits are conducted by the government and the contractor as part of the SEP 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 (T&E) phases.

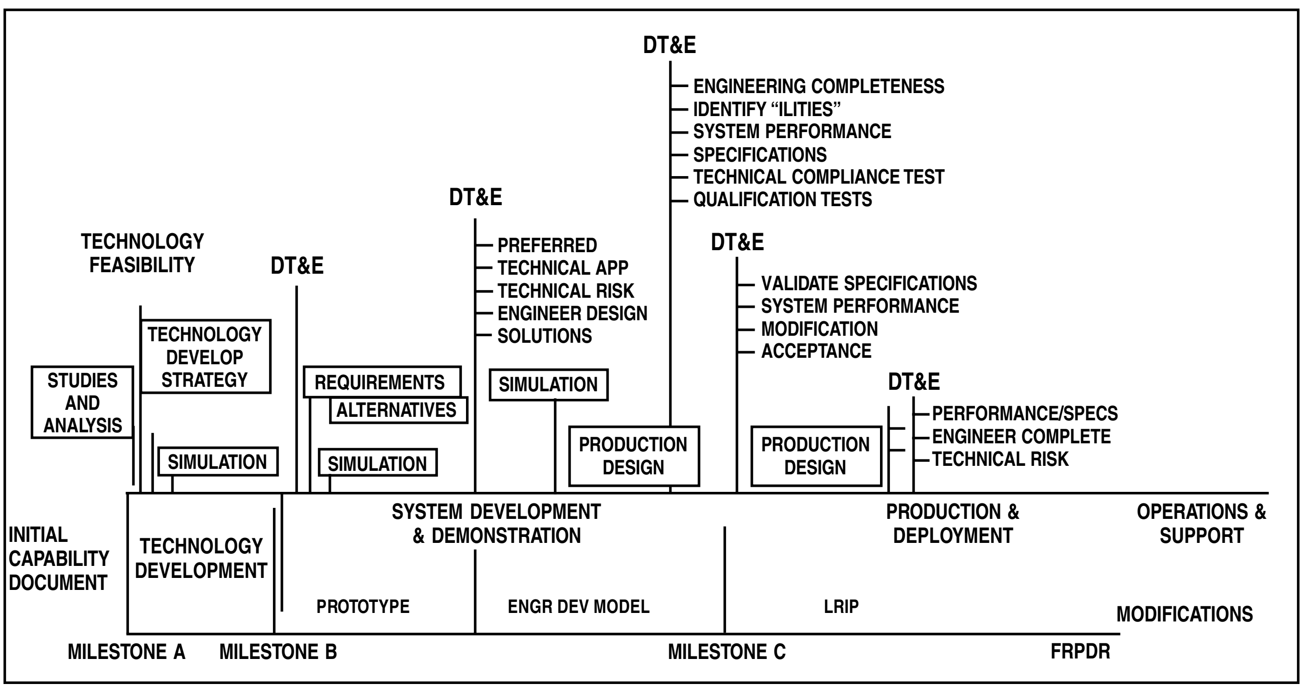


Figure 8-1. Relationship of DT&E to the Acquisition Process

Table 8-1. Technical Reviews and Audits

		WHEN	PURPOSE	DOCUMENTATION DATA
INITIAL TECHNICAL REVIEW	ITR	EARLY CONCEPT REFINEMENT	• EVALUATE INITIAL PROGRAM TECHNICAL BASELINE • ASSESS INITIAL PROGRAM COST ESTIMATES	• COST ANALYSIS REQUIREMENTS DESCRIPTION (CARD) LIKE DOCUMENT • ASSESSMENT OF TECHNICAL AND COSTS RISKS
ALTERNATIVE SYSTEM REVIEW	ASR	LATE CONCEPT REFINEMENT (PRE-MS A)	• ASSESS REQUIREMENTS AGAINST CUSTOMER NEEDS • ASSESS READINESS OF PREFERRED ALTERNATIVE TO ENTER TECHNOLOGY DEV	• TRADE STUDIES • PRELIMINARY SYSTEM SPEC • INPUT TO TDS • INPUT TO AOA • SYSTEMS ENGINEERING PLAN
SYSTEM REQUIREMENTS REVIEW	SRR	EARLY SYSTEMS INTEGRATION	• EVALUATE SYSTEM FUNCTIONAL REQUIREMENTS	• PRELIM PERF SPEC • PRELIM PLANNING DOCUMENTATION • FFBD, RAS, MBN ANALYSIS
SYSTEM FUNCTIONAL REVIEW	SFR	EARLY SYSTEMS INTEGRATION	• EVALUATE SYSTEM DESIGN • VALIDATE SYSTEM SPEC • ESTABLISH SYSTEM LEVEL FUNCTIONAL BASELINE	• PERFORMANCE SPEC • PRELIM ITEM (PERF) SPEC • DESIGN DOCUMENTS • RAS, TLS
SOFTWARE SPECIFICATION REVIEW	SSR	MID SYSTEMS INTEGRATION	• EVALUATE SW PERFORMANCE REQUIREMENTS • VALIDATE SW SPECS • ESTABLISH SW SPECS BASELINE	• SW SPEC • (SRS & IRS) • OPS CONCEPT DOC
PRELIMINARY DESIGN REVIEW	PDR	LATE SI OR EARLY SYSTEMS DEMO	• VALIDATE ITEM (PERF) SPECS • ESTABLISH HW ALLOCATED BASELINE • EVALUATE PRELIMINARY DESIGN HW & SW	• ITEM (PERF) SPEC • DES DOC TEST PLAN • ICD, ENGR DRAWINGS • PRELIMINARY SDD - IDD
CRITICAL DESIGN REVIEW	CDR	EARLY SYSTEMS DEMO	• EVALUATE CI DESIGN • DETERMINE READINESS FOR FABRICATION OF EDM	• PRELIM ITEM (DETAIL), PROCESS, MATERIAL SPECS • DETAIL DESIGN DOCUMENTS INCLUDE SDD - IDD
TEST READINESS REVIEW	TRR	MID/LATE SYSTEMS DEMO	• APPROVE SW TEST PROCEDURES • DETERMINE READINESS FOR FORMAL TEST	• SW TEST PLAN/PROCEDURES • INFORMAL SW TEST RESULTS
FUNCTIONAL CONFIGURATION AUDIT	FCA	LATE SYS DEMO OR EARLY P&D	• VERIFY CI ACTUAL PERFORMANCE COMPLIES WITH HW DEVELOPMENT OR SRS & IRS	• TEST PLANS & DESCRIPTIONS • SOFTWARE TEST REPORTS
FORMAL QUALIFICATION REVIEW	FQR	LATE LRIP	• VERIFY CI'S PERFORM IN SYSTEM ENVIRONMENT	• TEST REPORTS • SPECS • O&S DOCS
PRODUCTION READINESS REVIEW	PRR	INCREMENTAL SYSTEMS DEMO	• ASSESS RISK FOR PRODUCTION GO-AHEAD	• PROD PLANNING DOCUMENTS
PHYSICAL CONFIGURATION AUDIT	PCA	LRIP EARLY FRP	• FORMAT EXAMINATION OF THE AS-BUILT	• FINAL ITEM (DETAIL) SPEC • LISTINGS • LEVEL II & III DRAWING
IN-SERVICE REVIEW	ISR	AFTER SYSTEM IN OPERATIONAL USE (POST-FRPDR)	• ASSESS SYSTEM OPERATIONS IN FIELD CONDITIONS • ASSESS SUPPORT OF FIELDED SYS	• SYSTEM HAZARD RISK ASSESSMENT • OPERATIONAL READINESS ASSESSMENT • DISCREPANCY REPORTS

SOURCE: *Defense Acquisition Guidebook*, September 2004.

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. (Figure 8-2) The system specification is very important in this process. It is the document used as a benchmark to compare contractor progress in designing and developing the desired product. Guidelines for these formal technical reviews and audits can be found in *Processes for Engineering a System*, or *Application and Management of the Systems Engineering Process*.[1]

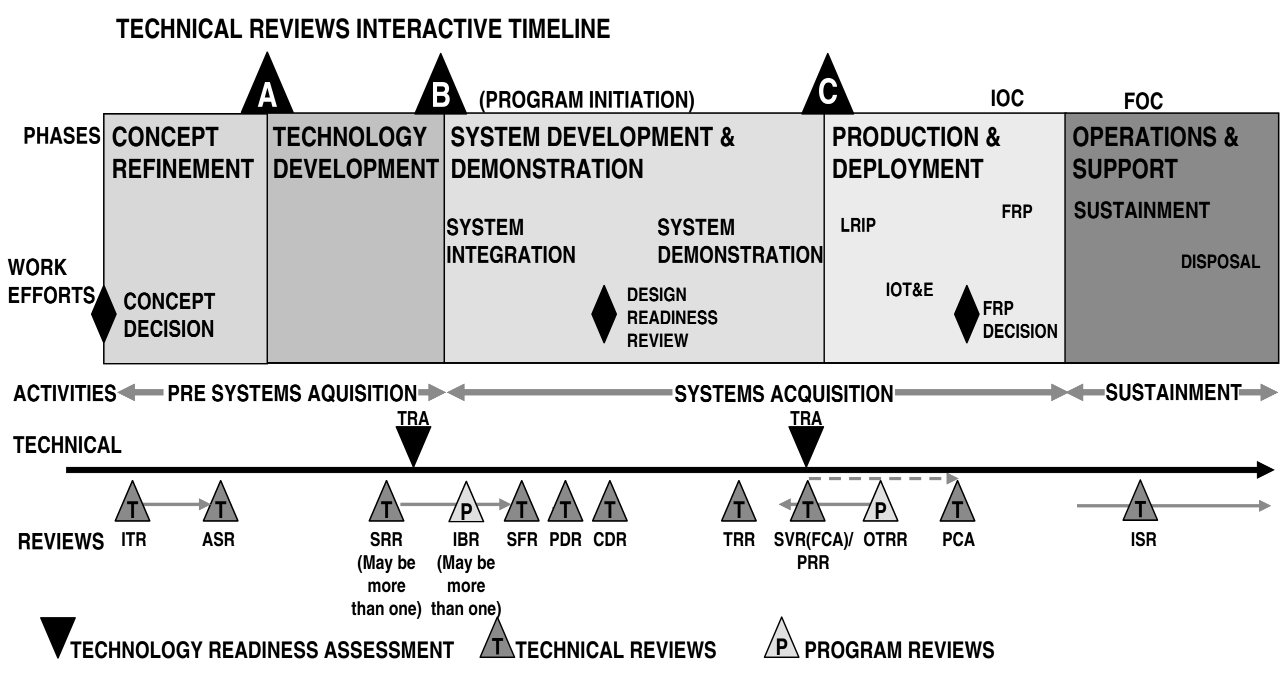
8.2.2 Testing in Support of Technical Reviews

The testing community must be continually involved in supporting the technical reviews of their systems. The timing of these events will vary somewhat from program to program, based upon the explicit and unique needs of the acquisition strategy. Lower level design reviews will lead to system-level reviews. 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 Fundamentals Guide*.[2] The reviews focus primarily on government technical specifications for the system. Figure 8-3 illustrates the program specifications and how they are developed in the system life cycle.

8.2.3 Design Reviews and Audits

8.2.3.1 Concept Refinement (CR) and Technology Development (TD)

The Alternative Systems Review (ASR) is conducted to demonstrate the preferred system concept(s). Technical reviews conducted during the early stages of development tend to focus on



SOURCE: Skalamera, Robert J. November 2004. "Implementing OSD Systems Engineering Policy" briefing to the USD(AT&L).

Figure 8-2. Technical Review Timeline

SPECIFICATION	WHEN PREPARED	PREPARING AGENT	APPROVING AGENT	CONTENT	BASELINE
SYSTEM SEGMENT (OLD TYPE A)	EARLY SI	DEV/PROG MGR INDUSTRY	DEV/PROG MGR USER	DEFINES MISSION AND TECH REQUIREMENTS; ALLOCATES REQUIREMENTS TO FUNCTIONAL AREAS; DOCUMENTS DESIGN CONSTRAINTS; DEFINES INTERFACES	FUNCTIONAL
ITEM PERFORMANCE (OLD TYPE B)	MID SD	INDUSTRY	DEV/PROG MGR	DETAILS DESIGN REQUIREMENTS; STATES, DESCRIBES PERFORMANCE CHARACTERISTICS OF EACH CI; DIFFERENTIATES REQUIREMENTS ACCORDING TO COMPLEXITY AND DISCIPLINE SETS	ALLOCATED "DESIGN TO"
ITEM DETAIL (OLD TYPE C)	LATE SD	INDUSTRY	DEV/PROG MGR	DEFINES FORM, FIT, FUNCTION, PERFORMANCE, AND TEST REQUIREMENTS FOR ACCEPTANCE	PRODUCT "BUILD UP"
PROCESS (OLD TYPE D)	LATE SD	INDUSTRY	DEV/PROG MGR	DEFINES PROCESS PERFORMED DURING FABRICATION	
MATERIAL (OLD TYPE E)	LATE SD	INDUSTRY	DEV/PROG MGR	DEFINES PRODUCTION OF RAW OR SEMI-FABRICATED MATERIAL USED IN FABRICATION	

Figure 8-3. Specifications Summary

top-level concepts and system definitions clarifying the requirements on which subsequent design and development activities will be based.

8.2.3.2 System Development and Demonstration (SDD)

The System Requirements Review (SRR) is normally conducted late in the system concept evaluation or shortly after program initiation. It consists of a review of the system/system segment specifications, previously known as the "A" specifications,[3] 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 (SOW) requirements. The primary function of the SRR is to ensure that the system's 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.

The System Functional Review (SFR) is conducted as a final review before submittal of the prototype design products. The system specification is validated to ensure that the most current specification is included in the System Functional Baseline and that they are adequate and cost-effective to satisfy validated mission requirements. The SFR 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 (OCD), Software Requirements Specifications (SRSs), and Interface Requirements Specifications (IRSs) during this review.

The Software Specification Review (SSR) is a formal review of the Computer System Configuration Item (CSCI) requirements, normally held after an SFR 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 SRSs, IRSs, and OCDs.

The Preliminary Design Review (PDR) is a formal technical review of the basic approach for a Configuration Item (CI). It is conducted at the CI and system level early in the SDD Phase to confirm that the preliminary design logically follows SFR findings and meets the system requirements. The review results in an approval to begin the detailed design. The draft item specifications (performance) 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 CI design approach; review Development Test (DT) and Operational Test (OT) activities to measure the performance of each CI; and establish the existence and compatibility of the physical and functional interface among the CI and other equipment.

The Critical Design Review (CDR) may be conducted on each CI and/or at the system level. It is conducted on the Engineering Development Model (EDM) design when the detailed design is essentially complete and prior to the Functional Configuration Audit (FCA). During the CDR, the overall technical program risks associated with each CI are also reviewed on a technical, cost, and schedule basis. It includes a review of the item specifications (detail) 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 Low Rate Initial Production (LRIP). Test plans are reviewed to assess if test efforts are being developed sufficiently to indicate a Test Readiness Review (TRR) would be successful. The CDR should precede the program DRR.

The TRR was a formal review of the contractor's readiness to begin CSCI testing, but has evolved into a review of both hardware and software. It is conducted after the software test procedures are available and computer software component testing has been completed. A government witness will observe the system demonstration to verify that the system is ready to proceed with CSCI testing. The purpose of the TRR is for the Program Management Office (PMO) to determine whether the test procedures are complete and verify their compliance with test plans and descriptions.

8.2.3.3 Production and Deployment (P&D)

The FCA is a formal review to verify that the CI's performance complied with its system specification. The item specifications are derived from the system requirements and baseline documentation. During the FCA, all relevant test data are 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. The information provided will be used during the FCA to determine the status of planned tests.

The Physical Configuration Audit (PCA) is a formal review that establishes the product baseline as reflected in an early production CI. It is the examination of the as-built version of hardware and software CIs 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 (QA) activities. It includes a detailed audit of engineering drawings, final Part II item specifications (detail), 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.

The System Verification Review (SVR) 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 (the production configuration), as determined through testing, complies with its item specifications (performance) and to document the results of the qualification tests. The SVR 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 SVR can be scheduled for a later time.

The Production Readiness Review (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 FRP decision. For more information, see Chapter 10, Production Related Testing Activities.

8.3 CONFIGURATION CHANGE CONTROL

Configuration Change Control is reviewed to assess the impact of engineering or design changes. It is conducted by the engineering, T&E, and Program Manager (PM) portions of the PMO. Most approved Class I Engineering Change Proposals (ECPs) 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 (ECR) 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. Guidelines for configuration control and engineering changes are discussed in C*onfiguration Control – Engineering Changes, Deviations, and Waivers*.[4]

8.4 SUMMARY

Design reviews are an integral and essential part of the SEP. 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 (WBS). 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.

ENDNOTES

1. EIA Standard 632, *Processes for Engineering a System*, January 1999; or IEEE 1220-1994, *Application and Management of the Systems Engineering Process,* 1994.

2. Defense Systems Management College, *Systems Engineering Fundamentals Guide*, January 2001.

3. Military Standard (MIL-STD)-1512A, *Technical Reviews and Audits for Systems, Equipment, and Computer Programs*. System Functional Block Diagram, Chapter 12, September 30, 1997.

4. EIA Standard 649, MIL-STD-480, C*onfiguration Control – Engineering Changes, Deviations, and Waivers,* August 1, 1998.

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 that 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 in which development testing and operational testing take place at the same time as two parallel, but separate and distinct, activities. In contrast, combined/integrated 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. (Table 9-1)

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 munition 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 Hardness and Survivability (NHS) testing, systems cannot be tested in a totally realistic operational environment; therefore, a single test program is often used to meet both DT and OT objectives.

A combined/integrated Development Test and Evaluation (DT&E) and Operational Test and Evaluation (OT&E) approach should be considered when there are time and cost savings.[1] 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 are 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 OT 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

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

COMBINED TESTING	
ADVANTAGES	**LIMITATIONS**
• 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 DATABASE. • 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 THE 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 DT DATA AVAILABLE FOR OT&E PLANNING AND EVALUATION. • CONTRACTOR PERSONNEL FREQUENTLY PERFORM MAINTENANCE FUNCTIONS IN A DT&E. LOGISTICS SUPPORT BY USER MUST BE AVAILABLE EARLIER FOR IOT&E. • LIMITED TEST ASSETS MAY RESULT IN LESS COVERAGE THAN PLANNED FOR OT&E OBJECTIVES.

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 OT 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.[2] 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.[3]

This approach placed a great deal of responsibility directly on the shoulders of the T&E Coordinator, and required the T&E Coordinator's staff to deal knowledgeably with a wide-ranging and complex test plan.

9.3 CONCURRENT TESTING

In 1983, a senior Department of Defense (DoD) T&E official testified that a concurrent testing approach is usually not an effective strategy.[4] 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 2 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, OT&E, and contractor 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.

ENDNOTES

1. As suggested by DoD Instruction (DoDI) 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003.

2. Hoivik, Thomas H., *The Navy Test and Evaluation Process in Major Systems Acquisition.* DTIC [Defense Technical Information Center] Accession Number AD A035935, November 1976.

3. *Ibid.*

4. Hearing before the Senate Committee on Governmental Affairs, June 23, 1983.

10

PRODUCTION-RELATED TESTING ACTIVITIES

10.1 INTRODUCTION

Most of the Test and Evaluation (T&E) discussed in this guide 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 (QA) sections of the program office and a representative of the government Defense Contract Management Agency (DCMA) oversee/perform many of these functions.

10.2 PRODUCTION MANAGEMENT

Production (manufacturing) management is 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. Production management begins early in the acquisition process, as early as the concept assessments, and is specifically addressed at each program milestone decision point. For instance, before program initiation production feasibility, costs and risks should be addressed. 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 for the program start decision. As development proceeds, the manufacturing strategy is developed; and detailed plans are made for production. Independent producibility assessments, conducted in preparation for the transition from development to production, are reviewed before entering Low Rate Initial Production (LRIP). Once production starts, 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 LRIP decision is supported by an assessment of the readiness of the system to enter production. The system cannot enter production 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 is the Production Readiness Review (PRR), which is conducted by the PM and staff.

10.3 PRODUCTION READINESS REVIEW

The following is an overview of PRRs:

> This review is intended to determine the status of completion of the specific actions that must be satisfactorily accomplished prior to executing a production go-ahead decision. The review is accomplished in an incremental fashion before commencing production. Usually two initial reviews and one final review are conducted 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 PRRs is a function of program posture and is not specifically locked into other reviews.

The conduct of a PRR (Table 10-1) 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.

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. Down line Production Qualification Tests (PQTs) 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

PQTs are a series of formal contractual tests conducted to ensure design integrity over the specified operational and environmental range. The tests are conducted on pre-Full Rate Production (FRP) items fabricated to the proposed production design drawings and specifications. The PQTs include all contractual Reliability and Maintainability (R&M) demonstration tests required prior to production release. For volume acquisitions, these tests are a constraint to production release.

10.4.2 First Article Tests (FATs)

FATs 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 online.[1]

10.5 TRANSITION TO PRODUCTION

In an acquisition process, often the first indication that a system will experience problems is during the transition from engineering design to LRIP. 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

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 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 2 TO 5 YEARS
- ASSOCIATED COMPUTER SOFTWARE DEVELOPED
- SKILLED PERSONNEL AND TRAINING PROGRAMS AVAILABLE

PRODUCTION ENGINEERING AND PLANNING (PEP)

- PRODUCTION PLAN DEVELOPED*
- 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 MATERIAL 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

*Military Standard (MIL-STD)-1528, *Manufacturing Master Plan.*

producibility was given early in the program's acquisition process.

10.5.1 Transition Planning

Producibility Engineering and Planning (PEP) is the common thread that guides a system from early concept to production. Planning 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. Planning 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 LRIP 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 (CI) 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 is the production of a system in limited quantity to provide articles for IOT&E and to demonstrate production capability. Also, it permits an orderly increase in the production rate sufficient to lead to FRP upon successful completion of operational testing. The decision to have an LRIP is made at the Milestone C approval of the program acquisition strategy. At that time, the PM must identify the quantity to be produced during LRIP and validate the quantity of LRIP articles to be used for IOT&E [Acquisition Category (ACAT) I approved by the Director, Operational Test and Evaluation (DOT&E), ACAT II and III approved by Service Operational Test Agency (OTA)].[2] When the decision authority thinks the systems will not perform to expectation, the PM may direct that it not proceed into LRIP until there is a program review. The DOT&E submits a Beyond LRIP (BLRIP) Report on all oversight systems to congressional committees before the FRP decision, approving the system to proceed beyond LRIP, is made.

10.7 PRODUCTION ACCEPTANCE TEST AND EVALUATION (PAT&E)

The PAT&E 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 pre-FRP models. The procured items or system must operate in accordance with system and item specifications. The PAT&E is usually conducted by the program office QA 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 QA 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 on-board 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 (MDAP). 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, conforms to the specifications of the contract, that the IOT&E is conducted with representative system configurations, and the system fielded meets the user's needs.

ENDNOTES

1. Federal Acquisition Regulation (FAR) Part 9.3, *First Article Testing and Approval,* September 2001.
2. DoD Instruction (DoDI) 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003.

III

MODULE

OPERATIONAL TEST AND EVALUATION

Operational Test and Evaluation (OT&E) is conducted to ensure that 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

OT&E is conducted for major programs by an organization that is independent of the developing, procuring, and using commands. Some form of operational assessment is normally conducted in each acquisition phase. Each assessment should be keyed to a decision review in the materiel acquisition process. It should include typical user operators, crews, or units in realistic combat simulations of 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 operational 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 to correct performance deficiencies;

(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 Operational Test Agency (OTA), 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, the Integrated Product Teams (IPTs), and test coordinating groups play important parts in the overall OT&E planning and execution process.

11.3.1 Service Operational Test Agencies

The OT&E agencies should become involved early in the system's life cycle, usually during the program's evaluation of concepts. At this time, they can begin to develop strategies for conducting operational tests (OT&E). As test planning continues, a more-detailed Test and Evaluation Master Plan (TEMP), Part IV (OT&E), 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 (OT) 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 (COIs). The OTA test directors, analysts, and evaluators design the OT&E so that the data collected will support answering the COIs. 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 Test and Evaluation Command (ATEC), Navy Operational Test and Evaluation Force (OPTEVFOR), Air Force Operational Test and Evaluation Center (AFOTEC), and Marine Corps Operational Test and Evaluation Activity (MCOTEA). Although policy only requires OTA within the Services, the Defense Information Systems Agency (DISA) has designated the Joint Interoperability Test Command (JITC) as the OTA for their Department of Defense (DoD) information systems and a Special Operations Command (SOCOM) element conducts many of its own IOT&E.

11.3.2 Test Personnel

Operational testing is conducted on materiel systems with "typical" user organizational units in a realistic operational environment. It uses personnel with the same Military Occupational Specialties (MOSs) 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 OT&E consist of force-on-force tests, the forces opposing the tested system must also be trained in the use of threat equipment, 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 Service school cadre who then train the participating organizational units. Once the system has entered Full-Rate Production (FRP), the Service will normally assume training responsibilities. Operational testing often requires a large support staff of data collectors and scenario controllers operating in the field with the user test forces and opposing forces.

11.4 TYPES OF OT&E

The OT&E can be subdivided into two phases: operational testing performed before FRP and the operational testing performed after the FRP decision. The pre-FRP OT&E includes Early Operational Assessments (EOAs), Operational Assessments (OAs), and IOT&E. OAs begin early in the program, frequently before program start and continue until the system is certified as ready for IOT&E. The IOT&E is conducted late during low rate production, when test assets are available, in support of the full rate decision review. The Navy uses the term "OPEVAL" (Operational Evaluation) to define IOT&E. After transition to FRP, all subsequent operational testing is usually referred to as Follow-on Operational Test and Evaluation (FOT&E). In the Air Force, if no Research and Development (R&D) funding is committed to a system, Qualification Operational Test and Evaluation (QOT&E) may be performed in lieu of IOT&E.

11.4.1 Early Operational Assessments

The EOAs are conducted primarily to forecast and evaluate the potential operational effectiveness and suitability of the weapon system during development. EOAs start during Concept Refinement (CR) and/or Technology Development (TD), may continue into system integration, and are conducted on prototypes of the developing design.

11.4.1.1 Operational Assessments

The OAs begin when the OTAs start their evaluations of system-level performance. The OTA uses any testing results, Modeling and Simulation (M&S), 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 OAs of performance requirements are conducted on Engineering Development Models (EDMs) or pre-production articles until the system performance is considered mature. The mature system can then be certified ready for its IOT&E (OPEVAL in the Navy). When using an evolutionary acquisition strategy, subsequent increments must have at least an OA before the new increment is issued to the user.

11.4.1.2 Initial Operational Test and Evaluation (Navy OPEVAL)

The IOT&E is the final dedicated phase of OT&E preceding an FRP 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 in compliance with United States Code (U.S.C.).[1] The IOT&E is conducted by an OT&E agency independent of the contractor, PMO, or developing agency. The test has been described as:

> All OT&E is conducted on production or production-representative articles, to support the decision to proceed beyond Low Rate Initial Production [LRIP]. 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 generally 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[2] by Congress. The results from this test are evaluated and presented to the Milestone Decision Authority (MDA) (i.e., the decision to enter FRP) to support the Beyond-Low-Rate Initial Production (BLRIP) decision. This phase of OT&E addresses the Key Performance Parameters (KPPs) identified in the Capability Production Document (CPD) and the Critical Operational Issues (COIs) in the TEMP. IOT&E test plans for ACAT I and IA and other designated programs must be approved by the OSD (Office of the Secretary of Defense), Director, Operational Test and Evaluation (DOT&E). Service IOT&E test reports provide the foundation for the DOT&E BLRIP Report.

11.4.2 Follow-On Operational Test and Evaluation (FOT&E)

The FOT&E is conducted after the FRP decision. The tests are conducted in a realistic tactical environment similar to that used in IOT&E, but fewer test items may be used. Normally FOT&E is conducted using fielded production systems with appropriate modifications, upgrades, or increments. Specific objectives of FOT&E include testing modifications that are to be incorporated into production systems, completing any deferred or incomplete IOT&E, evaluating correction of deficiencies found during IOT&E, and assessing reliability including spares support on deployed systems. 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 (QOT&E) (USAF)

Air Force QOT&E may be performed by the major command, user, or AFOTEC and is conducted on minor modifications or new applications of existing equipment when no R&D 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 (OT) 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. OT planning is conducted by the OTA before program start, and more detailed planning usually starts about 2 years before each OT event.

Operational planning can be divided into three phases: early planning, advanced planning, and detailed planning. Early planning entails developing COIs, formulating a plan for evaluations, determining the Concept of Operation (COO), 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

The COIs have been described as:

> A key operational effectiveness or operational suitability issue that must be examined in OT&E to determine the system's capability to perform its mission. A COI 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 Operational Requirement Document (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 OT. Identifying the issues is analogous to the first step in the Systems Engineering Process: defining the problem. When COIs are properly addressed, deficiencies in the system can be uncovered. They form the basis for a structured technique of analysis by which detailed sub-objectives or MOEs can be established. During the OT, each sub-objective is addressed by an actual test measurement (Measure of Performance (MOP)). After these issues are identified, the evaluation plans and test design are developed for test execution. (For more information, see Chapter 13.)

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 field test 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) Realistic combat 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, Nuclear, Biological and Chemical (NBC); Electronic Countermeasures (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: A Systems Engineering Process*, Roger Stevens stated, "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.[3]

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 roadmap used for planning T&E events. The DOT&E is briefed on test concepts for oversight programs and approves the test plan before IOT&E starts.

11.6 TEST EXECUTION

An OT 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. The test director 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. The test director's data management duties will encompass collecting raw data, creating a data status matrix, and ensuring data quality by processing and reducing, verifying, filing, storing, retrieving, and analyzing collected data. Once all tests have been completed and the data are reduced and analyzed, the results must be reported. A sample test organization used for the Army OT&E of the program is illustrated in Figure 11-1. (In the Army, the Deputy Test Director comes from the OTA and controls the daily OT activity.)

11.7 TEST REPORTING

The IOT&E test report is a very important document. It must communicate the results of

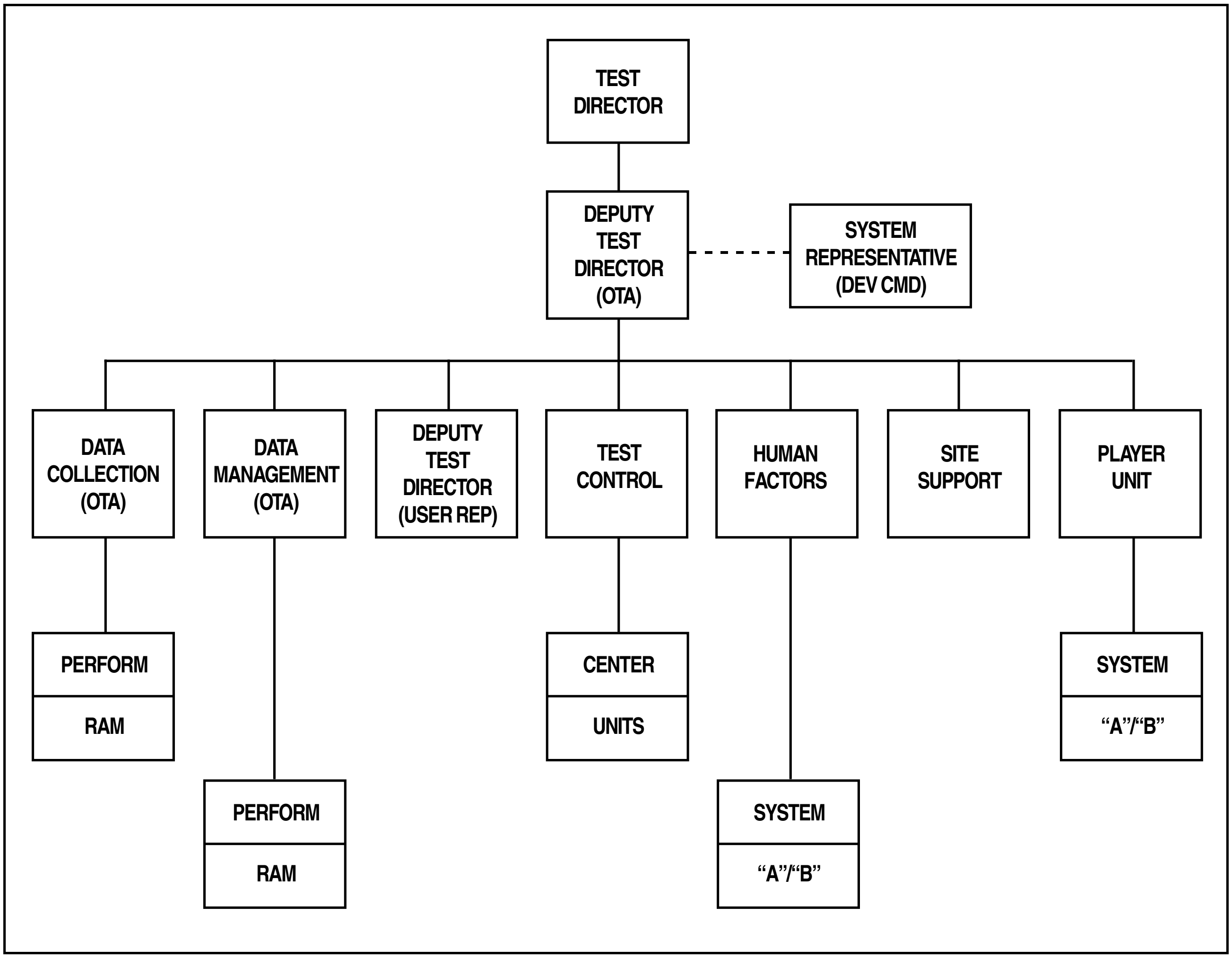


Figure 11-1. Organizational Breakdown of the I-81mm Mortar Operational Test Directorate

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 (Air Force, Navy) or in separate documents (Army, Marine Corps). The Army's Independent Evaluation Report (IER) advises the decision review principals and MDA concerning the adequacy of testing, the system's operational effectiveness and suitability, and survivability, as well as providing recommendations for future T&E and system improvements.

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 (discrete 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 milestones. The final T&E report (Air Force, Navy) or IER (Army, Marine Corps) presents the conclusions and recommendations including all supporting data and covering the entire IOT&E program. The Service OTA are required to provide reports of OT&E results in support of Milestones B and C, in addition to the IOT&E report required for the Full Rate Production Decision Review (FRPDR).[4]

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 Reliability and Maintainability (R&M) of the system; the effort and level of training required to maintain, support, and operate it; and any unique logistics 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.

ENDNOTES

1. Title 10, U.S.C., Section 2399, *Operational Test and Evaluation of Defense Acquisition Programs*.

2. Public Law 99-661, *National Defense Authorization Act (NDAA)*, Section 1207.

3. Stevens, Roger T., *Operational Test and Evaluation: A Systems Engineering Process*, John Wiley & Sons, New York: 1979, pp. 49-56.

4. As required by DoD Instruction (DoDI) 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003.

12

OT&E TO SUPPORT DECISION REVIEWS

12.1 INTRODUCTION

Mindful of principles of objectivity and impartial evaluation, Operational Test and Evaluation (OT&E) may be conducted before each decision review to provide the decision authority with the latest results from testing of Critical Operational Issues (COIs). 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 COIs.

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 Department of Defense (DoD) executive-level management decisions on major acquisition programs. OT&E 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 Beyond Low Rate Initial Production (BLRIP). Figure 12-1 illustrates how OT&E relates to the acquisition process.

12.2 CONCEPT REFINEMENT (CR) AND TECHNOLOGY DEVELOPMENT (TD)

The OT&E conducted during the first phase may be an Early Operational Assessment (EOA) 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 program initiation, the OT&E objectives are to assist in evaluating alternative concepts to solve the mission capability deficiencies and to assess the operational impact of the system. An early assessment also may provide data to support a decision on whether the technology is sufficiently mature to support entry into the next development phase. The OT&E conducted during this phase supports developing estimates of:

(1) The military need for the proposed system;

(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 (LCC);

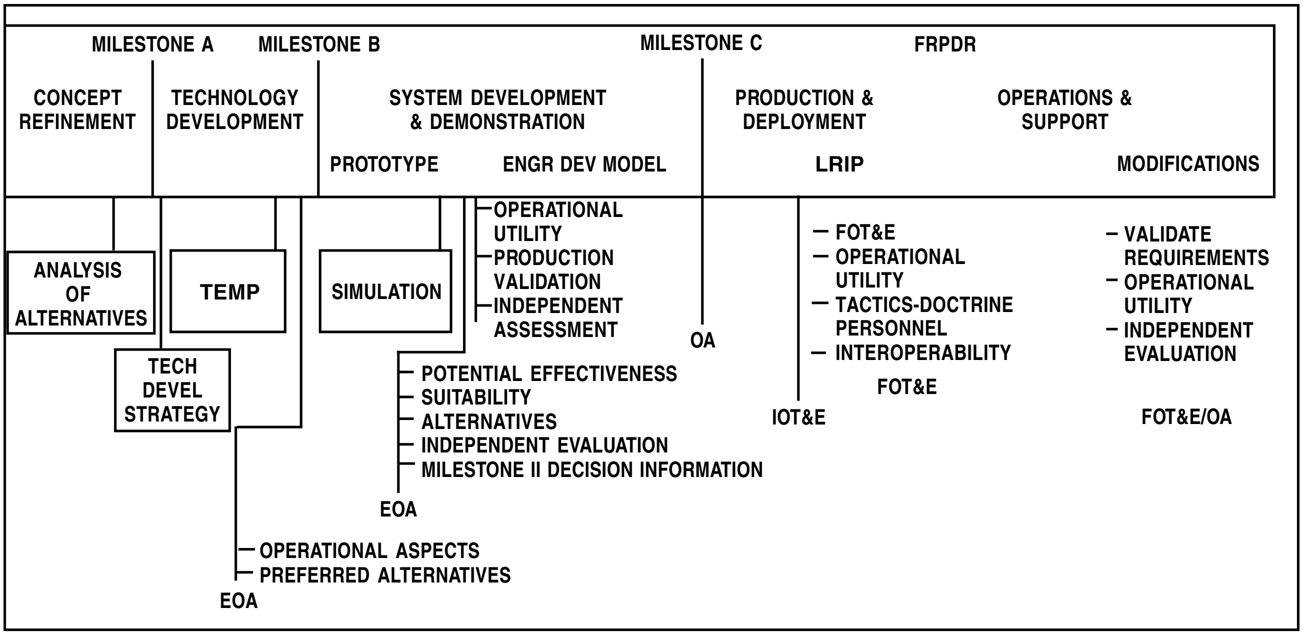


Figure 12-1. OT&E Related to the Milestone Process

(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.

During concept assessment, there is normally no hardware available for the operational tester. Therefore, the EOA is conducted from surrogate test and experiment data, breadboard models, factory user trials, mock-up/simulators, modeling/simulation, 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 planning in the Technology Development Strategy (TDS) that will later evolve into 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 SYSTEM DEVELOPMENT AND DEMONSTRATION (SDD)

After program initiation, combined DT/OT&E or an EOA may be conducted to support the EOA of prototype development and a decision regarding a system's readiness to move into the development of the Engineering Development Model (EDM). In all cases, appropriate T&E must be conducted on a mature system configuration, i.e., the EDM, before the production decision, thereby providing data for identification of risk before more resources are committed. As appropriate, Low Rate Initial Production (LRIP) will follow this phase of testing to verify system-level performance capability and to provide insight into test resources needed to conduct future interoperability, live fire, or operational testing.

12.3.1 Objectives of Operational Assessments

Operational Assessments (EOAs, OAs) 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. EOAs supporting decision reviews 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 versus military utility;

(3) Assess the adequacy of the concept for employment, supportability, and organization; 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 COIs that need to be resolved before construction of EDMs is initiated;

(7) Assess the need during LRIP of long lead hardware to support Initial Operational Test and Evaluation (IOT&E) prior to the Full-Rate Production (FRP) decision;

(8) Provide data to support test planning for this 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 Experimentation Program (CEP) 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

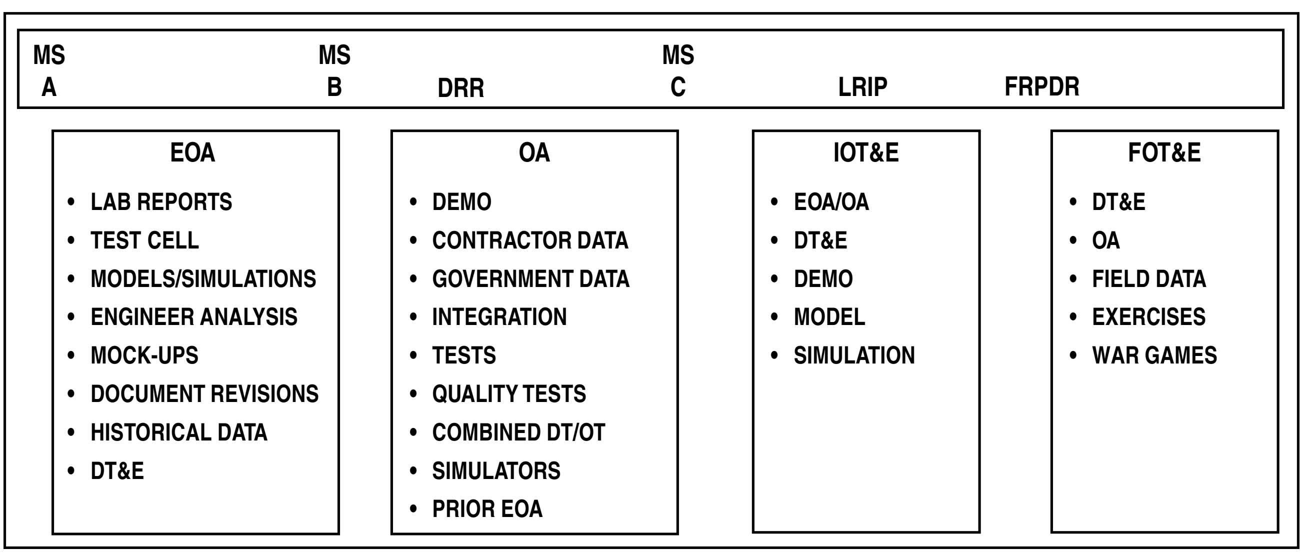


Figure 12-2. Sources of Data

documented in an Operational Assessment (EOA) or end of phase OT&E report. These data, along with the mission needs and requirements documentation and TEMP, assist in the review of performance for the next decision review.

The OAs during the system demonstration are conducted on EDMs. These operational evaluations estimate the operational effectiveness and suit-ability and provide data on whether the system meets minimum operational thresholds.

12.4 PRODUCTION AND DEPLOYMENT (P&D)

Just before the FRP decision, the dedicated T&E is conducted on equipment that has been formally certified by the Program Manager (PM) 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 IOT&E conducted is characterized by testing performed by user organizations in a field exercise to examine the organization and doctrine, Integrated Logistics Support (ILS), threat, communications, Command and Control (C^2), and tactics associated with the operational employment of the unit during tactical operations. This includes estimates that:

(1) Assess operational effectiveness and suitability;

(2) Assess the survivability of the system;

(3) Assess the systems reliability, maintainability, and plans for ILS;

(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 FRP.

12.5 OPERATIONS AND SUPPORT (O&S)

After the FRP decision and deployment, 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 Operational Test and Evaluation (FOT&E) to refine the effectiveness and suitability estimates made during earlier OT&E, assess performance not evaluated during IOT&E, evaluate new tactics and doctrine, and assess the impacts of system modifications or upgrades.

The FOT&E is performed with production articles in operational organizations. It is normally funded with Operations 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 Milestone 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 OA 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 Logistics Support Systems

The testing objectives to evaluate post-production logistics readiness and support are to:

(1) Assess the logistics readiness and sustainability;

(2) Evaluate the weapon support objectives;

(3) Assess the implementation of logistics support planning;

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

(5) Determine the disposition of displaced equipment;

(6) Evaluate the affordability and LCC of the system.

12.6 SUMMARY

The OT&E is that T&E (OAs, IOT&E, or FOT&E) conducted to provide feedback on system design and its potential to be operationally effective and operationally suitable. OT&E events in each phase of development will identify needed modifications; provide information on tactics, doctrine, organizations, and personnel requirements; and evaluate the system's logistics supportability. The acquisition program structure should include feedback from OAs or evaluations at critical decision points/technical reviews beginning early in the development cycle and continuing throughout the system's life cycle.

IV

MODULE

TEST AND EVALUATION PLANNING

Many program managers face several T&E 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 program managers when they develop the integrated strategy for the T&E program.

13

EVALUATION

13.1 INTRODUCTION

This chapter describes the evaluation portion of the Test and Evaluation (T&E) 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 (MOEs) 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.

13.2 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 any "program or procedure that is designed to obtain, verify, or provide data for the evaluation of any of the following: 1) progress in accomplishing developmental objectives; 2) the performance, operational capability, and suitability of systems, subsystems, components, and equipment items; and 3) the vulnerability and lethality of systems, subsystems, components, and equipment items."[1]

"Evaluation" denotes the process whereby data are logically assembled, analyzed, and compared to expected performance to aid in making systematic decisions.

To summarize, evaluation is the process for review and analysis of qualitative or quantitative data obtained from design review, hardware inspection, Modeling and Simulation (M&S), testing, or operational usage of equipment.

13.3 THE EVALUATION PROCESS

The evaluation process requires a broad analytical approach with careful focus on the development of an overall T&E plan that will provide timely answers to critical issues and questions required by decision authorities throughout all the acquisition phases. (Table 13-1) Evaluations should focus on Key Performance Parameters (KPPs); i.e., those "minimum attributes or characteristics considered most essential for an effective military capability." An attribute is a testable or measurable characteristic that describes an aspect of a system or capability.[2]

A functional block diagram of a generic (i.e., not Service-specific) evaluation process is shown in Figure 13-1. The Army's Independent Evaluation Plan (IEP) is written very early and applies to the system and not a specific test event, so it is not as detailed as the format provided in Table 13-1. 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 be established to define

Table 13-1. 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
      - 2.2.1.5.1.X MOPx
  - 2.2.2 ISSUE 2
  - THROUGH
  - 2.2.m ISSUE n
- 2.3 OPERATIONAL SUITABILITY
  - 2.3.1 ISSUE n+1
  - THROUGH
  - 2.3.n ISSUE n+x
- 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)
- 2.7 EVALUATION DATABASE STRUCTURE
  - 2.7.1 IDENTIFICATION OF REQUIRED FILES
  - 2.7.2 DESCRIPTION OF FILE RELATIONSHIPS
  - 2.7.3 DATA ELEMENTS 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*, Department of the Army (DA) Pamphlet 71-3.

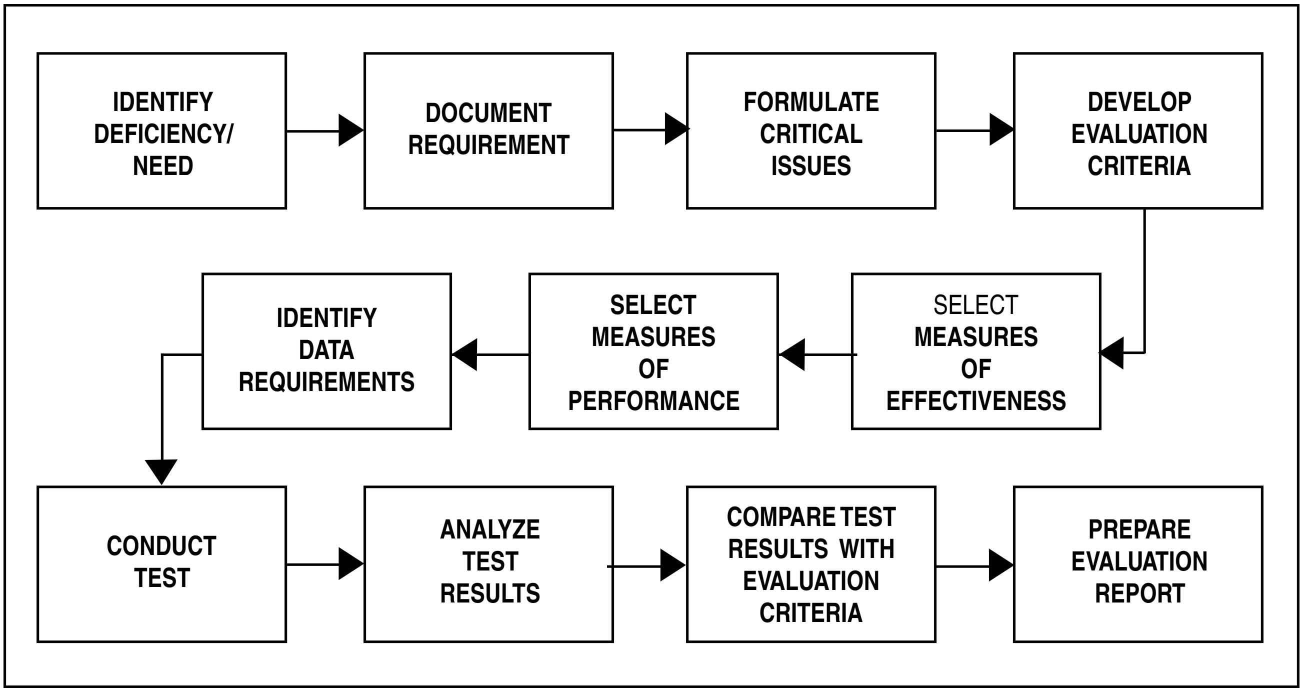


Figure 13-1. Functional Block Diagram of the Evaluation Process

required performance or supportability parameters and to evaluate progress in reaching them. T&E 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.

13.4 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 (MDRs). 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.

13.4.1 Key Performance Parameters (KPPs)/ Critical Issues

The KPPs often can support the development 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 Test and Evaluation Master Plan (TEMP) may be derived from the KPPs found in the capability requirements documents—Capability Development Document (CDD) and

Capability Production Document (CPD). The system requirements and baseline documentation will provide many of the performance parameters required to develop the hierarchy of issues.

13.4.2 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 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, inspection, demonstration, or testing) to resolve the issue. Issues requiring test data are further referred to as test issues.

13.4.3 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 may be partitioned into technical issues that are addressed by the Development Test and Evaluation (DT&E) community (contractors and government) 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.

13.4.4 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."

13.4.4.1 Test of Thresholds and Objectives

A threshold for a performance parameter lists a minimum acceptable operational value below which the utility of the system becomes questionable.[3] Thresholds are stated quantitatively whenever possible. Specification of minimally acceptable performance in measurable parameters is essential to selecting appropriate MOEs, 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. The Program Manager (PM) must budget for the attainment of threshold values.

Objectives are "the desired operational goal associated with a performance attribute, beyond which any gain in utility does not warrant additional expenditure. The objective value is an operationally significant increment above the threshold. An objective value may be the same as the threshold when an operationally significant increment above the threshold is not significant or useful."[4] Objectives are not normally addressed by the operational tester, whose primary concern is to "determine if thresholds in the approved CPD and critical operational issues have been satisfied."[5]

Going into system demonstration, 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 evaluation of Engineering Development Models (EDMs). Along with the CPD, they should remain relatively stable through production.

13.5 MEASURES OF EFFECTIVENESS (MOEs)

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, sub-objectives, and MOEs (sometimes partitioned into MOEs and measures of suitability). For example, an MOE (airspeed) may have an associated evaluation criterion (450 knots) against which the actual performance (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. 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. The MOEs should be so specific that the system's effectiveness during developmental and operational testing can be assessed using some of the same effectiveness criteria as the Analysis of Alternatives [AoAs].[6]

13.6 EVALUATION PLANNING

13.6.1 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.[7]

13.6.1.1 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.

13.6.1.2 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.

13.6.1.3 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" versus "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.

13.6.1.4 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, MOEs, Measures of Performance (MOPs), and data requirements in a rootlike structure as depicted in Figure 13-2. 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. The MOEs 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.[8] Each test objective and MOE is also linked to one or more MOPs (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.

13.6.2 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 Test (DT), Operational Test (OT), and comparable systems.

13.7 EVALUATING DEVELOPMENT AND OPERATIONAL TESTS

Technical and operational evaluations employ different techniques and have different evaluation criteria. The DT&E 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.[9] The DT&E uses technical criteria for evaluating system performance. These criteria are usually parameters that can be measured during controlled DT&E tests. They are particularly important to

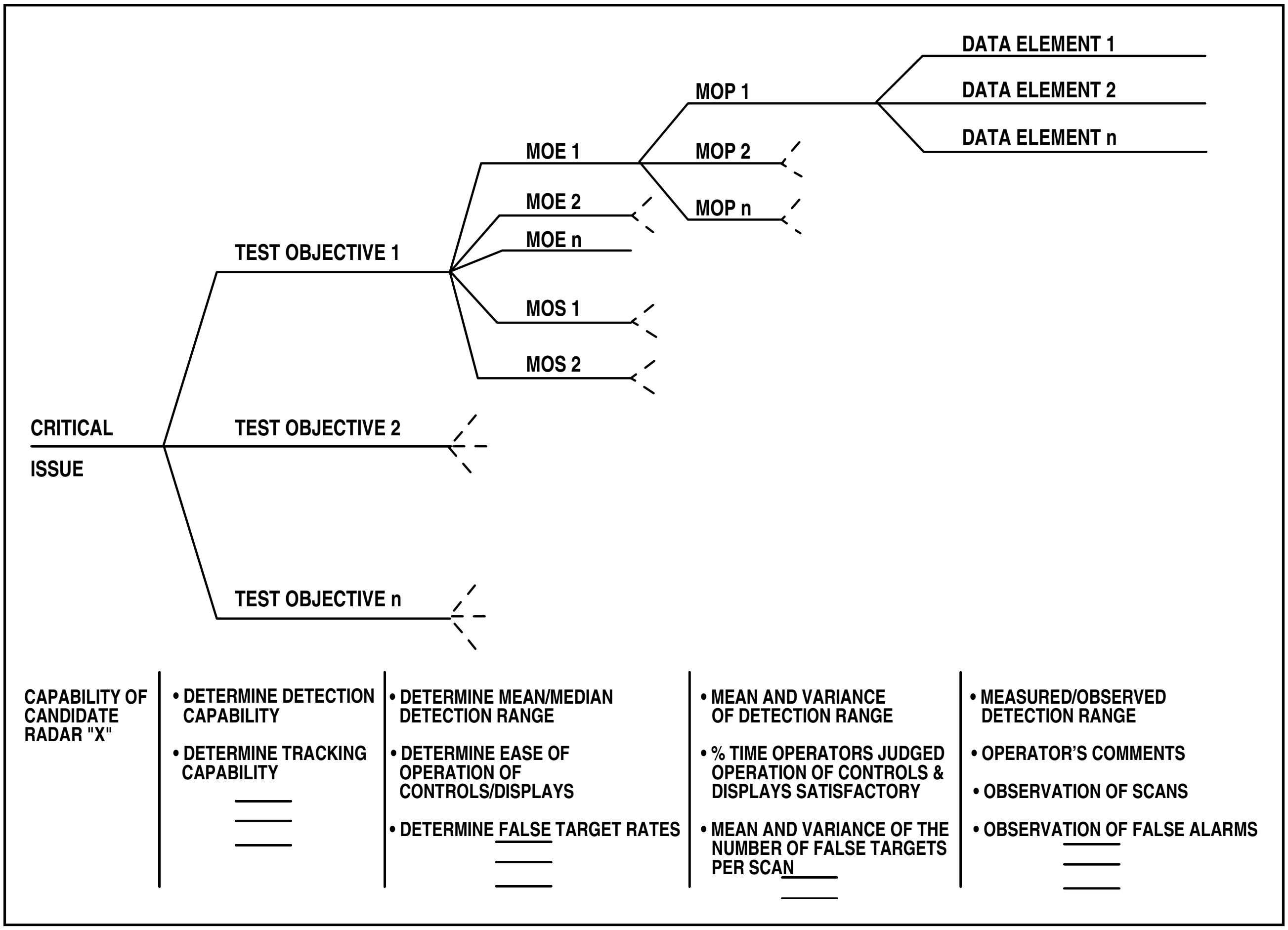


Figure 13-2. Dendritic Approach to Test and Evaluation

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 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 take 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.

13.7.1 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, interoperability, vulnerability, lethality, transportability, Reliability, Availability, and Maintainability (RAM); manpower and personnel; system safety; Integrated Logistics Support (ILS); 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 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.[10]

13.7.2 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 in ensuring 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.[11]

13.8 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.

ENDNOTES

1. Defense Acquisition University Press, *Glossary of Defense Acquisition Acronyms & Terms*, 11th Edition, September 2003.

2. Chairman of the Joint Chiefs of Staff Instruction (CJCSI) 3170.01D, *Joint Capabilities Integration and Development System (JCIDS)*, March 12, 2004, GL (glossary) -9.

3. *Ibid*.

4. *Ibid,* GL -11.

5. Department of Defense Instruction (DoDI) 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003, p. 37.

6. *Ibid.*

7. Department of the Army (DA) Pamphlet 73-1, *Test and Evaluation in Support of Systems Acquisition,* May 30, 2003.

8. Air Force Regulation (AFR) 80-38, *Management of the Air Force Systems Survivability Programs*, May 27, 1994.

9. NAVSO-P-2457, *RDT&E/Acquisition Management Guide*, 10th Edition.

10. Army Regulation (AR) 73-1, *Test and Evaluation Policy*, January 7, 2002.

11. AFR 800-2, *Acquisition Program Management*, July 21, 1994.

14

MODELING AND SIMULATION SUPPORT TO T&E

14.1 INTRODUCTION

This chapter discusses the applications of Modeling and Simulation (M&S) in Test and Evaluation (T&E). The need for M&S 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, Department of Defense Instruction (DoDI) 5000.2 encourages the use of M&S as a source of T&E data.[1] For instance, the Armored Family of Vehicles (AFV) program used more than 60 models, simulations, and other test data to support system concept exploration. The reliance on M&S by this and other acquisition programs provides the T&E community with valuable information that can increase confidence levels, decrease field test time and costs, and provide data for pre-test prediction and post-test validation. The Defense Modeling and Simulation Office (DMSO), working for the Director Defense Research and Engineering (DDR&E), is developing Office of the Secretary of Defense (OSD) guidance on the application of M&S to the acquisition process. The DMSO operates the Modeling and Simulation Information Analysis Center (MSIAC) to provide assistance to all Department of Defense (DoD) M&S users, including program offices and the acquisition community at large.

This chapter discusses using M&S 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.

14.2 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, Hardware-in-the-Loop (HWIL) simulations, 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 (DDT&E),[2] simulations used in T&E can be divided into three categories:

Constructive Simulations: 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 T&E.

Virtual Simulations: 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.

Another type of virtual 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. HWIL simulators or full-up man-in-the-loop 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.

Live Simulations: Some say that everything except global combat is a simulation, even limited regional engagements. Live exercises where troops use equipment under actual environmental conditions approach real life in combat while conducting peacetime operations. Training exercises and other live simulations provide a testing ground with real data on actual hardware, software, and human performance when subjected to stressful conditions. These data can be used to validate the models and simulations used in an acquisition program.

As illustrated in Figure 14-1, 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. As you rely less on live forces in exercises, the ability to clearly understand the live-force exercise's representativeness decreases. At the other end of the spectrum, as you rely more on virtual simulations (software-generated activities), more of the models' internal complexities and increasing complexity impacts the accuracy of results. Examples of uncertainty would be defining why a battle was lost in a war

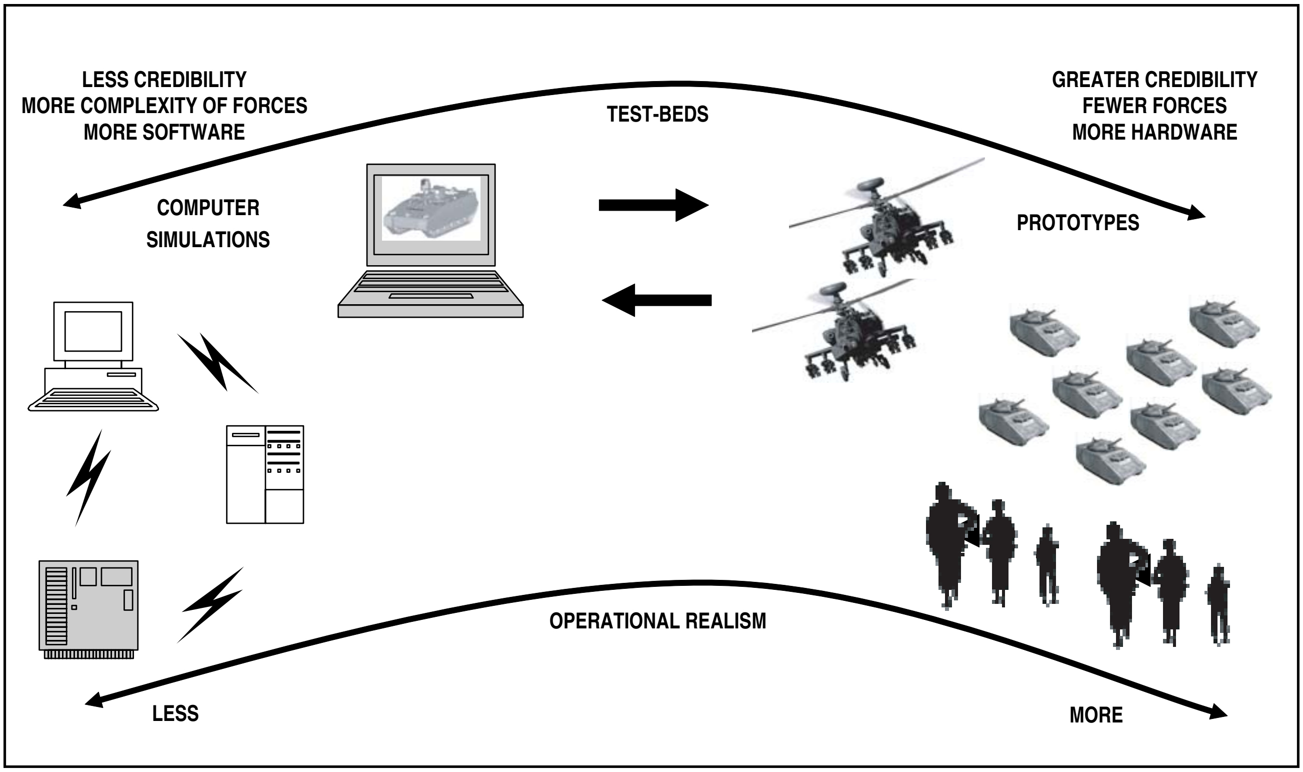


Figure 14-1. The Simulation Spectrum

game simulation, or a system failed to perform as expected in a prototype or full system simulation, or why in a flight simulator, the aircraft seemed to become inverted when crossing the equator. Determining accuracy is a factor of the technical level of confidence or credibility of the simulation.

14.3 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, including the decision process and the proficiency of personnel in the performance of their functions. Therefore, models and simulations are not a total substitution for physical T&Es. Simulations, manual and computer-designed, can complement and increase the validity of live T&Es by proper selection and application. Figure 14-2 contrasts the test criteria that are conducive to M&S versus 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.

Conversely, certain highly complex simulations can be more effective analytical tools than limited live testing, generally due to the significance of the external constraints. For example, nuclear weapons effects testing can only be done in simulation because of international treaties. Testing depleted uranium ammunition is very expensive with its related environmental considerations. Finally, many confidence-building iterations of multimillion dollar missile defense tests can only be executed using simulated events in conjunction with a few instances of real firings.

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. Models and simulations must be approved for use through

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 14-2. Values of Selected Criteria Conducive to Modeling and Simulation

verification, validation, and accreditation processes.[3] Verification is the process of determining that a model implementation accurately represents the developer's conceptual description and specifications. Validation is the process of determining: (a) the manner and degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model; and (b) the confidence that should be placed on this assessment. Accreditation is the official certification that a model or simulation is acceptable for use for a specific purpose.

14.4 SUPPORT TO TEST DESIGN AND PLANNING

14.4.1 Modeling and Simulation in T&E Planning

The M&S can assist in the T&E planning process and can reduce the cost of testing. In Figure 14-3, areas of particular application include scenario development and the timing of test events; the development of objectives, essential elements of analysis, and MOEs; 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. The test designer can also predict the effects of various assumptions and constraints and evaluate candidate MOEs to help formulate the test design. In the live fire vulnerability testing of the F/A-22 wing panel, a mechanical deformation model was used extensively in pre-test predictions. The pre-test analysis was used to design the experiment that assisted in shot-line selection and allowed elimination of the aft boom testing saving schedule and about $100,000. Real data in post-test analysis verified the capability to predict impulse damage within acceptable limits, resulting in greater use of the model in future applications.

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

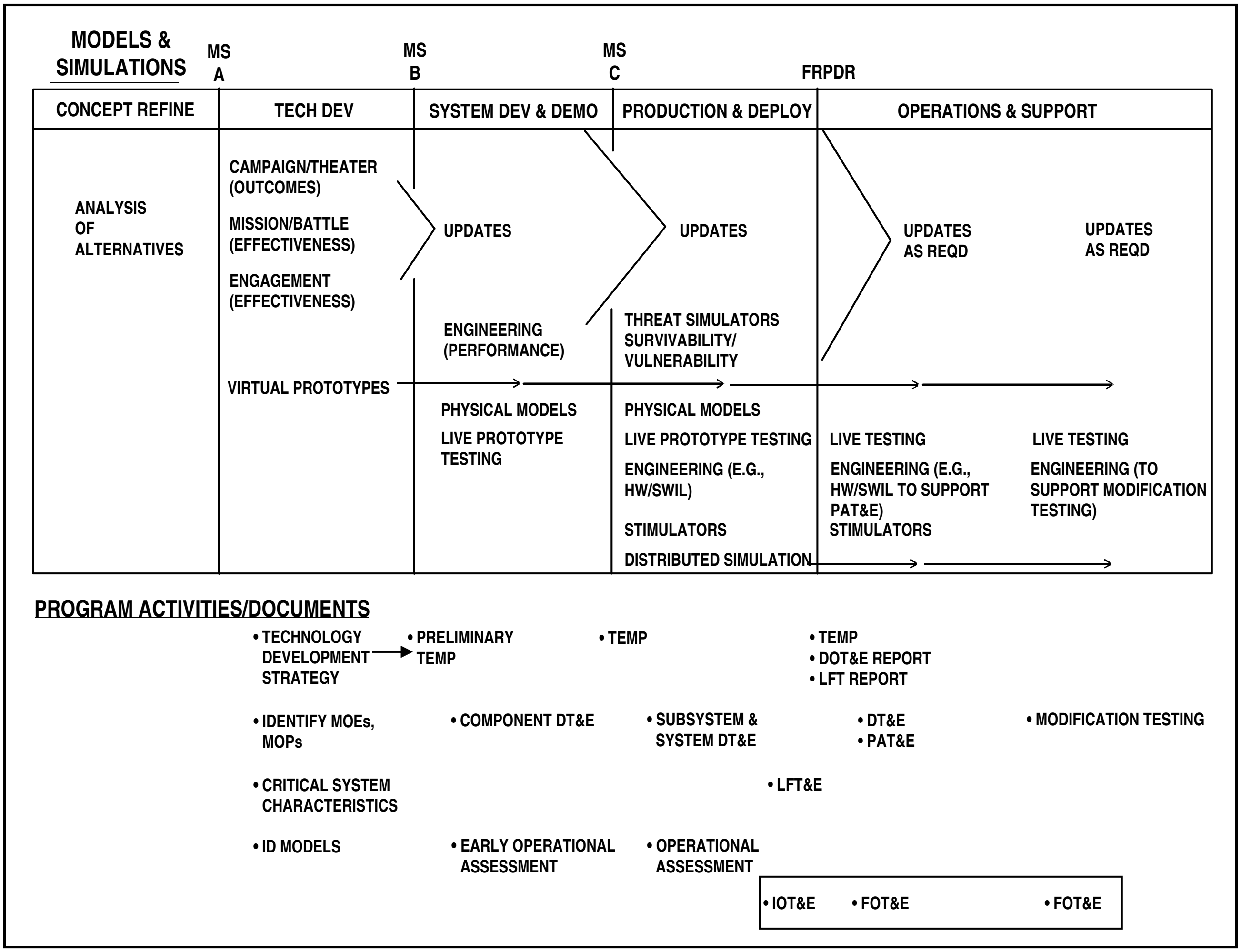


Figure 14-3. Modeling and Simulation Application in Test and Evaluation

"cure-all," they should be explored and used whenever feasible as another source of data for the evaluator to consider during the test evaluation. Simulation fidelity is improving rapidly with enhancements to computer technology, and it is no longer cost prohibitive to develop credible M&S applications for future weapon systems.

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 MOEs 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 pre-test planning is to preview the test to evaluate ways to make test results more effective. Pre-testing attempts to optimize test results by pointing out potential trouble spots. It constitutes a test setup analysis, which can encompass a multitude of areas. The model-test-model process is an integrated approach to using models and simulations in support of pre-test analysis and planning; conducting the actual test and collecting data; and post-test analysis of test results along with further validation of the models using the test data.

As an example of simulations used in test planning, consider a model that portrays aircraft versus 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 tradeoffs can 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.

14.4.2 Simulation, Test and Evaluation Process (STEP)

The STEP has been proposed in DoD guidance of the recent past and is still a valid concept. In STEP, simulation and test are integrated, each depending on the other to be effective and efficient. Simulations provide predictions of the system's performance and effectiveness, while tests are part of a strategy to provide information regarding risk and risk mitigation, to provide empirical data to validate models and simulations, and to determine whether systems are operationally effective, suitable, and survivable for intended use. A by-product of this process is a set of models and simulations with a known degree of credibility providing the potential for reuse in other efforts (Figure 14-4).

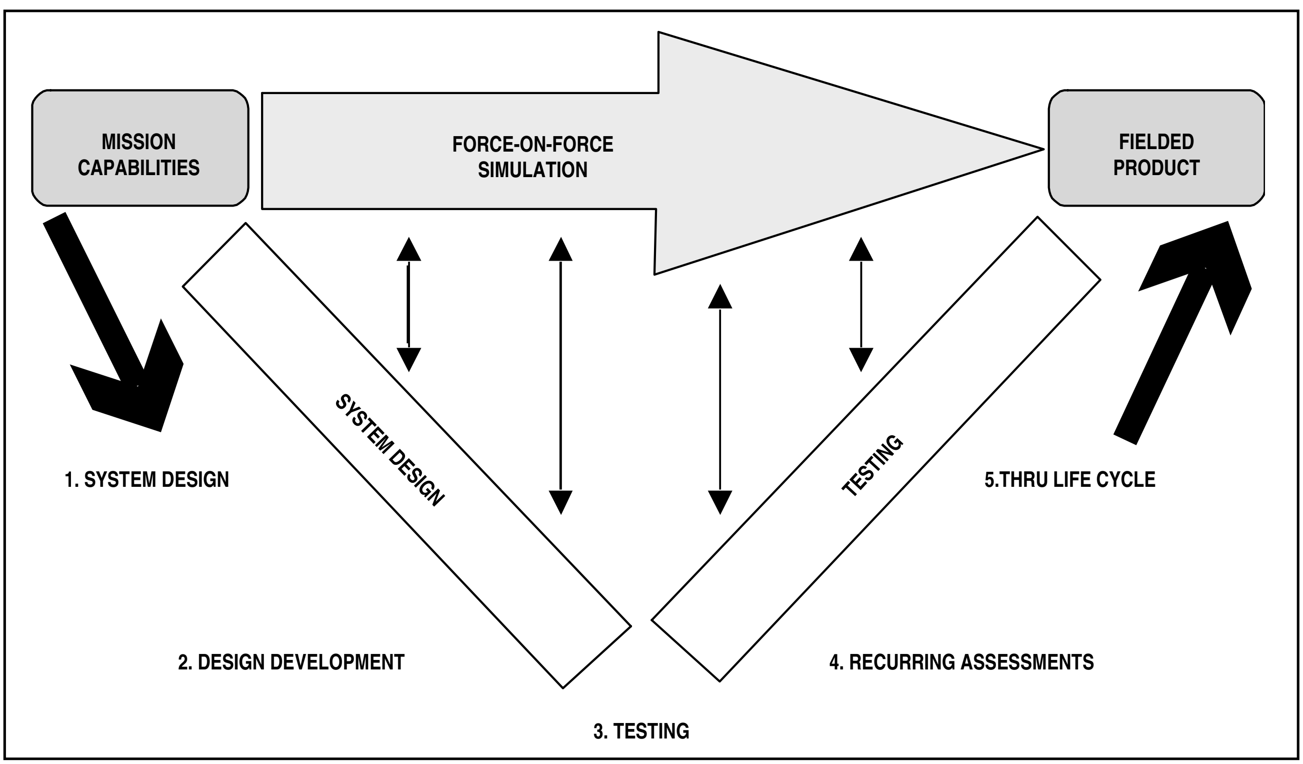


Figure 14-4. STEP Process

STEP is driven by mission and system requirements. The product of STEP is information. The information supports acquisition program decisions regarding technical risk, performance, system maturity, operational effectiveness, suitability, and survivability. STEP applies to all acquisition programs, especially Major Defense Acquisition Programs (MDAPs) and Major Automated Information Systems (MAISs).

Throughout STEP, tests are conducted to collect data for evaluating the system and refining and validating models. Through the model-test-model iterative approach, the sets of models mature, culminating in accurate representations of the system with appropriate fidelity, which can be used to predict system performance and to support the acquisition and potentially the training communities.

1. STEP begins with the Missions Needs Analysis and continues through the life cycle. Top-level requirements are used to develop alternative concepts and select/develop digital models that are used to evaluate theater/campaign and mission-/battle-level simulations. Mission-/battle-level models are used to evaluate the ability of a multiple platform force package to perform a specific mission. Mission and functional requirements continue to be refined, and the system reaches the preliminary design stage.

2. M&S is used both as a predictive tool and with test in an iterative process to evaluate the system design. The consequences of design changes are evaluated and help translate the most promising design approach into a stable, interoperable, and cost-effective design.

3. System components and subsystems are tested in a laboratory environment. Data from this hardware are employed in the model-test-model process. M&S is used in the planning of tests to support a more efficient use of resources. Simulated tests can be run on virtual ranges to conduct rehearsals and determine if test limitations can be resolved. STEP tools are used to provide data for determining the real component or subsystem's performance and interaction with other components. M&S is used during both Development Testing (DT) and Operational Testing (OT) to increase the amount of data and supplement the live test events that are needed to meet test objectives.

4. Periodically throughout the acquisition process the current version of the system under development should be reexamined in a synthetic operational context to reassess its military worth. This is one of the significant aspects of STEP, understanding the answer to the question: What difference does this change make in the system's performance?

5. STEP does not end with fielding and deployment of a system, but continues to the end of the system's life cycle. STEP results in a thoroughly tested system with performance and suitability risks identified. A by-product is a set of models and simulations with a known degree of credibility with the potential for reuse in other efforts. New test data can be applied to models to incorporate any system enhancements and further validate its models.

14.5 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. The test director 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 pre-test 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), New Mexico, 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 non-testable 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. These effects can be simulated repeatedly with no environmental impacts, costing only for the simulation operations.

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 positive or negative synergistic effect on test results, and could allow evaluation more quickly through a credible simulation designed with the flexibility to replicate operational variations.

14.6 SUPPORT TO ANALYSIS AND TEST REPORTING

M&S may be used in post-test analysis to extend and generalize results and to extrapolate to other conditions. The difficulty of instrumentation 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.

14.7 SIMULATION INTEGRATION

Simulations have become so capable and widespread in their application, and the ability to network dissimilar and/or geographically separate simulators in real time to synergistically simulate more than ever before, that whole new applications are being discovered. Simulations are no longer stove-piped tools in distinct areas, but are increasingly crossing disciplines and different uses. The F-35 Joint Strike Fighter (JSF) program uses nearly 200 different models and simulations throughout its many development and testing activities, in both government centers and contractor facilities. Increasingly, operational issues are being determined through operational analysis simulations, which will subsequently impact OT much later in a program. Elements of the T&E community should be involved increasingly earlier in every program to insure such results are compatible with the OT objectives and requirements. For example, the F-35 program networked eight different war game simulations together to develop the initial operational requirements that subsequently became documented in their Operational Requirements Document (ORD) for the many different required Service applications, i.e., Navy, Air Force, Marine Corps, and British Royal Navy. Only through this extensive "Virtual Strike Warfare Environment" were the broad and concurrent range of operational requirements able to be evaluated and established. New capabilities were achieved, but the extensive multi-simulation network demanded new techniques for establishing the technical confidence in the results.

14.8 SIMULATION PLANNING

With M&S becoming increasingly more complex, more expensive, and more extensive, testers must be thorough in planning their use and subsequent technical confidence. Testers are expected to be involved increasingly earlier, including the Operational Test Agencies (OTAs), if the subsequent T&E results are to be accepted. Simulation Support Plans (SSPs) are program documents that span the many simulations, their purpose, and their expected credibility. The Army, Air Force, and Marine Corps have policies requiring the early establishment of an SSP process, and stand-alone SSP documents. Usually this starts with a program office-level simulation support group of Service M&S experts who advise the program on M&S opportunities, establish the Verification, Validation, and Accreditation (VV&A) process, and document these procedures in the SSP, which extends across the life cycle of the system development, testing, and employment. It is vital that the SSP be fully coordinated with the Test and Evaluation Master Plan (TEMP). Without early planning of what each model or simulation is for, their expense, impact on system design as well as testing, earlier programs have ended up with a volume of different data for different purposes, which could not be analyzed, as subsequently found not to be credible, or delayed testing or program development.

14.9 SUMMARY

M&S 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 M&S 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. It takes appropriate time and funding to bring models and simulations along to the point that they are useful during an acquisition.

ENDNOTES

1. DoDI 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003.

2. Watt, Charles K. *The Role of Test Beds in C^3I*, Armed Forces Communications and Electronics Association (AFCEA) 35th Annual Convention, June 15-17, 1982.

3. DoD Directive (DoDD) 5000.59, *DoD Modeling and Simulation Management*, January 20, 1998, and DoDI 5000.61, *DoD Modeling and Simulation (M&S) Verification, Validation and Accreditation (VV&A)*, May 13, 2003.

15

TEST RESOURCES

15.1 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 the *Defense Acquisition Guidebook*, 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.[1] 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 planning and conduct shall take full advantage of existing investment in DoD [Department of Defense] ranges, facilities, and other resources, wherever practical, unless otherwise justified in the Test and Evaluation Master Plan [TEMP]."

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.

15.2 OBTAINING TEST RESOURCES

15.2.1 Identify Test Resources and Instrumentation

As early as possible, but not later than program start, the test facilities and instrumentation requirements to conduct program Test and Evaluation (T&E) should be identified and a tentative schedule of test activities prepared. This information is recorded in the TEMP and Service test resource documentation.

15.2.2 Require Multi-Service OT&E

Multi-Service Operational Test and Evaluation (MOT&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 low rate production decision.

15.2.3 Military Construction Program Facilities

Some programs cannot be tested without Military Construction Program (MCP) 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.

15.2.4 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.

15.2.5 Test Termination

Testers should not hesitate to terminate a test before its completion if it becomes clear that the main objective of the test has already been achieved, is unachievable (due to hardware failure, unavailability of resources, etc.), or if additional samples will not change the outcome and conclusions of the test.

15.2.6 Budget for Test

The Acquisition Strategy, TEMP, and budgeting documents should be reviewed regularly to ensure that there are adequate identified testing funds relative to development and fabrication funds.

The Acquisition Strategy, 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.)

15.2.7 Test Articles

A summary of important test planning items that were identified by the Defense Science Board (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.

15.2.8 Major Range and Test Facility Base (MRTFB)

The National Defense Authorization Act (NDAA) of Fiscal Year (FY) 2003 directed the DoD establish a Test Resource Management Center (DTRMC), with the director (three star or senior civilian) reporting directly to the Under Secretary of Defense for Acquisition, Technology, and Logistics (USD(AT&L)). The DTRMC provides oversight for T&E strategic planning and budgets for the Department's T&E activities, including the MRTFB, Central Test and Evaluation Investment Program (CTEIP), and the T&E Science and Technology (S&T) Program.

All Services operate ranges and test facilities for test, evaluation, and training purposes. These activities constitute the *DoD Major Range and Test Facility Base*.[2] This MRTFB is described as "a national asset which shall be sized, operated, and maintained primarily for DoD T&E support missions, but also is available to all users having

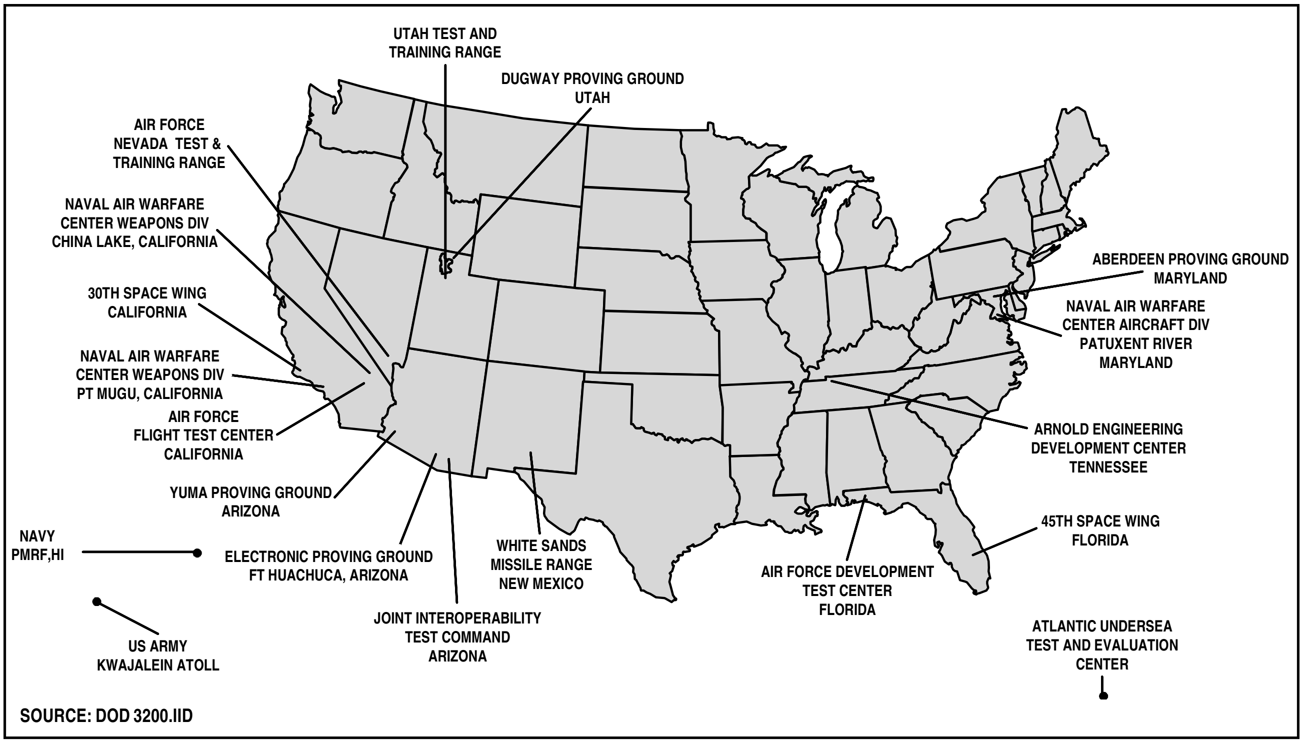


Figure 15-1. DoD Major Range and Test Facility Base

a valid requirement for its capabilities. The MRTFB consists of a broad base of T&E activities managed and operated under uniform guidelines to provide T&E support to DoD Components responsible for developing or operating materiel and weapon systems."[3] The list of MRTFB activities and their locations are shown on Figure 15-1. Summaries of the capabilities of each of these activities (with points of contact listed for further information) may be found in *Major Range and Test Facility Base Summary of Capabilities*.[4]

The MRTFB facilities are available for use by all the Services, other U.S. government 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.[5] The Director, Test Ranges Management Center 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 DTRMC sponsors a Joint Test Assets Database, which lists MRTFB and Operational Test Agency (OTA) test facilities, test area and range data, instrumentation, and test systems.

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 *Universal Documentation System Handbook*.[6] The requirements must be stated in the 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.

15.2.9 Central Test and Evaluation Investment Program (CTEIP)

In 1994 the Principal Deputy Under Secretary of Defense for Acquisition and Technology

(PDUSD(A&T)) chartered the Director, Test, Systems Engineering, and Evaluation (DTSE&E) to establish and chair a steering group that would oversee the acquisition and integration of all training and associated test range instrumentation and develop related policy. As a result of the reorganization of the Office of the Secretary of Defense (OSD) T&E functions, the Director, Operational Test and Evaluation (DOT&E) and subsequently the DTRMC, manages the implementation of the Joint Training and Test Range Roadmap (JTTR) and executes the CTEIP. The CTEIP provides OSD funding and a mechanism for the development and acquisition of new test capabilities to satisfy multi-Service testing requirements.

15.2.10 Service Test Facilities

There are other test resources available besides MRTFB. Frequently, National Guard or Reserve units, commercial or international test facilities, and war reserve assets are available to support DoD T&E. Testers can determine resources available by contacting their Service headquarters staff element. Within the Army, consult documents include: Army Test and Evaluation Command (ATEC)'s database on existing Army major test facilities, major instrumentation, and test equipment; the Project Manager for Instrumentation, Targets, and Threat Simulators (PM, ITTS) Targets Information Manual for a descriptive catalogue of Army targets and foreign ground assets available (or in development) for support of T&E and training; and, PM ITTS' Threat Inventory Database on available assets for both hardware simulators and software simulation systems. 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.

15.3 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.

15.3.1 TEMP Resource Requirements

The Program Manager (PM) 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 first TEMP must be prepared for program initiation, initial test resource planning must be accomplished very early. Refinements and reassessments of test resource requirements are included in each TEMP update. The guidance for the content of the test resource summary (Part V) of the TEMP is in Appendix 2, Test and Evaluation Master Plan, of the *Defense Acquisition Guidebook* (see Table 15-1). Once identified, the PM must then work within the Service headquarters and range management structure to assure the assets are available when needed.

15.3.2 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), test scenarios, and specific test resource requirements.

Table 15-1. TEMP Test Resource Summary Section

PART V – TEST AND EVALUATION RESOURCE SUMMARY

PROVIDE A SUMMARY (PREFERABLY IN A TABLE OR MATRIX FORMAT) OF ALL KEY TEST AND EVALUATION RESOURCES, BOTH GOVERNMENT AND CONTRACTOR, THAT WILL BE USED DURING THE COURSE OF THE ACQUISITION PROGRAM.

THE TEMP SHOULD PROJECT THE KEY RESOURCES NECESSARY TO ACCOMPLISH DEVELOPMENTAL AND VALIDATION TESTING AND OPERATIONAL TEST AND EVALUATION. THE TEMP SHOULD ESTIMATE, TO THE DEGREE KNOWN AT MILESTONE B, THE KEY RESOURCES NECESSARY TO ACCOMPLISH DEVELOPMENTAL TEST AND EVALUATION, LIVE FIRE TEST AND EVALUATION, AND ALL OPERATIONAL TEST AND EVALUATION. THESE SHOULD INCLUDE ELEMENTS OF THE NATIONAL TEST FACILITIES BASE (WHICH INCORPORATES THE MAJOR RANGE AND TEST FACILITY BASE (MRTFB), CAPABILITIES DESIGNATED BY INDUSTRY AND ACADEMIA, AND MRTFB TEST EQUIPMENT AND FACILITIES), UNIQUE INSTRUMENTATION, THREAT SIMULATORS, AND TARGETS. AS SYSTEM ACQUISITION PROGRESSES, THE PRELIMINARY TEST RESOURCE REQUIREMENTS SHALL BE REASSESSED AND REFINED AND SUBSEQUENT TEMP UPDATES SHALL REFLECT ANY CHANGED SYSTEM CONCEPTS, RESOURCE REQUIREMENTS, OR UPDATED THREAT ASSESSMENT. ANY RESOURCE SHORTFALLS THAT INTRODUCE SIGNIFICANT TEST LIMITATIONS SHOULD BE DISCUSSED WITH PLANNED CORRECTIVE ACTION OUTLINED. SPECIFICALLY, IDENTIFY THE FOLLOWING TEST RESOURCES:

- TEST ARTICLES
- TEST SITES AND INSTRUMENTATION
- TEST SUPPORT EQUIPMENT
- THREAT REPRESENTATION
- TEST TARGETS AND EXPENDABLES
- OPERATIONAL FORCE TEST SUPPORT
- SIMULATORS, MODELS, AND TEST-BEDS
- SPECIAL REQUIREMENTS
- TEST AND EVALUATION FUNDING REQUIREMENTS
- MANPOWER/PERSONNEL TRAINING

Source: *Defense Acquisition Guidebook*, September 2004.

15.3.2.1 Army Test Resource Planning

In the Army, the independent evaluator, with the assistance of the developmental and operational testers, prepares input to the TEMP and develops the System Evaluation Plan (SEP), the primary planning documents for integrated T&E of the system. These documents should be prepared early in the acquisition cycle (at the beginning of system acquisition activities). They describe the entire T&E strategy including critical issues, test methodology, MOEs, and all significant test resources. The TEMP and SEP 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 ATEC 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, ATEC, and consists of a general officer or equivalent representatives from the Army staff and major commands. The TSARC meets semiannually 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 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 Headquarters, Department of the Army (HQDA) Deputy Chief of Staff for G-3 (DCS G-3), 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 National Aeronautics and Space Administration (NASA)), or a contractor are required, the request is coordinated by ATEC. For example, the request for a range must be made at least 2 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 United States, such as in Canada, Germany, or other North Atlantic Treaty Organization (NATO) countries, is also handled by ATEC.

15.3.2.2 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 PM and Operational Test Director (OTD) refer to documents such as the Mission Need Statement (MNS), Acquisition Strategy, Navy Decision Coordinating Paper (NDCP), Operational Requirement Document (ORD), threat assessments, Secretary of the Navy Instruction (SECNAVINST) 5000.2B, and the *Operational Test Director's Guide* (Commander, Operational 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 OTD. The TEMP defines all T&E (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 5 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 by letter or message to coordinate with the Operational Test and Evaluation Force (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 OPTEVFOR is authorized direct liaison with other Service-independent 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 (IR&D).

15.3.2.3 Air Force Test Resource Planning

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

During the Advanced Planning Phase of a weapon system acquisition (5 to 6 years before OT&E), AFOTEC prepares the OT&E section of the first full TRP, coordinates the TRP with all supporting organizations, and assists the Resource Manager (RM) in programming required resources. The resource requirements listed in the Resource Information Network TRP are developed by the test manager, RM, and test support group, using sources such as the ORD and threat assessments. The TRP should specify, in detail, all the resources necessary to successfully conduct a test once it is entered in the Test Resource Information Management System (TRIMS).

The TRP 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 TRP, it is likely the resource will not be available for the test. The TRP is revised and updated on a continuous basis, since the test resource requirements become better defined as the OT&E plans mature. The initial TRP serves as a baseline for comparison of planned OT&E resources with actual expenditures. Comparisons of the initial TRP with subsequent updates provide an audit trail of changes in the test program and its testing requirements. The AFOTEC maintains all TRPs on TRIMS; this permits immediate response to all queries regarding test resource requirements.

The AFOTEC/RM consolidates the resource requirements from all TRPs 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.

All requests for test resources are coordinated by HQ AFOTEC as part of the TRP 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 (JTCG) 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 RM 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 RM and the TM 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 Air Force, Chief of Staff/Directorate of Test and Evaluation (AF/TE) and based on formal Memoranda of Understanding (MOU). The U.S. Air Force-Europe/Directorate of Operations-Operations (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.

15.4 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 (PE): R&D 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, sub-claimants, administering offices, commanding generals, etc.). (See DoD 7000.14-4, *Financial Management Regulation*, Vol. 2A.)

The Planning, Programming, Budgeting and Execution (PPBE) system is a DoD internal system used to develop input to the Congress for each year's budget while developing future-year budgets. The PPBE is calendar oriented. There are concurrent 2-year PPBE cycles ongoing at one time. These cycles are: planning, programming, budgeting, and execution. At any one time there are three budgets being worked by the Services. The current 2-year budget is being executed. The next 6 years of defense planning are being programmed, and long-range program plans and planning guidance are being reviewed for updating.

There are various types of funding in the PPBE: R&D funding for maintaining the technology base; exploratory development funding for conducting the concept assessments; advanced development funding for conducting both the concept development and the early prototyping; engineering development funding for demonstrating the Engineering Development Model (EDM); procurement funding for conducting LRIP, Full Rate Production (FRP), system deployment and operational support. RDT&E management and support funding is used throughout the development 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 R&D intended to improve performance, including test items, DT&E and test support of OT&E of the system, or equipment test items.

Funding that is planned, programmed, and budgeted through the PPBE 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); 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 O&M appropriations. Testing of Product Improvements (PIs), block upgrades, and major modifications is funded from the same appropriations as the program development. Follow-on Test and Evaluations (FOT&Es) 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:

Development Test (DT) Funding. Funds required for conduct of engineering and development tests are programmed and budgeted by the materiel developer, based upon the requirements of the TEMP. These costs may include, 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 PM 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. The Air Force funds OT&E separate from the program office through a dedicated PE for AFOTEC conducted operational testing.

15.4.1 Army Funding

Test resources are developed and funded under various Army appropriations. The AMC 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 the TRADOC or FORSCOM. Weapon system PMs use RDT&E funds to reimburse these supporting commands for costs directly related to their tests. Weapon system materiel developers are 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 PM ITTS threat simulators program—many weapon system developers still have to fund the cost of new threat systems that are specifically needed to test their weapon system. Funding for Army operational testing is through the PM's Program Element (PE) and is given to ATEC for direct control of funds for each program. Funding requirements are developed in consonance with the Outline Test Plan (OTP).

15.4.2 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 the Office of the Chief of Naval Operations (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.

15.4.3 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 SPO to the designated test agency. The RDT&E appropriation funds the cost associated with R&D, 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 PE and has direct control of OT&E funds for all programs. The IOT&E manager prepares a TRP that summarizes the resource requirements for IOT&E and related test support. All pre-test 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.

15.5 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 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 PPBE 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. 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. The Air Force OTA develops and directly controls OT&E funds.

ENDNOTES

1. *Defense Acquisition Guidebook*, http://akss.dau.mil/DAG/.

2. Department of Defense Directive (DoDD) 3200.11, *DoD Major Range and Test Facility Base*, May 1, 2002.

3. DoD 5000.3-M-1, *Test and Evaluation Master Plan (TEMP) Guidelines*, October 1986. Cancelled.

4. DoD 3200.11-D, *Major Range and Test Facility Base Summary of Capabilities*, June 1983.

5. *Ibid.*

6. Range Commanders Council (RCC), *Universal Documentation System Handbook*, Document 501-84, Vol. 3, August 1989.

16

TEST AND EVALUATION MASTER PLAN

16.1 INTRODUCTION

Guidance contained in *Operation of the Defense Acquisition System* and the *Defense Acquisition Guidebook* stipulates that a Test and Evaluation Master Plan (TEMP) format shall be used for all Acquisition Category (ACAT) I and IA or Office of the Secretary of Defense (OSD)-designated oversight acquisition programs.[1] This reinforces the philosophy that good planning supports good operations. For effective engineering development and decision-making processes, an early Test and Evaluation (T&E) strategy in the Technology Development Strategy (TDS) must be evolved into an overall integrating master plan detailing the collection and evaluation of test data on required performance parameters. Less than ACAT I programs are encouraged to tailor their T&E strategy using the TEMP format as a guide. The TEMP relates program schedule, test management strategy and structure, and required resources to: Critical Operational Issues (COIs); Critical Technical Parameters (CTPs); Minimum Acceptable Values (MAVs) (thresholds); acquisition strategy; 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.

16.2 TEMP DEVELOPMENT

The development of system T&E strategy begins during the Technology Development (TD) effort with the test planning in the TDS document. This evolves as the system is better defined and T&E events are consolidated in the TEMP. Effective management of the various test processes are the primary functions of a Program Management Office (PMO) T&E Working-level Integrated Product Team (WIPT). The T&E strategy is highly contingent on early system concept(s) that are deemed appropriate for satisfying user requirements. As outlined in *The Defense Acquisition System*,[2] the priority for selecting a solution is:

(1) A non-materiel solution, such as changes to Doctrine, Organization, Training, Leadership and education, Personnel, and Facilities (DOT_LPF);

(2) A Materiel (M) alternative, chosen according to the following sequence:

(a) The procurement or modification of commercially available products, services, and technologies from domestic or international systems, or the development of dual-use technologies.

(b) The additional production or modification of previously developed U.S. and/or allied military systems or equipment.

(c) A cooperative development program with one or more allied nations.

(d) A new, joint, Department of Defense (DoD) Component or government agency development program.

(e) A new DoD Component-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 versus 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 (T&E Integrated Product Teams (IPTs)) and reviews by panels of senior T&E management officials— "gray beards." As the tempo of actual test activities begins to build concept to prototype to Engineering Development Model (EDM) to pre-Low Rate Initial Production (LRIP), internal T&E management staff are needed to control the processes and evaluate results.

16.2.1 Program Management Office Responsibilities

The 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. This task is a "difficult do" prior to program start since some Services do not formulate or staff a PMO until formal program initiation. An additional complicating factor is the nebulous condition of other program source documents (Capability Development Document (CDD), Technical Management Plan (TMP) , Acquisition Strategy, System Threat Assessment (STA), Logistics Support Plan (LSP), etc.) that are also in early stages of development/updating for the milestone review. Since the TEMP must conform to the acquisition strategy and other program management documents, it frequently lags in the development process and does not receive the attention needed from PMO or matrix support personnel. PMO emphasis on early formulation of the test planning teams (T&E WIPT) 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 exceeds the "streamlined" approval process signatories; however, it must be coordinated. An early start in getting Service-level concurrence is important so the milestone decision 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.

16.2.2 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 an incremental or spiral approach with moderate- to low-risk technologies that should reduce the intensity and duration of the T&E program. It does, however, include a requirement for post-production test activities as the system is modified to accommodate previously unknown new technologies, new threats, or other performance enhancements.

A revolutionary acquisition strategy (single step) 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 workload increases in direct proportion to the technology risk and difficulty in fixing problems. There is a much higher potential for extended schedules with iterative test-fix-test cycles.

16.2.3 General T&E Planning Issues

The Defense Science Board (DSB) report presented guidance on T&E at two levels.[3] 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.

16.2.3.1 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. 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.

16.2.3.2 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;
- Not be repeated if failures occur without first conducting a detailed analysis of the failure.

16.2.3.3 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.

16.2.3.4 Requirement for Test Rehearsals

Test rehearsals should be conducted for each new phase of testing.

16.2.4 Scheduling

Specific issues associated with test scheduling are listed below.

16.2.4.1 Building Block Test Scheduling

The design of a set of tests to demonstrate feasibility prior to testing the system-level EDM 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.

16.2.4.2 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.

16.2.4.3 Phasing of DT&E and IOT&E

Problems that become apparent in operational testing can often be evaluated faster with the instrumented DT&E hardware. The Integrated Test Plan (ITP) should provide time and money to investigate test failures and eliminate causes of failures before other, similar tests take place.

16.2.4.4 Schedule IOT&E to Include System Interfaces with Other Systems

Whenever possible, the IOT&E/FOT&E (Follow-on Operational Test and Evaluation) of a weapon system should be planned to include other systems that must have a technical interface with the new system. For example, missiles should be tested on most of the platforms for which they are programmed.

A Preplanned Product Improvements (P^3I) or increment strategy is a variant of the evolutionary development process in which PMs and other acquisition managers recognize the high-risk technologies/subsystems and put them on parallel development tracks. The testing strategy should anticipate the requirements to evaluate P^3I item technology maturity and then test the system during the integration of the additional capability.

Advanced Technology Demonstrations (ATD) or Advanced Concept Technology Demonstrations (ACTD) may provide early insights into available technologies for incorporation into developmental or mature, post-prototype systems. Using proven, mature technology provides a lower-risk strategy and may significantly reduce the development testing workload. To assess and manage risk, PMs and other acquisition managers should use a variety of techniques, including technology demonstrations, prototyping, and T&E. The process for verifying contract performance and item specifications, T&E of threshold performance requirements in the Operational Requirements Document (ORD), exit criteria, or the Acquisition Program Baseline (APB) 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. OT&E 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 an overshadowing evaluation plan, outlining methodologies, models, simulations, and test data required at periodic decision points.

The TEMP should: (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.

A reliability growth (Test, Analyze, Fix, and Test (TAFT)) program should be developed to satisfy the reliability levels required at Full Rate Production (FRP). Reliability tests and demonstrations will be based on actual or simulated operational conditions.[4]

Maintainability will be verified with a maintainability demonstration before FRP.[5]

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 objective will be validated.

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.

The TEMP will address health hazard and safety critical issues to provide data to validate the results of system safety analyses.

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.

The TEMP shall provide a roadmap for integrated simulation, test, and evaluation plans, schedules, and resource requirements necessary to accomplish the T&E program. T&E planning should address Measures of Effectiveness/Suitability (MOEs/MOSs) 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 threshold values specified in the Capability Production Document (CPD).

16.3 TEMP FORMAT

The format specified in the *Defense Acquisition Guidebook*, Appendix 2, is required for all acquisition category I, IA, and OSD-designated oversight programs (Table 16-1). 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 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., Missile Defense Agency (MDA), Family of Tactical Vehicles, Future Combat Systems (FCSs)). 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 milestone decision process to assess the adequacy of planned T&E, 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 16-1.

Table 16-1. Test and Evaluation Master Plan Format

PART I	SYSTEM INTRODUCTION	
		MISSION DESCRIPTION
		SYSTEM DESCRIPTION
		SYSTEM THREAT ASSESSMENT
		MEASURES OF EFFECTIVENESS AND SUITABILITY
		CRITICAL TECHNICAL PARAMETERS
PART II	INTEGRATED TEST PROGRAM SUMMARY	
		INTEGRATED TEST PROGRAM SCHEDULE
		MANAGEMENT
PART III	DEVELOPMENT TEST AND EVALUATION OUTLINE	
		DEVELOPMENT TEST AND EVALUATION OVERVIEW
		FUTURE DEVELOPMENTAL TEST AND EVALUATION LIMITATIONS
PART IV	OPERATIONAL TEST AND EVALUATION OUTLINE	
		OPERATIONAL TEST AND EVALUATION OVERVIEW
		CRITICAL OPERATIONAL ISSUES
		FUTURE OPERATIONAL TEST AND EVALUATION LIMITATIONS
		LIVE FIRE TEST AND EVALUATION
PART V	TEST AND EVALUATION RESOURCE SUMMARY	
		TEST ARTICLES
		TEST SITES AND INSTRUMENTATION
		TEST SUPPORT EQUIPMENT
		THREAT REPRESENTATION
		TEST TARGETS AND EXPENDABLES
		OPERATIONAL FORCE TEST SUPPORT
		SIMULATIONS, MODELS, AND TEST BEDS
		SPECIAL REQUIREMENTS
		TEST AND EVALUATION FUNDING REQUIREMENTS
		MANPOWER/PERSONNEL TRAINING
APPENDIX A		BIBLIOGRAPHY
APPENDIX B		ACRONYMS
APPENDIX C		POINTS OF CONTACT
ATTACHMENTS		(AS APPROPRIATE)

SOURCE: *Defense Acquisition Guidebook*, September 2004.

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, and Maintainability (RAM), logistics objectives and requirements, and major decision points. It should summarize the testing accomplished to date and explain the relationship of the various models and simulations, subsystem tests, integrated system development 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 TEMP 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 FOT&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 stresses (as early as possible) the weapon system in a variety of realistic environments?

16.4 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 manhours and PM guidance. The interactions of the various T&E players and support agencies in the T&E WIPT must be woven into the fabric of the total system acquisition strategy. Cost and schedule implications must be negotiated to ensure a viable T&E program that provides timely and accurate data to the engineering and management decision makers.

ENDNOTES

1. DoD Instruction (DoDI) 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003, and *Defense Acquisition Guidebook,* http://akss.dau.mil/DAG.

2. DoD Directive (DoDD) 5000.1, *The Defense Acquisition System*, May 12, 2003.

3. Defense Science Board, *Report of Task Force on Test and Evaluation*, April 1, 1974.

4. DoD 3235.1-H, *Test and Evaluation of System Reliability, Availability and Maintainability – A Primer*, March 1982.

5. *Ibid.*

V

MODULE

SPECIALIZED TESTING

Many program managers face several T&E 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 program managers when they develop the integrated strategy for the T&E program.

17

SOFTWARE SYSTEMS TESTING

17.1 INTRODUCTION

Software development presents a major development risk for military weapons and National Security Systems (NSS). Software is found in Major Automated Information Systems (MAISs) and weapon system software. High-cost Software systems, such as personnel records management systems, financial accounting systems, or logistics records, that are the end-item solution to user requirements, fall in the MAIS category. Performance requirements for the MAIS typically drive the host hardware configurations and are managed by the Information Technology Acquisition Board (ITAB) chaired by the Assistant Secretary of Defense (Networks and Information Integration (ASD(NII))/Chief Information Officer (CIO). The Director, Operational Test and Evaluation (DOT&E) is a principal member of the ITAB.

Software developments, such as avionics systems, weapons targeting and control, and navigation computers, that are a subset of the hardware solution to user requirements fall in the weapon-system software category. Performance requirements for the system hardware are flowed down to drive the functionality of the software resident in onboard computers. The effectiveness of the weapon system software is reviewed as part of the overall system review by the Defense Acquisition Board (DAB), chaired by the Under Secretary of Defense for Acquisition, Technology, and Logistics (USD(AT&L)). The DOT&E is a principal member and the Deputy Director, Development, Test and Evaluation (DT&E) is an advisor to the DAB. No Milestone A, B, or Full Rate Production (FRP) decision (or their equivalent) shall be granted for a MAIS until the Department of Defense (DoD) CIO certifies that the MAIS program is being developed in accordance with the Clinger-Cohen Act provisions.[1] For MDAP and MAIS programs, the DoD Component CIO's confirmation (for Major Defense Acquisition Programs (MDAPs)) and certification (for MAISs) shall be provided to both the DoD CIO and the Milestone Decision Authority (MDA).

Software development historically has escalated the cost and reduced the reliability of weapon systems. Embedded computer systems that are physically incorporated into larger weapon systems, have a major data processing function. Typically, the output of the systems is information, control signals, or computer data required by the host system to complete its mission. Although hardware and software often contribute in equal measure to successful implementation of system functions, there have been relative imbalances in their treatment during system development.

Automated Information Systems (AISs), once developed, integrated, and tested in the host hardware, are essentially ready for production. Software in weapon systems, once integrated in the host hardware, continues to be tested as a component of the total system and is not ready for production until the total system has successfully demonstrated required performance. Any changes to weapon system hardware configuration may stimulate changes to the software. The development of all software 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 may be erroneously specified, or later in the development cycle, when Systems Integration (SI) is implemented. This chapter addresses the use of testing to obtain insights into the development risk of AIS and weapon system software, particularly as it pertains to the software development processes.

17.2 DEFINITIONS

The term Automated Information System (AIS) is defined as a combination of computer hardware and software, data, or telecommunications that performs functions such as collecting, processing, transmitting, and displaying information.[2] Excluded are computer resources, both hardware and software, that are: physically part of, dedicated to, or essential in real time to the mission performance of weapon systems.[3]

The term weapon system software includes Automated Data Processing Equipment (ADPE), 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 weapon system;

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

Acquisition of software for DoD is described in Military Standard (MIL-STD)-498, *Software Development and Documentation*, which although rescinded has been waived for use until commercial standards such as EIA 640 or J-Std-016 (*Software Life Cycle Processes, Software Development*, September 1995) become the guidance for software development.[4] Guidance may also be found in the *Defense Acquisition Guidebook*.

17.3 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 17-1, errors may occur at the inception of the process. These errors may occur during requirements definition, when objectives may 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 (LCCs). 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 LCCs 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 Test and Evaluation (T&E) reveals that half of all maintenance costs are incurred in the correction of previously undetected errors. Approximately one-half of the operational LCCs 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

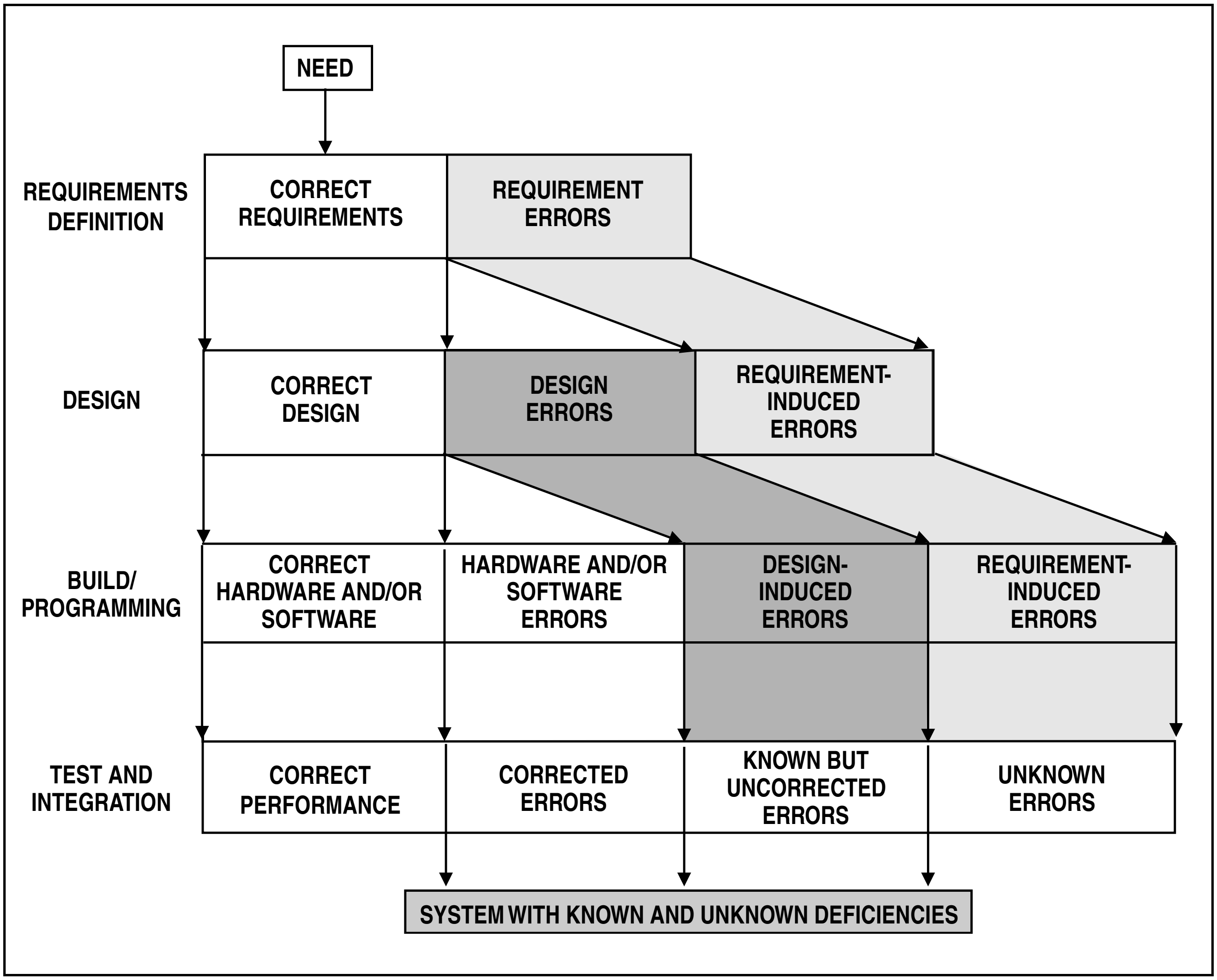


Figure 17-1. The Error Avalanche

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 T&E planning is critical to the successful development of a computer system.

17.4 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 a commercial 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 maturity metrics from the written specifications that will lead to 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.

17.5 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 (Figures 17-2, 17-3).

17.5.1 Testing Approach

The integration of software 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, non-computer-based methods of evaluating architectures, designs, and interfaces. It can consist of:

- **Inspections:** Programmers explain their 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 programmers of their own work. There is a low

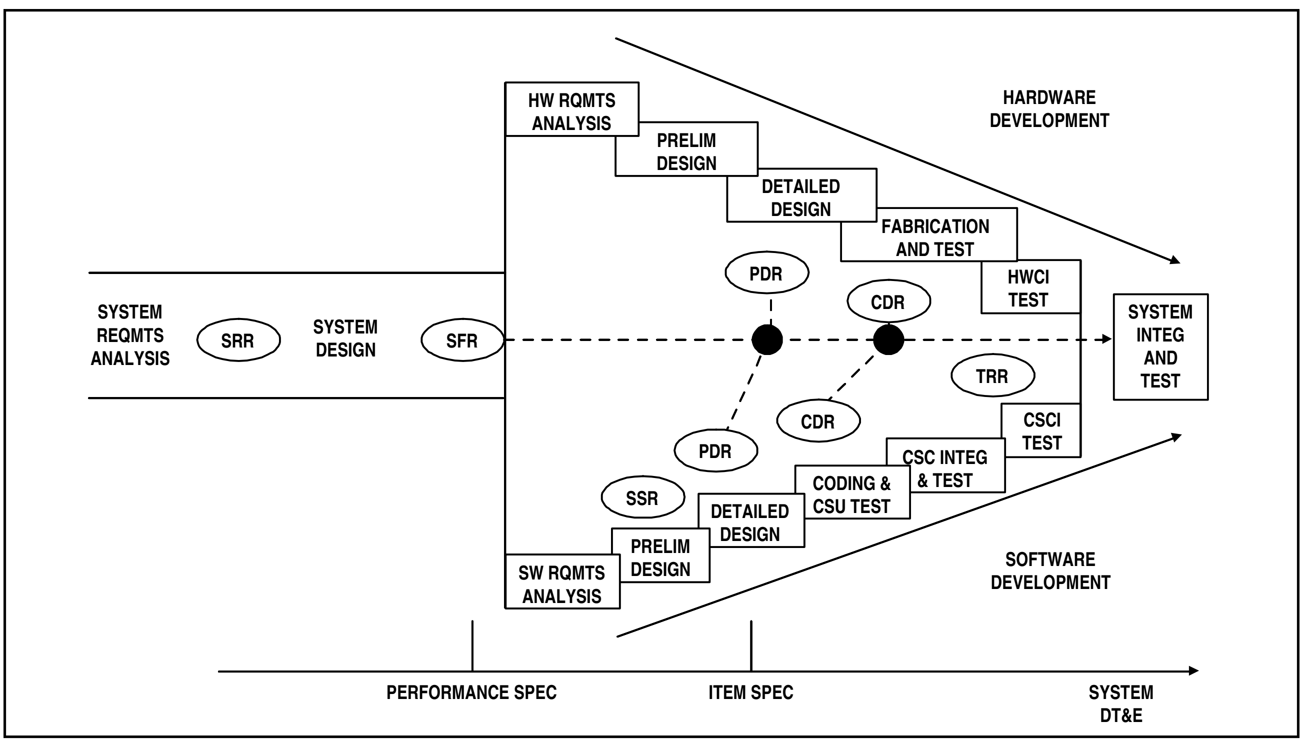


Figure 17-2. System Development Process

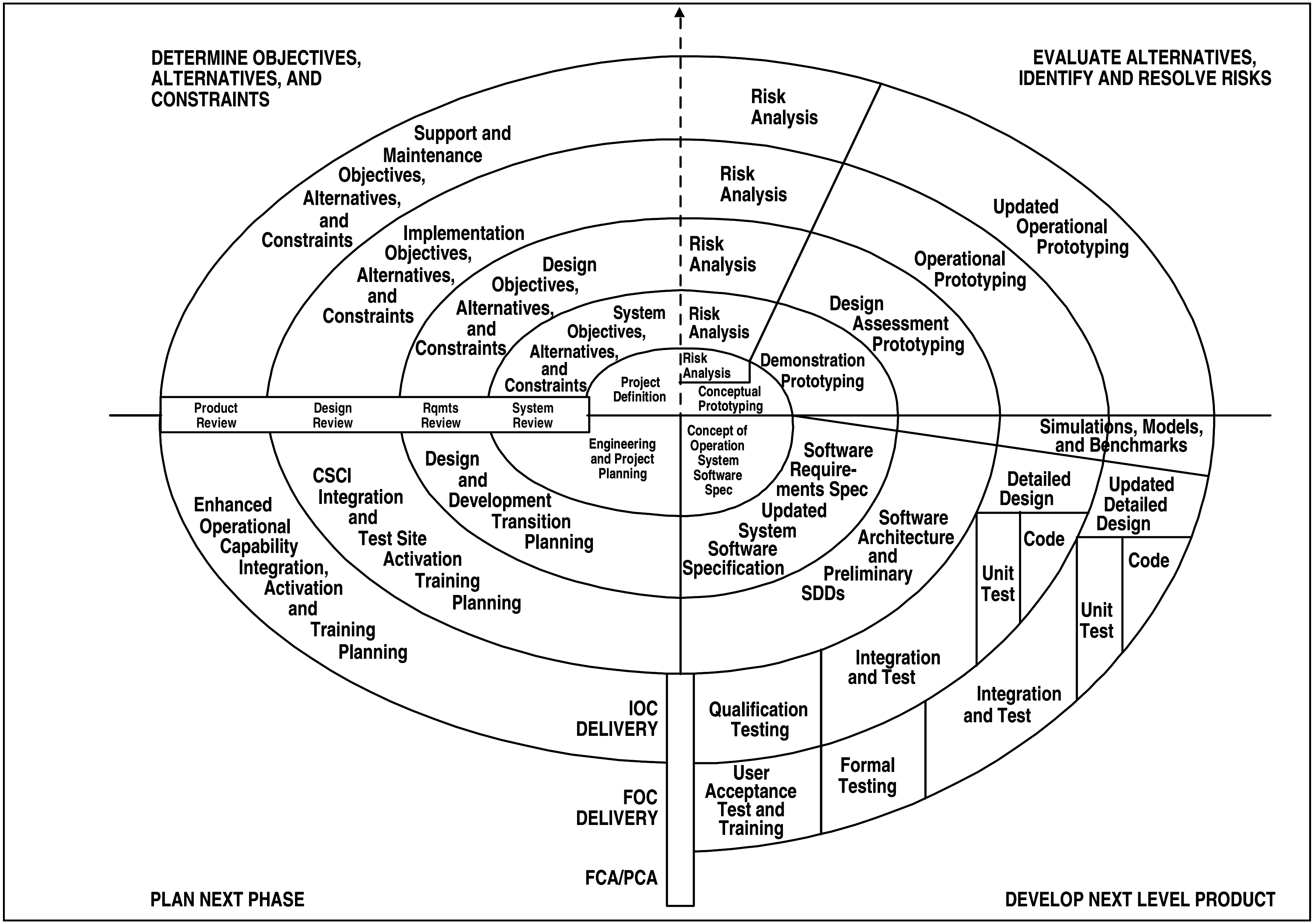


Figure 17-3. Spiral Model for AIS Development Process

probability of identifying their 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 the 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 (OA (Operational Assessment), IOT&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 personnel tasked for future maintenance of the software system should be brought on board to familiarize them with the system operating characteristics and documentation. Their resultant expertise should be included in the software T&E 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 objectives 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. Information assurance testing must be conducted on any system that collects, stores, transmits, or processes classified or unclassified information. In the case of Information Technology (IT) systems, including NSS, DT&E should support the DoD Information Technology Security Certification and Accreditation Process (ITSCAP) and Joint Interoperability Certification (JIC) processes.[5] The test environments must be chosen on a careful analysis of characteristics to be demonstrated and their relationship to the development, operational, and support environments. In addition, environments must be representative of those in which the software will be maintained. At Milestone C, for MAIS, the MDA shall approve, in coordination with the DOT&E, the quantity and location of sites for a limited deployment for IOT&E.[6]

17.5.2 Independent Verification and Validation

IV&V 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 (PM).

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 proceeds with minimal delays and misunderstandings. Frequent informal reviews, design and code walk-through and audits ensure that the programming standards, software engineering standards, Software Quality Assurance (SQA), 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.

17.6 SUMMARY

There is a useful body of software testing technologies that can be applied to testing of AIS and weapon system software. As a technical discipline, though, software testing is still maturing. There is a growing foundation of guidance documents to guide the PM in choosing one testing technique over another. One example is the U.S. Air Force Software Technology Support Center's (STSC's) *Guidelines for Successful Acquisition and Management of Software-intensive Systems*. The Air Force Operational Test and Evaluation Center (AFOTEC) has also developed a course on software OT&E. It is apparent that systematic T&E techniques are far superior to ad-hoc testing techniques. Implementation of an effective software T&E plan requires a set of strong technical and management controls. Given the increasing amount of AIS and weapon system software being acquired, there will be an increased emphasis on tools and techniques for software T&E. Another resource is the Software Program Manager's Network that publishes guides on Best Practices in Software T&E, Vol. I & II. (http://www.spmn.com)

ENDNOTES

1. DoD Instruction (DoDI) 5000.2, *Operation of the Defense Acquisition Program*, May 12, 2003.

2. DoD 5000.2-R, *Mandatory Procedures for Major Defense Acquisition Programs (MDAPs) and Major Automated Information System (MAIS) Acquisition Programs*, June 2001.

3. There is some indication that DoD Directive (DoDD) 8000.1, *Management of DoD Information Resources and Information Technology*, February 27, 2002, Defense Information Management Program, which provides guidance on AIS development, will be incorporated in a future change to the 5000 Series on acquisition management.

4. MIL-STD-498, *Software Development and Documentation*, December 4, 1994, rescinded; EIA 640 or J-Std-016 (*Software Life Cycle Processes, Software Development*, September 1995).

5. DoDD 4630.5, *Interoperability and Supportability of Information Technology (IT) and National Security Systems (NSS)*, May 5, 2004, and Chairman of the Joint Chiefs of Staff Instruction (CJCSI) 6212.01C, *Interoperability and Supportability of Information Technology and National Security Systems*, November 20, 2003.

6. DoDI 5000.2.

18

TESTING FOR VULNERABILITY AND LETHALITY

18.1 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 survivability and lethality of developing systems. (Table 18-1) 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*.[1]

18.2 LIVE FIRE TESTING

18.2.1 Background

In March 1984, the Office of the Secretary of Defense (OSD) chartered a joint T&E program designated "The Joint Live Fire Program." This program was to assess the vulnerabilities and lethality of selected U.S. and threat systems already fielded. The controversy over joint Live Fire Testing (LFT) 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 (NDAA) of Fiscal Year (FY) 87. This act required an OSD-managed LFT program for major acquisition programs fitting certain criteria. Subsequent amendments to legislative guidance have dictated the current program. The Department of Defense (DoD) implementation of congressional guidance in *Defense Acquisition Guidebook*, requires that "covered systems, major munitions programs, missile programs, or Product Improvements (PIs) to these" (i.e., Acquisition Category (ACAT) I and II programs) must execute survivability and lethality testing before Full Rate Production (FRP). The legislation dealing with the term "covered system" has been clarified as a system "that DOT&E, acting for the Secretary of Defense [SECDEF], has determined to be: a major system within the meaning of that term in Title 10, United States Code (U.S.C.) 2302(5) that is (1) user-occupied and designed to provide some degree of protection to its occupants in combat; or (2) a conventional

Table 18-1. Relationships Between Key Concepts

	PERSPECTIVE		
TERMINOLOGY	DEFENSIVE	OFFENSIVE	MEANING
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, GAO/PEMD-87-17, August 1987, page 15.

munitions program or missile program; or (3) a conventional munitions program for which more than 1,000,000 rounds are planned to be acquired; or (4) a modification to a covered system that is likely to affect significantly the survivability or lethality of such a system."

The Secretary of Defense (SECDEF) has delegated the authority to waive requirements for the full-up, system-level Live Fire Test and Evaluation (LFT&E) before the system passes the program initiation milestone, to the Under Secretary of Defense (Acquisition, Technology, and Logistics) [USD(AT&L)] for ACAT ID and the CAE for ACAT II programs, when it would be unreasonably expensive and impractical. An alternative vulnerability and lethality T&E program must still be accomplished. Programs subject to LFT or designated for oversight are listed on the OSD annual T&E oversight list. The DoD agent for management of the Live Fire Test program is the Director, Operational Test and Evaluation (DOT&E). 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 Full Rate Production Decision Review (FRPDR) and congressional committees. The Service-detailed Live Fire Test Plan must be reviewed and approved by the DOT&E, and LFT must be addressed in Part IV of the program Test and Evaluation Master Plan (TEMP). The OSD had previously published guidelines, elements of which have subsequently been incorporated into the latest revision to the 5000 Series *Defense Acquisition Guidebook*.

18.2.2 Live Fire Tests

There are varying types and degrees of live fire tests. The matrix in Table 18-2 illustrates the various possible combinations. Full-scale, full-up testing is usually considered to be the most realistic.

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.[2]

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

Table 18-2. 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)
SUB SCALE	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 Program*, General Accounting Office, GAO/PEMD-87-17, August 1987.

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. LFT 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).[3]

18.2.3 Use of Modeling and Simulation (M&S)

Survivability and lethality assessments have traditionally relied largely on the use of modeling and simulation techniques. The LFT Program does not replace the need for such techniques; in fact, the LFT Guidelines issued by OSD in May 1987 (Figure 18-1) required that no shots be conducted until pre-shot model predictions were made concerning the expected damage. Such predictions are useful for several reasons. First, they

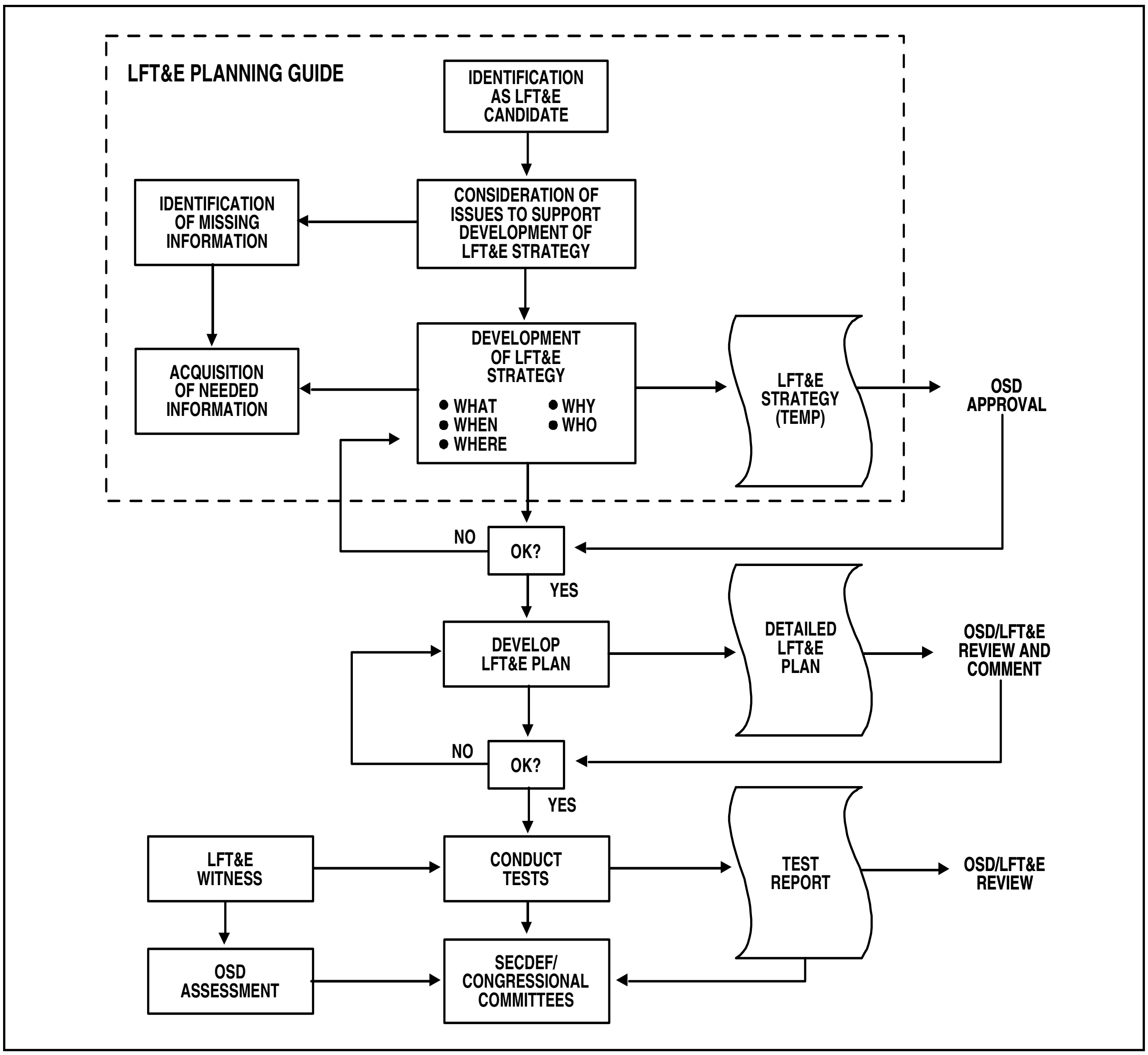


Figure 18-1. Live Fire Test and Evaluation Planning Guide

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 shot lines 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.

18.2.4 Live Fire Test Best Practices

The *Defense Acquisition Guidebook* states that plans for LFT must be included in the TEMP. Key points covered in the LFT guidelines include the following:

- The LFT&E Detailed T&E Plan is the basic planning document used by OSD and the Services to plan, review, and approve LFT&E. Services will submit the plan to the DOT&E for comment at least 30 days prior to test initiation.

- 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.

- 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 120 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[4] will be produced by the DOT&E, approved by the Secretary of Defense, and transmitted to Congress. The conclusions of the report will be independent of the conclusions of the Service report. Reporting on LFT&E may be included in the weapon system's Beyond Low Rate Initial Production (BLRIP) Report completed by the DOT&E.

- 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 Deputy DOT&E-Live Fire may elect to supplement such funds for the acquisition of targets or target simulators, although the ultimate responsibility rests on the concerned Service.

The Survivability, Vulnerability Information Analysis Center (SURVIAC) is the DoD focal point for nonnuclear survivability/vulnerability data, information, methodologies, models, and analysis relating to U.S. and foreign aeronautical and surface systems. Data on file for conventional missiles and guns includes information on acquisition, detection, tracking, launch, fly-out and fuzing characteristics, countermeasures and counter-countermeasures, and terminal effects. Other weapons systems information includes physical and functional characteristics; design, performance, and operational information; acoustics, infrared, optical, electro-optical, and radar signature; combat damage and repair; and system, subsystem, and component probability of kill given a hit (pk/h) functions. SURVIAC is sponsored by the Joint Logistics Commanders' Joint Aircraft Survivability Program Office (http://jas.jcs.mil) and the Joint Technical Coordinating Group on Munitions Effectiveness (http://jtcg.amsaa.army.mil/index.html).

18.3 TESTING FOR NUCLEAR HARDNESS AND SURVIVABILITY

18.3.1 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. Program Managers (PMs) and their 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.*[5]

18.3.2 Assessing Nuclear Survivability Throughout the System Acquisition Cycle

The PM must ensure that nuclear survivability issues are addressed throughout the system acquisition cycle. During assessment of concepts, the survivability requirements stated in the Service requirements document should be verified, refined, or further defined. During the system's early design stages, tradeoffs between hardness levels and other system characteristics (such as weight, decontaminability, and compatibility) should be described quantitatively. Tradeoffs between hardness, avoidance, proliferation, and reconstitution as a method for achieving survivability should also be considered at this time. During advanced engineering development, the system must be adequately tested to confirm that hardness objectives, criteria, requirements, and specifications are met. Plans for NHS testing should be addressed in the TEMP. 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 FRP, deployment, and operational support, 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 reevaluated 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, logistics support, maintenance actions, or natural causes.

18.3.3 Test Planning

The *Nuclear Survivability Handbook for Air Force OT&E* describes the following challenges associated with NH&S testing:

(1) The magnitude and range of effects from a nuclear burst are much greater than those from conventional explosions that may be

Table 18-3. Nuclear Hardness and Survivability Assessment Activities

CONCEPT ASSESSMENT

- 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

PROTOTYPE TESTING

- INCREASE TEST PLANNING
- TEMP UPDATE
- CONDUCT OF NH&S TRADE STUDIES
  - NH&S REQUIREMENTS VERSUS 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 PERFORMANCE SPECIFICATIONS THAT INCLUDE QUANTITATIVE HARDNESS LEVELS

ENGINEERING DEVELOPMENT MODEL (EDM)

- 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 TEST
  - DEVELOPMENTAL 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 (LRIP, FULL RATE), DEPLOYMENT AND OPERATIONAL SUPPORT

- IMPLEMENTATION OF PROGRAM TO ENSURE SYSTEM RETAINS ITS NH&S PROPERTIES
- SCREENING OF PRODUCTION HARDWARE FOR HARDNESS
- IMPLEMENTATION BY USER OF PROCEDURES TO ENSURE SYSTEM'S OPERATION, LOGISTICS SUPPORT AND MAINTENANCE DO NOT DEGRADE HARDNESS FEATURES
- REASSESSMENT OF SURVIVABILITY THROUGHOUT SYSTEM LIFE CYCLE

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 (Operational Test and Evaluation) test program than normally found in other types of testing.

PMs and Test Managers (TMs) 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, PMs/TMs must exercise caution when developing test objectives and specifying measures of merit related to nuclear survivability.

18.3.4 Test Execution

For NH&S 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.

T&E 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 18-3 outlines the major activities relevant to the assessment of NH&S and the phases of the acquisition cycle in which they occur.

18.4 SUMMARY

The survivability and lethality aspects of certain systems must be evaluated through live fire tests. These tests are used to provide insights into the weapon system’s ability to continue to operate/fight after being hit by enemy threat systems. It provides a way of examining the damages inflicted not only on materiel but also on personnel. LFT 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. Tradeoffs 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 resulted in a variety of interpretations. Therefore, PMs must exercise caution when developing test objectives related to nuclear survivability.

ENDNOTES

1. Air Force Operational Test and Evaluation Center (AFOTEC), *Nuclear Survivability Handbook for Air Force OT&E*.

2. DoD Instruction (DoDI) 5030.55, *Joint AEC-DoD Nuclear Weapons Development Procedures,* January 4, 1974.

3. Office of the Deputy Director, Test and Evaluation (Live Fire Test), *Live Fire Test and Evaluation Planning Guide*, June 1989.

4. *Defense Acquisition Guidebook,* Appendix 3. http://akss.dau.mil/DAG/.

5. AFOTEC.

19

LOGISTICS INFRASTRUCTURE T&E

19.1 INTRODUCTION

In all materiel acquisition programs, the logistics support effort begins in the capabilities assessments of mission needs, leads to the development of the Initial Capabilities Document (ICD) before program initiation, continues throughout the acquisition cycle, and extends past the deployment phase. Logistics support system 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. Discussion of Performance Based Logistics (PBL) can be found at the Defense Acquisition University (DAU) Web site (http://www.dau.mil), Acquisition, Technology, and Logistics (AT&L) Knowledge Sharing Site (AKSS) for the Logistics Community of Practice (CoP). 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 logistics manager's participation in the Test and Evaluation Master Plan (TEMP) development process should be stressed. The logistics manager must ensure the logistics system Test and Evaluation (T&E) objectives are considered and that adequate resources are available for logistics support system T&E.

Logistics system support planning is 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 deployment and operational support at affordable cost.

Logistics support systems are usually categorized into 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.

19.2 PLANNING FOR LOGISTICS SUPPORT SYSTEM T&E

19.2.1 Objectives of Logistics System T&E

The main objective of logistics system T&E is to verify that the logistics support being developed for the materiel system is capable of meeting the required objectives for peacetime and wartime employment. This T&E consists of the usual Development Test and Evaluation (DT&E) and Operational Test and Evaluation (OT&E) but also

ACQUISITION PHASE / TEST TYPE	CONCEPT REFINEMENT	TECHNOLOGY DEVELOPMENT	SYSTEM DEVELOPMENT AND DEMONSTRATION	PRODUCTION & DEPLOYMENT	OPERATIONS & SUPPORT
DEVELOPMENTAL T&E	• ASSESS PREFERRED SYSTEM SUPPORT CONCEPTS • INPUT TO AOA • INPUT TO TDS • SUPPORTABILITY M&S	• ASSESS TECHNICAL APPROACH, LOGISTICS RISKS, AND ALTERNATIVE SUPPORTABILITY SOLUTIONS • EVALUATE LOGISTICS-RELATED TECHNOLOGIES • TDS LOGISTICS ASSESSMENT STRATEGY EXPANDED FOR TEMP • ASSESS TESTABILITY OF POTENTIAL LOGISTICS FOOTPRINT ELEMENTS • ASSESS THRESHOLDS/ OBJECTIVES FOR LOGISTICS PARAMETERS IN CDD	• IDENTIFY DESIGN PROBLEMS AND ASSESS SOLUTIONS FOR SYSTEM LOGISTICS ELEMENTS • EVALUATE SYSTEM CDD LOGISTICS PARAMETERS, AND RAM PREDICTIONS • CONDUCT EARLY MAINTENANCE DEMOS, LEVEL OF REPAIR ASSESSMENTS, AND MAINTENANCE TASK ANALYSES • CONDUCT FAILURE MODE EFFECTS, AND CRITICALITY ANALYSIS (FMECA) • SUPPORT SYSTEMS ENGINEERING DESIGN REVIEWS • RELIABILITY GROWTH T&E	• VERIFY CPD LOGISTICS PARAMETERS FOR PRODUCTION SYSTEMS • PROVIDE INSIGHT INTO READINESS FOR OPERATIONAL SUITABILITY TESTING OF LOGISTICS ELEMENTS • CONDUCT ASSESSMENTS OF MAINTENANCE CONCEPTS AND PROCESSES • PROVIDE RAM DATA TO SUPPORT LCCE	• EVALUATE IMPACT OF ENGINEERNIG CHANGES ON LOGISTICS PARAMETERS • ASSESS ADEQUACY OF FIXES FOR LOGISTICS DEFICIENCIES IDENTIFIED BY FIELD UNITS • EVALUATE FIXES FOR IOT&E DEFICIENCIES • DEMONSTRATE ATTAINMENT OF RAM/READINESS CRITERIA
OPERATIONAL T&E	• ASSESS POTENTIAL FOR OPERATIONAL SUITABILITY OF PREFERRED CONCEPT • ASSIST IN SELECTING PREFERRED SYSTEM SUPPORT CONCEPTS • INPUT TO TDS	• EXAMINE OPERATIONAL ASPECTS OF SUPPORTABILITY TECHNICAL SOLUTIONS • DEVELOP MEASUES AND COI FOR FUTURE SUITABILITY ASSESSMENTS-INPUT FOR TEMP • ESTIMATE POTENTIAL OPERATIONAL SUITABILITY OF LOGISTICS SYSTEMS AND SUSTAINMENT PLANNING	• ASSESS ABILITY OF LOGISTICS SYSTEM TO PROVIDE SUPPORT IN A RELEVANT OPERATIONAL ENVIRONMENT • ASSESS IMPROVEMENT IN OPERATIONAL SUITABILITY FROM PROTOTYPE TO EDM SYSTEMS • REFINE CONCEPTS FOR EVALUATION OF CPD LOGISTICS PARAMETERS	• ASSESS OPERATIONAL SUITABILITY FOR PRODUCTION SYSTEMS • PROVIDE RELIABILITY DATA FOR RELIABILITY GROWTH MODEL	• EVALUATE OPERATIONAL SUITABILITY AND SUPPORTABILITY OF DESIGN CHANGES • IDENTIFY IMPROVEMENTS FOR SUPPORTABILITY PARAMETERS • PROVIDE DATA FOR ADJUSTING SUPPORTABILITY MODELS • EVAUATE LOGISTICS PARAMETERS WAIVED FOR IOT&E

Figure 19-1. Logistics Supportability Objectives in the T&E Program

includes post-deployment supportability assessments. The formal DT&E and OT&E begin with subcomponent assembly and prototype testing, and continue throughout advanced engineering development, production, deployment, and operational support. Figure 19-1 describes the specific Development Test (DT), Operational Test (OT), and supportability assessment objectives for each acquisition phase. "The PM shall apply Human Systems Integration (HSI) to optimize total system performance (hardware, software, and human), operational effectiveness, and suitability, survivability, safety, and affordability."[1]

19.2.2 Planning for Logistics Support System T&E

19.2.2.1 Logistics Support Planning

"The PM shall have a comprehensive plan for HSI in place early in the acquisition process to optimize total system performance, minimize Total Ownership Costs [TOCs], and ensure that the system is built to accommodate the characteristics of the user population that will operate, maintain, and support the system." HSI includes design for: human factors engineering; personnel; habitability; manpower; training; Environment, Safety, and Occupational Health (ESOH); and survivability.[2] The logistics support manager for a materiel acquisition system is responsible for developing the logistics support planning, which documents planning for and implementing the support of the fielded system. It is initially prepared during preparation for program initiation, and progressively developed in more detail as the system moves through the acquisition phases. Identification of the specific logistics support test issues related to the individual logistics elements and the overall system support and readiness objectives are included.

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

19.2.2.2 Supportability Assessment Planning

Based upon suitability objectives, the logistics support manager, in conjunction with the system's Test Manager (TM), develops suitability assessment planning. This planning identifies the testing approach and the evaluation criteria that will be used to assess the supportability-related design requirements (e.g., Reliability and Maintainability (R&M)) and adequacy of the planned logistics support resources for the materiel system. Development of the suitability T&E planning begins during concept refinement; the planning is then updated and refined in each successive acquisition phase. The logistics support manager may apply the best practices of logistics support analysis as described in Military Standard (MIL-STD)-1388-1A.[3]

The T&E strategy is formulated, T&E program objectives and criteria are established, and required test resources are identified. The logistics support 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 logistics support 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.

19.2.2.3 Test and Evaluation Master Plan

The Program Manager (PM) must include suitability T&E information in the TEMP as specified in the *Defense Acquisition Guidebook*. The input, derived from logistics supportability planning with the assistance of the logistics support 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 post deployment supportability) be identified in the TEMP because the TEMP provides a long-range alert upon which test resources are budgeted and obtained for testing.

19.2.3 Planning Guidelines for Logistics Support System T&E

The following guidelines were selected from those listed in the *Acquisition Logistics Guide*:[4]

(1) Develop a test strategy for each logistics support-related objective. Ensure that OT&E planning encompasses all logistics support elements. The general objectives shown in Figure 19-1 must be translated into detailed quantitative and qualitative requirements for each acquisition phase and each T&E program.

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

(3) Identify logistics support 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 logistics support items for formal DT/OT and separate logistics support system 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 logistics T&E designated 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.

Additional guidance may be found in the *Logistics Test and Evaluation Handbook* developed by the 412 Logistics Group, Edwards Air Force Base, California.

19.3 CONDUCTING LOGISTICS SUPPORT SYSTEM T&E

19.3.1 The Process

The purposes of logistics support system 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 become available, and to assess the operational suitability of the planned support system. It also evaluates the system's ability to achieve planned readiness objectives for the system/equipment being developed. Specific logistics support system T&E tasks (guidance in MIL-STD-1388-1A) include:[5]

- 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, logistics support resources, or operational tactics;

- Projection of changes in costs, readiness, and logistics 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.

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

19.3.2 Reliability and Maintainability

System availability is generally considered to be composed of two major system characteristics—Reliability and Maintainability (R&M). Reliability, Availability, and Maintainability (RAM) requirements should be based on operational requirements and Life Cycle Cost (LCC) considerations; stated in quantifiable, operational terms; measurable during developmental and operational test and evaluation; and defined for all elements of the system, including support and training equipment. Relevant definitions from the *DAU Glossary*:[6]

> Reliability (R) is the ability of a system and its parts to perform its mission without failure, degradation, or demand on the support system. A common MOE has been Mean Time Between Failure (MTBF).
>
> 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. A common Measure of Effectiveness (MOE) has been Mean Time To Repair (MTTR).
>
> 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. R&M objectives should 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 *T&E of System Reliability, Availability and Maintainability*.[7] 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.[8]

> Reliability, Availability, and Maintainability (also known as RAM), are requirements imposed on acquisition systems to insure they are operationally ready for use when needed, will successfully perform

assigned functions, and can be economically operated and maintained within the scope of logistics concepts and policies. RAM development programs are applicable to materiel systems, test measurement and diagnostic equipment, training devices, and facilities supporting weapon systems.

Availability is a measure of the degree to which an item is in an operable state and can be committed at the start of a mission when the mission is called for at an unknown (random) point in time.

Availability commonly uses different MOE based on the maturity of the weapon system technology. Achieved availability (Aa) is measured for the early prototype and EDM systems developed by the system contractor when the system is not operating in its normal environment. Inherent Availability (Ai) is measured with respect only to operating time and corrective maintenance, ignoring standby/delay time and mean logistics delay time. It may be measured during DT&E or when testing in the contractor's facility under controlled conditions. Operational Availability (Ao), measured for mature systems that have been deployed in realistic operational environments, is the degree to which a piece of equipment works properly when it is required for a mission. It is calculated by dividing uptime by uptime plus all downtime. The calculation is sensitive to low utilization rates for equipment. Ao is the quantitative link between readiness objectives and supportability.[9]

19.3.2.1 Reliability

Guidelines for reliability evaluation are to be published in a non-government standard to replace MIL-STD-785:

- **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 simulate 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-Test (TAFT) where equipment is tested under actual, simulated, or accelerated environments. It is an iterative methodology intended to rapidly and steadily improve reliability. Test articles are usually subjected to ESS prior to beginning RDT/RGT to eliminate those with manufacturing defects.

- **Reliability Qualification Test (RQT)** is to verify that threshold 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 percent sampling is not recommended.

19.3.2.2 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:[10]

- **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, a maintainability demonstration should be conducted prior to Full Rate Production (FRP).[11]

19.3.3 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 onsite 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 logistics support 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.

19.3.4 Data Collection and Analysis

The logistics support manager must coordinate with the testers to ensure that the methods used for collection, storage, and extraction of logistics support system T&E data are compatible with those used in testing the materiel system. As with any testing, the test planning must ensure that all required data are 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.

19.3.5 Use of Logistics Support System Test Results

The emphasis on the use of the results of testing changes as the program moves from the concept assessments to post-deployment. During early phases of a program, the evaluation results are used primarily to verify analysis and develop future projections. As the program moves into advanced engineering development and hardware becomes available, the evaluation addresses design, particularly the R&M aspects; training programs; support equipment adequacy; personnel skills and availability; and technical publications.

The logistics support system manager must make the 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.

19.4 LIMITATIONS TO LOGISTICS SUPPORT SYSTEM 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. An acquisition strategy may stipulate that support resources such as operator and maintenance manuals, tools, support equipment, training devices, etc., for major weapon system components would not be procured before the weapons system/component hardware and software design stabilizes. This shortage of equipment is often the reason that shelf-life and service-life testing is incomplete, leaving the logistics support system 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 logistics support system testing will continue post-production to obtain required sample sizes for verifying performance criteria. Many aspects of the logistics support system may not be available for Initial Operational Test and Evaluation (IOT&E) and become testing limitations. The PMO must develop enough logistics support to ensure the user can maintain the system during IOT&E without requiring system contractor involvement (legislated constraints). Any logistics support system limitations upon IOT&E will likely be evaluated during Follow-on Operational Test and Evaluation (FOT&E).

19.5 SUMMARY

T&E 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 logistics support system T&E is based upon the completeness and timeliness of the planning effort.

The logistics support system 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 the Full Rate Production Decision Review (FRPDR), decision makers can expect that some logistics system performance parameters will not have finished testing because of the large sample sizes required for high confidence levels in statistical analysis.

ENDNOTES

1. DoD Directive (DoDD) 5000.1, *The Defense Acquisition System*, May 12, 2003, p. 11.

2. DoD Instruction (DoDI) 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003, p. 43.

3. MIL-STD-1388-1A, *Logistics Support Analysis.* Cancelled.

4. Defense Systems Management College, *Acquisition Logistics Guide,* December 1997.

5. MIL-STD-1388-1A.

6. Defense Acquisition University, *Glossary of Defense Acquisition Acronyms & Terms*, 11th edition, September 2003.

7. DoD 3235.1-H, *Department of Defense T&E of System Reliability, Availability and Maintainability,* March 1982.

8. MIL-STD-781, *Reliability Testing for Engineering Development, Qualification, and Production*. Cancelled.

9. DoD 3235.1-H

10. *Ibid.*

11. Guidelines in MIL-HDBK-470, *Designing and Developing Maintainable Products and Systems,* 1998.

20

EC/C^{4}ISR TEST AND EVALUATION

20.1 INTRODUCTION

Testing of Electronic Combat (EC) and Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance (C^{4}ISR) systems poses unique problems for testers because of the difficulty in measuring their operational performance. Compatibility, Interoperability, and Integration (CII) are key performance areas for these systems. Special testing techniques and facilities are normally required in EC and C^{4}ISR testing. This chapter discusses the problems associated with EC and C^{4}ISR testing and presents methodologies the tester can consider using to overcome the problems.

20.2 TESTING EC SYSTEMS

20.2.1 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 Air Force Base (AFB), Florida, or the desert at Nellis AFB, Nevada, 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 the tester 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.

20.2.2 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 20-1. 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.

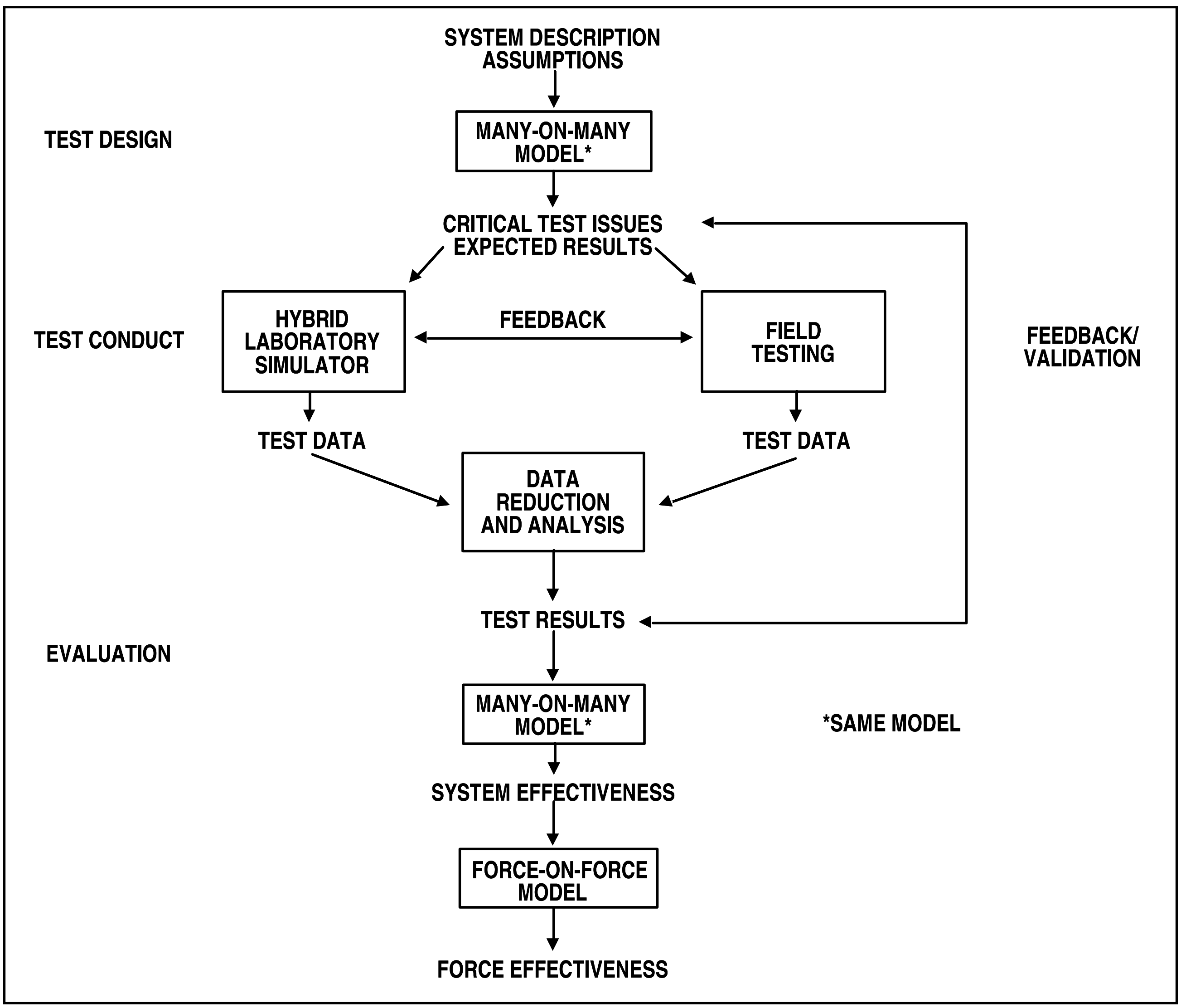


Figure 20-1. Integrated EC Testing Approach

Another alternative is the electronic combat test process.[1] The six-step process described here is graphically represented by Figure 20-2:

(1) Deriving test requirements,

(2) Conducting pre-test 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 20-3, 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 tradeoffs to obtain the necessary test data. Many more iterations of a computer simulation can be run for the cost of an open-air test.

20.3 TESTING OF C[4]ISR SYSTEMS

20.3.1 Special Considerations When Testing C[4]ISR Systems

The purpose of a C[4]ISR system is to provide a commander with timely and relevant information

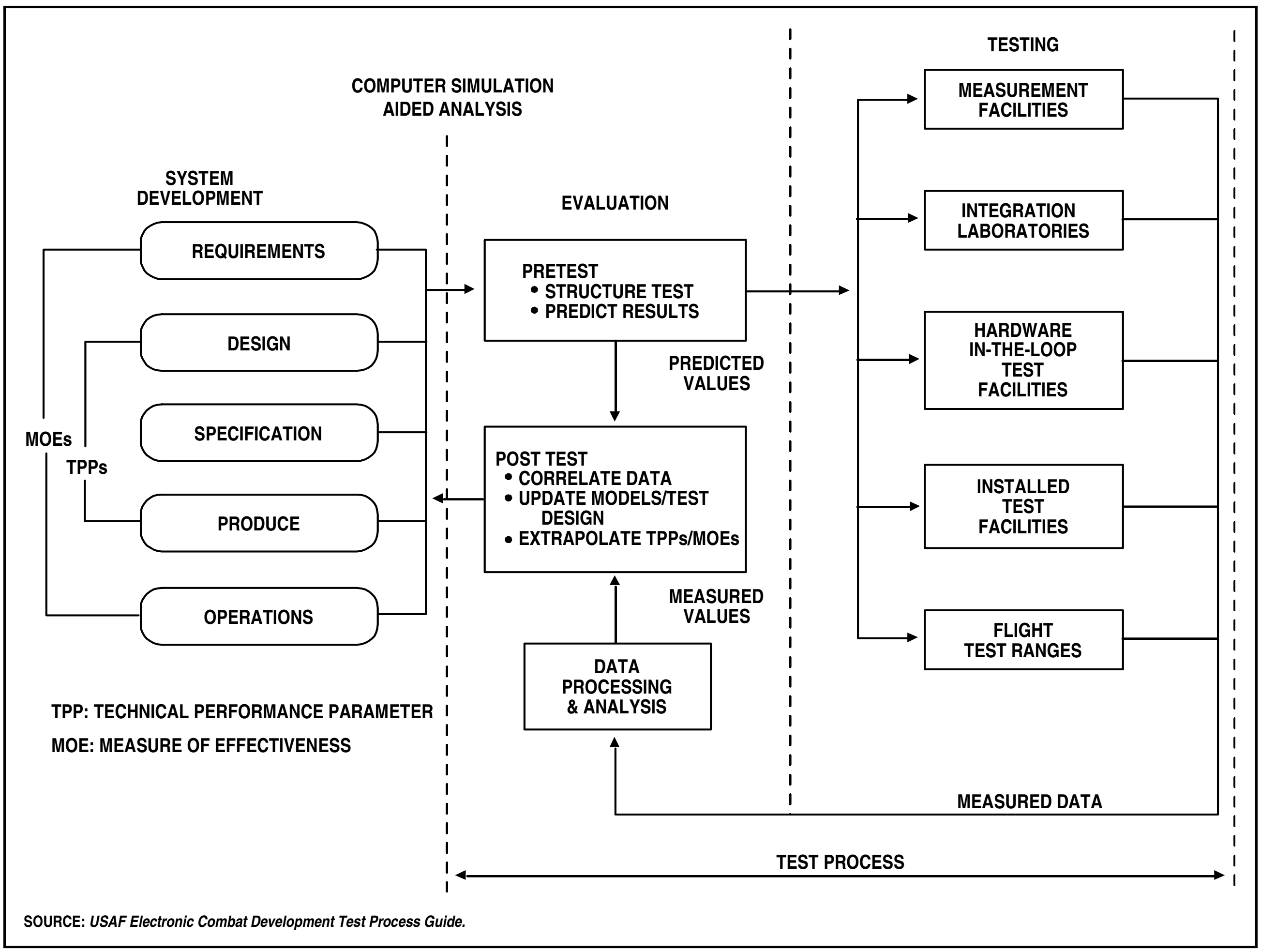


Figure 20-2. EC Test Process Concept

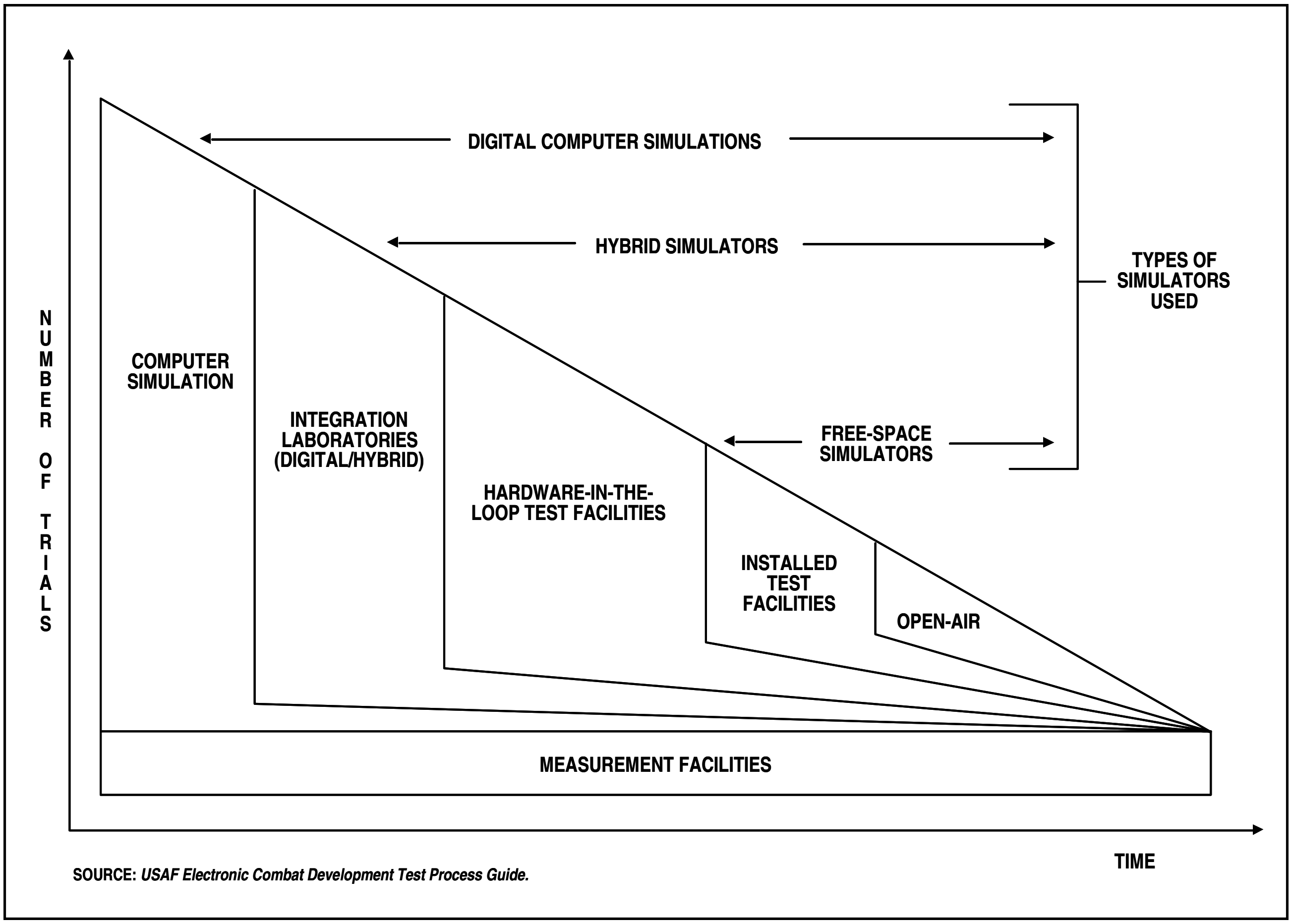


Figure 20-3. EC Test Resource Categories

to support sound warfighting decisionmaking. A variety of problems faces the C[4]ISR system tester. However, in evaluating command effectiveness, it is difficult to separate the contribution made by the C[4]ISR system from the contribution made by the commander's innate cognitive processes. To assess a C[4]ISR system in its operational environment, it must be connected to the other systems with which it would normally operate, making traceability of test results difficult. Additionally, modern C[4]ISR systems are software-intensive and highly interactive, with complex man-machine interfaces. Measuring C[4]ISR 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 C[4]ISR systems of defense agencies and the Services (Army, Navy, Air Force, and Marines Corps) are expected to interoperate with each other and with those of the North Atlantic Treaty Organization (NATO) forces; hence, the tester must also ensure inter-Service and NATO CII.

20.3.2 C[4]I Test Facilities

Testing of Command, Control, Communications, Computers and Intelligence (C[4]I) systems will have to rely more on the use of computer simulations and C[4]I test-beds to assess their overall effectiveness. The Defense Information Systems Agency (DISA), which is responsible for ensuring interoperability among all U.S. tactical C[4]I systems that would be used in joint or combined operations, directs the Joint Interoperability Test Command (JITC) at Ft. Huachuca, Arizona. The JITC is a test-bed for C[4]I systems CII.

20.4 TRENDS IN TESTING C^4I SYSTEMS

20.4.1 Evolutionary Acquisition of C^4I 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 C^4I 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[2] and the International Test and Evaluation Association (ITEA) have addressed the issues involved in the EA and testing of Command, Control, and Communications (C^3) systems. The ITEA study illustrated EA in Figure 20-4 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.[3]

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 PM must insist that the testing for EA systems be carefully planned to ensure the test data are shared by all and there is a minimum of repetition or duplication in testing.

20.4.2 Radio Vulnerability

The Radio Vulnerability Analysis (RVAN) methodology is for assessing the anti-jam capability and limitations of Radio Frequency (RF) data links when operating in a hostile Electronic Countermeasures (ECM) 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 anti-jam capabilities and vulnerabilities of an RF data link operating in a hostile 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.

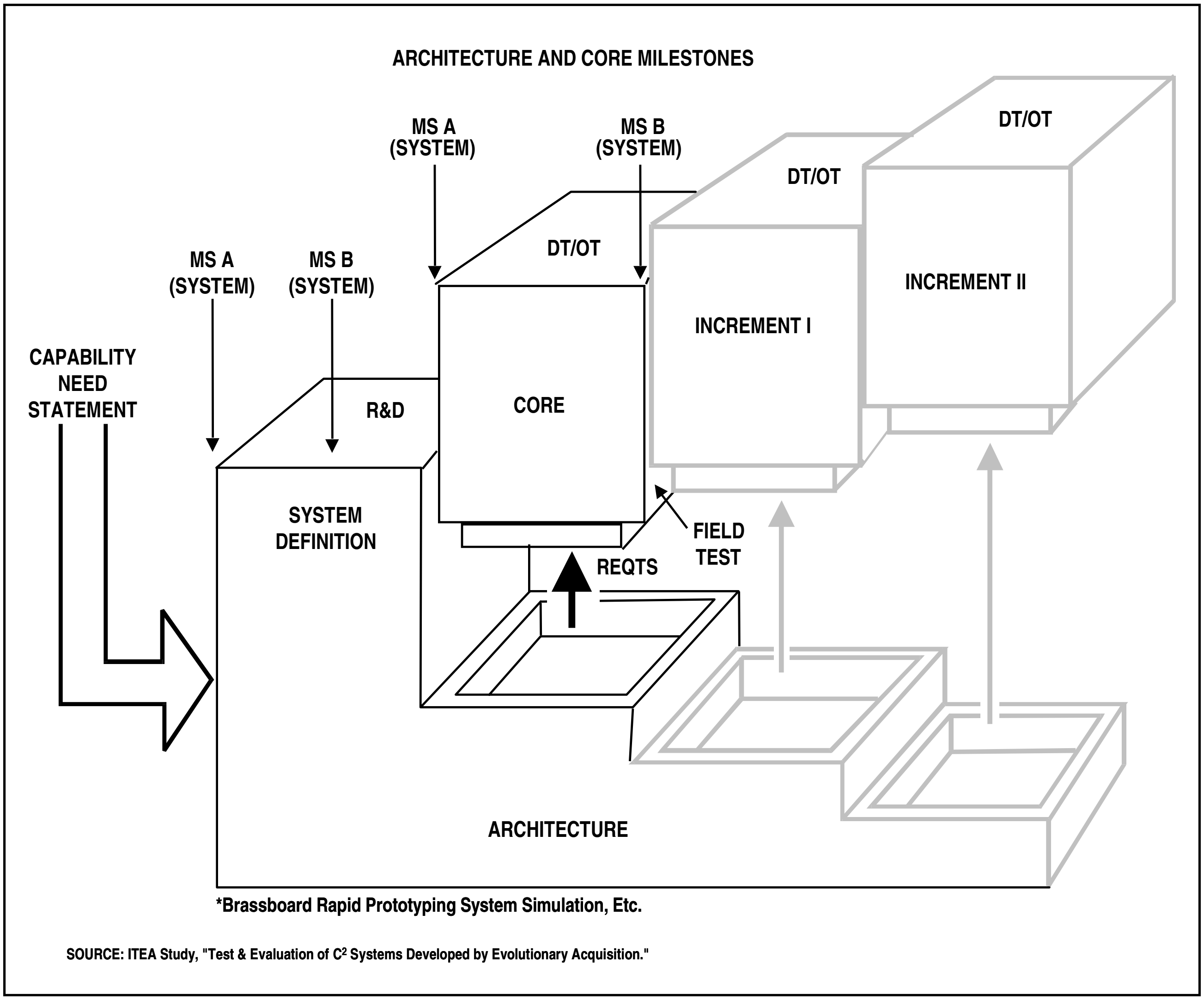


Figure 20-4. The Evolutionary Acquisition Process

The analyst applying the DVAL methodology will require test data; and the Test Manager (TM) of the C^3I 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 Air Force Operational Test and Evaluation Center (AFOTEC), Kirtland AFB, New Mexico.

20.5 T&E OF SURVEILLANCE AND RECONNAISSANCE SYSTEMS

Intelligence, Surveillance, and Reconnaissance (ISR) capabilities provide the requisite battlespace awareness tools for U.S. Forces to take and hold the initiative, increase operating tempo, and concentrate power at the time and place of their choosing. These vital capabilities are achieved through highly classified sensor systems ranging from satellites, aircraft, maritime vessels, electronic intercept, and the soldier in the field to the systems required to analyze that data, synthesize it into usable information, and disseminate that information in a timely fashion to the warfighter. As a general rule, ISR systems are considered to be Joint Systems.

Because of the multifaceted nature of ISR programs, the classified nature of how data are gathered in the operational element, test planning to verify effectiveness and suitability can be complex. Adding to that inherent complexity is the variable nature of organizational guidance directives upon the tester. While the broad management principles enunciated by Department of Defense Directive (DoDD) 5000.1, *The Defense Acquisition System*, May 12, 2003, will apply to highly sensitive classified systems and cryptographic and intelligence programs, the detailed guidance contained in *Defense Acquisition Guidebook* applies to Major Defense Acquisition Programs (MDAPs) and Major Automated Information Systems (MAISs). Many ISR programs fall below this threshold and the wise TM should anticipate that several agencies will have taken advantage of this opening to tailor organizational guidance.

Key issues for the T&E of ISR systems to consider include compliance verification with the CII requirements contained in DoDD 4630.5, *Interoperability and Supportability of Information Technology (IT) and National Security Systems (NSS)*, Chairman of the Joint Chiefs of Staff Instruction (CJCSI) 3170.01D, *Joint Capabilities Integration and Development System*, and CJCSI 6212.01C, *Interoperability and Supportability of Information Technology and National Security Systems*,[4] as certified by the JITC; completion of the system security certification required prior to processing classified or sensitive material; verification; and documentation of required support from interfaced C^4ISR systems in the Information Support Plan (ISP) to ensure the availability and quality of required input data, characterization of the maturity of mission critical software, finalization of the range of human factors analysis, and consideration of information operations vulnerabilities/capabilities. In addition to this partial listing, many of these systems will operate inside a matrix of ISR system architectures, which must be carefully considered for test planning purposes. As a final issue, Advanced Concept Technology Demonstration (ACTD) programs are being used to quickly deliver capability to the user because of the critical and time-sensitive nature of many ISR requirements. The TM must carefully consider structuring the T&E effort in light of the inherent nature of ACTD programs.

20.6 SUMMARY

The EC systems must be tested under conditions representative of the dense threat signal environments in which they will operate. The C^4ISR 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 C^4ISR T&E.

ENDNOTES

1. 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.

2. Defense Systems Management College, *Joint Logistics Commanders Guidance for the Use of an Evolutionary Acquisition (EA) Strategy in Acquiring Command and Control (C^2) Systems*, March 1987.

3. International Test and Evaluation Association (ITEA), *Test and Evaluation of Command and Control Systems Developed Evolutionary Acquisition*, May 1987.

4. DoDD 4630.5, *Interoperability and Supportability of Information Technology (IT) and National Security Systems (NSS)*, May 5, 2004; CJCSI 3170.01D, *Joint Capabilities Integration and Development System*, May 12, 2004; and CJCSI 6212.01C, *Interoperability and Supportability of Information Technology and National Security Systems*, November 20, 2003.

21

MULTI-SERVICE TESTS

21.1 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 (Joint program). A multi-Service test program should not be confused with the Office of the Secretary of Defense (OSD)-sponsored, non-acquisition-oriented Joint Test and Evaluation (JT&E) program. A brief description of the JT&E program is provided in Chapter 6. The *Defense Acquisition Guidebook* provides guidance on management of joint acquisition programs.[1]

21.2 BACKGROUND

The *Defense Acquisition Guidebook* states: The Component Acquisition Executive (CAE) of a designated acquisition agent given acquisition responsibilities shall utilize the acquisition and test organizations and facilities of the Military Departments to the maximum extent practicable; A designated joint program shall have…one integrated test program and one set of documentation (TEMP [Test and Evaluation Master Plan]) and reports; The MDA [Milestone Decision Authority] shall designate a lead OTA [Operational Test Agency] to coordinate all OT&E [Operational Test and Evaluation]. The lead OTA shall produce a single operational effectiveness and suitability report for the program; unless statute, the MDA, or a MOA [Memorandum of Agreement] signed by all DoD [Department of Defense] components directs otherwise, the lead executive component shall budget for and manage the common RDT&E [Research, Development, Test and Evaluation] funds for assigned joint programs; and, individual components shall budget for their unique requirements.

Formulation of a multi-Service Test and Evaluation (T&E) program 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 joint acquisition program. A "Supporting Service" is a Service designated as a participant in the system acquisition.)

A multi-Service T&E program may include either Development Test and Evaluation (DT&E) or Operational Test and Evaluation (OT&E) or both. The Service's OTAs have executed a formal MOA on multi-Service OT&E that provides a framework for the conduct of a multi-Service Operational Test (OT) program.[2] The Defense Information Systems Agency (DISA) has signed the MOA in relation to the Joint Interoperability Test and certification for multi-Service T&E. The MOA includes a glossary of common and Service-peculiar terms used in multi-Service T&E.[3] Generally the process includes:

(1) In a multi-Service 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 may be attached to this single report.

(3) Deviations from the Lead Service T&E regulations may be accommodated by mutual agreement among the Services involved.

21.3 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 multi-Service T&E Integrated Product Team (IPT) should be established for each multi-Service test program. Its membership consists of one senior representative from each participating Service or agency headquarters. The T&E IPT 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.

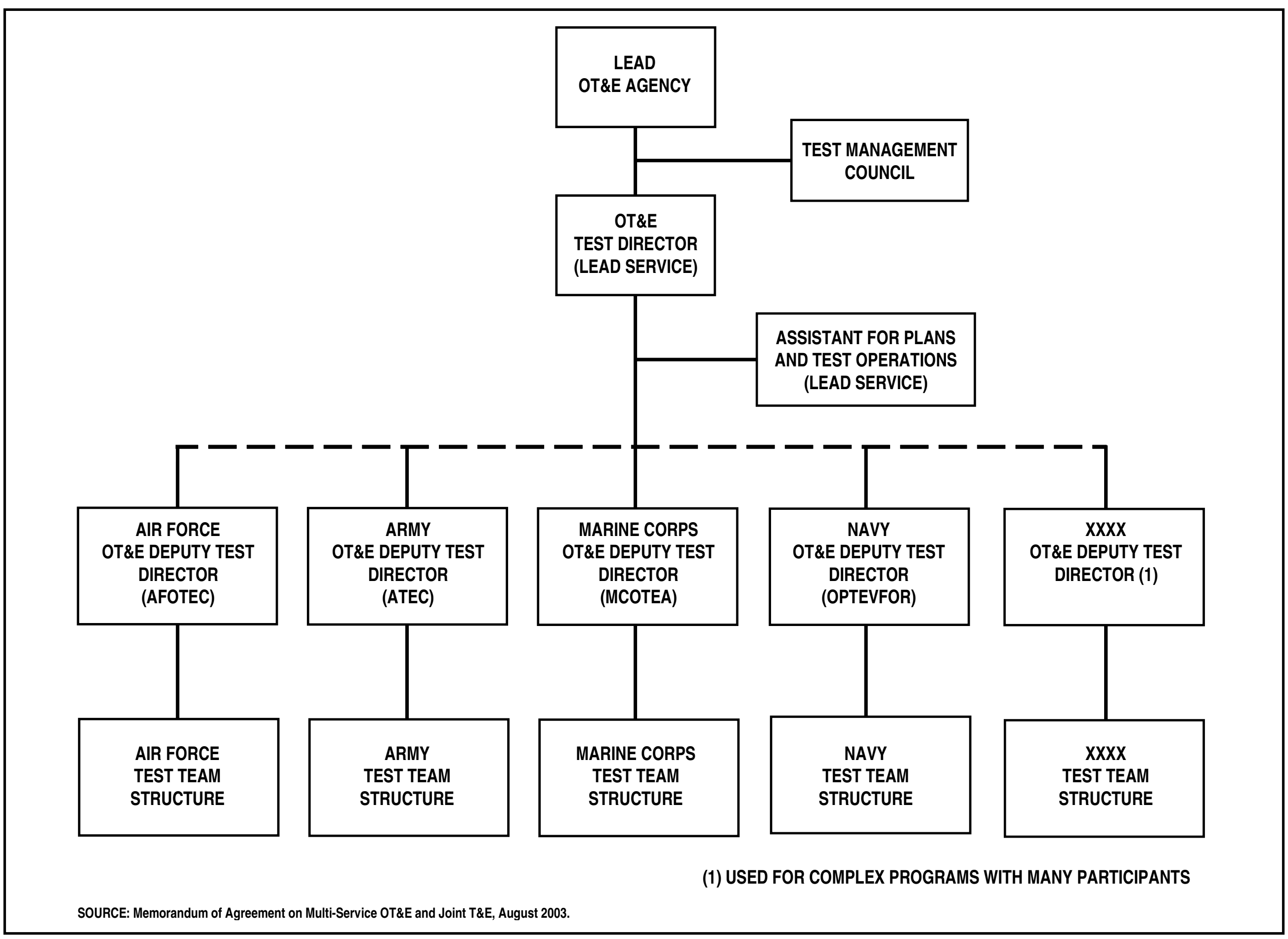


Figure 21-1. Simple Multi-Service OT&E Test Team Composition

Resource requirements are documented in the 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. Funding will be in accordance with Public Law and Department of Defense (DoD) 7000.14-R.[4]

21.4 TEST TEAM STRUCTURE

A sample test team structure is shown in Figure 21-1. 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.

21.5 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 [MOT&E] and Joint T&E:"[5]

(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.

(6) The Lead Service T&E agency prepares the multi-Service T&E test plan(s), consolidating the input from all participating agencies.

21.6 DEFICIENCY REPORTING

In a multi-Service T&E program, a deficiency 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 deficiency reporting system of the Lead Service is normally used. All members of the multi-Service test team will report deficiencies 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 deficiency considered disqualifying by one Service is not necessarily disqualifying for all Services. Deficiency reports of a disqualifying nature must include a statement by the concerned Service of why the deficiency has been so classified. It also includes statements by the other Services as to whether or not the deficiency affects them significantly.

If one of the participating Services identifies a deficiency that it considers as warranting termination of the test, the circumstances are reported immediately to the test director.

21.7 TEST REPORTING

The following test-reporting policy applies to multi-Service IOT&E:

(1) The Lead OTA will prepare and coordinate the report reflecting the system's operational effectiveness and suitability for each Service. Interim reports will not normally be prepared.

(2) It will synergize the different operational requirements and operational environments of the involved Services.

(3) It will state findings, put those findings into perspective, and present rationale why there is or is not consensus on the utility of the system.

(4) All participating OTAs will sign the report.

(5) Each participating OTA may prepare an Independent Evaluation Report (IER) or final test report, as required, in its own format and process that report through normal Service channels.

(6) The Lead OTA will ensure that all separate participating Service IERs/test reports are appended to the overall final report prepared by the Lead OTA for submission to the MDA.

(7) A single integrated multi-Service report will be submitted 90 calendar days after the official end of test is declared, but not later than 45 days prior to a milestone decision.

21.8 SUMMARY

Multi-Service test programs are conducted by two or more Services when a system is to be acquired for employment 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. The MOA on OT&E procedures provides guidance to the OTA. 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.

ENDNOTES

1. *Defense Acquisition Guidebook,* Chapter 7, "Acquiring Information Technology and National Security Systems," http//akss.dau.mil/DAG.

2. Memorandum of Agreement of Multi-Service OT&E, with changes, August 2003.

3. AFOTEC 99-103, *Capabilities Test and Evaluation*, October 22, 2001, and the Department of the Army Pamphlet (DA PAM) 73-1, *Test and Evaluation Policy*, January 7, 2002, provide guidance for procedures followed in a multi-Service T&E program.

4. DoD Financial Management Regulation (FMR) 7000.14-R, Vol. 02B, *Budget Formulation and Presentation*, Chapter 5, "Research, Development, Test, and Evaluation Appropriations," June 2004.

5. Memorandum.

22

INTERNATIONAL TEST AND EVALUATION PROGRAMS

22.1 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 (FCT) Program[1] and the International Test Operations Procedures (ITOPs). Factors that bear on the T&E of multinational acquisition programs are discussed also.

22.2 FOREIGN COMPARATIVE TEST PROGRAM

22.2.1 Program Objective

The 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 Forces. 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 support the U.S. national policy of encouraging international armaments cooperation and to reduce the costs of Research and Development (R&D). Policy and procedures for the execution of the program were originally documented in the Department of Defense (DoD), but now can be found in the *Foreign Comparative Test Handbook* and DoD Instruction (DoDI) 5000.2.[2]

22.2.2 Program Administration

Foreign Weapons Evaluation (FWE) activities and responsibilities are assigned to the Director, Comparative Test Office (CTO), Director of Defense Research and Engineering (DDR&E), OSD. Each year, sponsoring military Services forward Candidate Nomination Proposals (CNPs) for systems to be evaluated under the FCT program to the Director, CTO. 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 the *FCT Handbook.*[3]

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 (MOA), private industry licensing agreements, data exchange agreements, and/or cooperative technology exchange programs.

FWE 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. The Under Secretary of Defense for Acquisition, Technology, and Logistics (USD(AT&L)) shall notify Congress a minimum of 30 days prior to the commitment of funds for initiation of new FCT evaluations.[4]

22.3 NATO COMPARATIVE TEST PROGRAM

The North Atlantic Treaty Organization (NATO) Comparative Test Program has been integrated with the FCT program. It was created by an act of the Congress in the Fiscal Year (FY) 86 Defense Authorization Act. The program supported the evaluation of NATO nations' weapons systems, equipment and technology, and assessed their suitability for use by U.S. forces. The selection criteria for the NATO Comparative Test Program were essentially the same as for the FCT program. The exception was that the equipment must be produced by a NATO member nation and be considered as an alternative to a system that was either in a late stage of development in the United States or that offered a cost, schedule, or performance advantage over U.S. equipment. In addition, the NATO Comparative Test Program required that notification be sent to the Armed Services and Appropriations Committees of the House of Representatives and Senate before funds were obligated. With this exception, the NATO Comparative Test Program followed the same nomination process and administrative procedures. Guidelines for the program were also contained in the *FCT Handbook*.

Examples of proposals funded under the NATO Comparative Test Program included T&E of a German mine reconnaissance and detection system for the Army, a United Kingdom-designed mine sweeper for the Navy, and the Norwegian Penguin missile system for the Air Force. According to the FY88 Report of the Secretary of Defense (SECDEF) to the Congress, the program generated considerable interest among NATO allied nations and became 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.

22.4 T&E MANAGEMENT IN MULTINATIONAL PROGRAMS

22.4.1 Compatibility With Allies

Rationalization, Standardization, and Interoperability (RSI) have become increasingly important elements in the materiel acquisition process. Public Law 94-361[5] 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." Program Managers (PMs) and Test Managers (TMs) must, therefore, be fully aware of any potential international applications of the systems for which they are responsible. Two guidebooks published by the Defense Acquisition University (DAU) are valuable compendiums of information for the PM of a developing system with potential multinational applications.[6] Additionally, on file at the DAU David D. Acker Library, is the research report "Operational Test and Evaluation in the IDEA [International Defense Education Association] Nations" (1999) comparing the Operational Test and Evaluation (OT&E) processes of the United States with those of Great Britain, Germany and France. This report concluded that there were significant differences in what was considered OT&E by the four countries. The *Defense Acquisition Guidebook* states that Foreign Military Sales (FMS) of U.S. major systems (ACAT [Acquisition Category] I, II) prior to IOT&E must be approved by the USD(AT&L).[7]

Representatives of the United States, United Kingdom, France, and Germany have signed a MOA 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). The European concept of OT&E is significantly different from that used by the U.S. DoD.

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 (OT) scenario and battlefield representation that will be used.

22.4.2 International Test Operations Procedures

The *Defense Acquisition Guidebook*[8] states "results of T&E of systems using approved International Test Operations Procedures (ITOPs) may be accepted without repeating the testing." The ITOPs are documents containing standardized state-of-the-art test procedures prepared by the cooperative efforts of France, Germany, the United Kingdom, and the United States. Their use assures high quality, efficient, accurate, and cost-effective testing. The Director, Operational Test and Evaluation (DOT&E) is the OSD sponsor for providing basic guidance and direction to the ITOPs' processes. The ITOPs program is intended to shorten and reduce costs of the materiel development and acquisition cycle, minimize duplicate testing, improve interoperability of U.S. and allied equipment, promote the cooperative development and exchange of advanced test technology, and expand the customer base. Each Service has designated an ITOPs point of contact. The Army uses the Test and Evaluation Management Agency (TEMA); in the Navy it is the Director, Navy T&E Division (N-912); and the Air Force has the Chief, Policy and Program Division (AF/TEP). The Army, who initiated the program in 1979, is the lead Service. A total of 75 ITOPs have been completed and published in six technical areas under the Four-Nation Test and Evaluation MOU. Additional ITOPs are under development by active working committees.[9] Completed documents are submitted to the Defense Technical Information Center (DTIC) for official distribution.

22.5 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 (SOW), data requirements, operational test planning, and test schedule formulation must be included. Also, the U.S. Army operational user, Forces Command (FORSCOM), must be involved in the OT program. Usually, a Multinational Memorandum of Understanding 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) Electronic Support Measures (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 OT&E. Testing will be conducted in the U.S. and European the Program Management Office (PMO), 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 Program Office, and the NATO program is managed by the NAPMA.

22.6 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 (R&D) costs, faster initial operational capability, improved interoperability with friendly nations, and lower procurement costs because of economies of scale. This is normally a preferred solution to user requirements before attempting to start a new development. Testing such systems presents specific challenges to accommodate the needs of all users. OT&E conducted by allied and friendly nations is not as rigorous as that required by U.S. public law and DoD acquisition regulations. 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.

ENDNOTES

1. Title 10, United States Code (U.S.C.), Section 2350A, *Cooperative research and development agreements: NATO organizations; allied and friendly foreign countries,* March 18, 2004.

2. DoD 5000.3-M-2, *Foreign Comparative Testing Program Procedures Manual*, January 1994; DoDI 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003, Enclosure 5.

3. Office of the Under Secretary of Defense (Acquisition, Technology and Logistics) Director, Strategic & Tactical Systems, *The Foreign Comparative Testing (FCT) Handbook,* March 2001, http://www.acq.osd.mil/cto.

4. DoDI 5000.2.

5. Public Law 94-361, "DoD Authorization Act, 1977," (Title 10, U.S.C. 2457), July 14, 1976.

6. Kausal, Tony, editor. *Comparison of the Defense Acquisition Systems of France, Great Britain, Germany, and the United States*, Defense Acquisition University, September 1999; Kausal, Tony, editor and Stefan Markowski. *Comparison of the Defense Acquisition Systems of Australia, Japan, South Korea, Singapore, and the United States*, Defense Acquisition University, July 2000.

7. *Defense Acquisition Guidebook*, http://akss.dau.mil/DAG.

8. *Ibid.,* Chapter 2, C2.9.2, "International Cooperation."

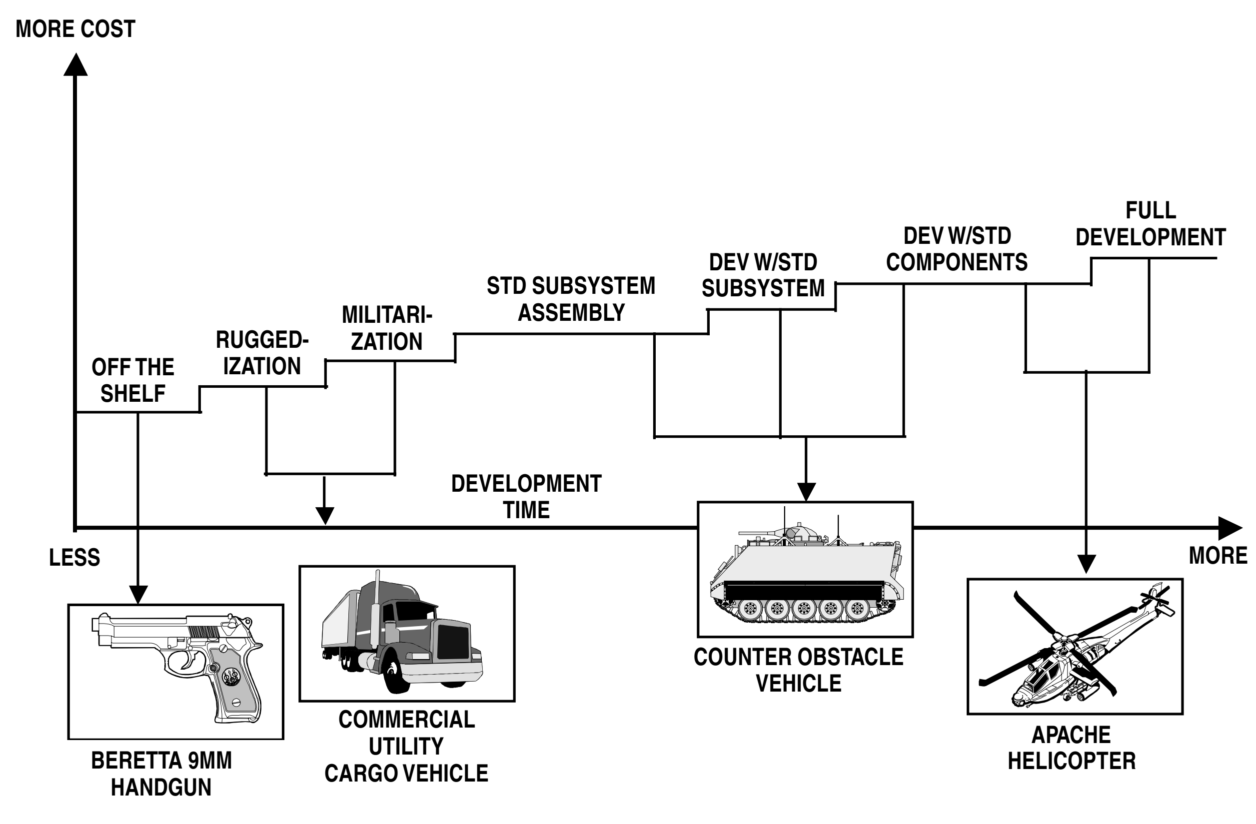
9. Check the Developmental Test Command (DTC) Web site at http://www.dtc.army.mil/publications/topsindex.aspx for updates.

23

COMMERCIAL AND NON-DEVELOPMENT ITEMS

23.1 INTRODUCTION

Many options are available when an acquisition strategy calls for a material solution before electing to develop a new system. They range from the preferred solution of using commercially available products from domestic or international sources to the least preferred option of a traditional, single-Service new Research and Development (R&D) program. Between these two extremes are other acquisition strategies that call for using modified systems at different component levels and unmodified or ruggedized cooperative developments to various extents.[1] Figure 23-1 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. The PM [Program Manager] is required to consider this spectrum of options for all levels of the system design.[2]



Source: Army Materiel Command Pamphlet 70-2, *AMC-TRADOC Materiel Acquisition Handbook*. March 26, 1987.

Figure 23-1. The Spectrum of Technology Maturity

23.1.1 Definitions

A commercial item is generally defined as any item, other than real property, that is of a type customarily used for non-governmental purposes and that: (1) has been sold, leased, or licensed to the general public; (2) has been offered for sale, lease, or license to the general public; or any item that evolved through advances in technology or performance and that is not yet available in the commercial marketplace, but will be available in the commercial marketplace in time to satisfy the delivery requirements under a government solicitation. Also included in the definition are services in support of a commercial item, or a type offered and sold competitively in substantial quantities in the commercial marketplace based on established catalog or market prices for specific tasks performed under standard commercial terms and conditions; this does not include services that are sold based on hourly rates without an established catalog or market price for a specific service performed.

A Non-Developmental Item (NDI) is considered: (1) any previously developed item of supply used exclusively for governmental purposes by a federal agency, a state, or local government, or a foreign government with which the United States has a mutual defense cooperation agreement; (2) any item described in (1) that requires only minor modification or modifications of the type customarily available in the commercial marketplace in order to meet the requirements of the procuring department or agency; (3) any item described in (1) or (2) solely because the item is not yet in use.

All such systems are required to undergo technical and Operational Test and Evaluation (OT&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.[3] See *Federal Acquisition Regulation (FAR)*, Section 2.101 for more precise definitions of commercial and NDIs; and *SD-2, Buying Commercial and Nondevelopmental Items: A Handbook,* April 1, 1996.

23.1.2 Advantages and Disadvantages of Commercial and NDI Approaches

The use of commercial and NDI offers the following advantages:

- The time to field a system can be greatly reduced, providing a quicker response to the user's needs;
- R&D costs or Total Ownership Costs (TOCs) may be reduced;
- State-of-the-art technology may be available sooner.

Commercial and NDIs offer the following disadvantages:

- Acquisitions are difficult to standardize/integrate with the current fleet equipment;
- Acquisitions create logistics support difficulties;
- Acquisitions tend not to have comparable competition; therefore, there are fewer second sources available;
- With commercial and NDI acquisitions, engineering and test data often are not available.

23.1.3 Application of Commercial and NDIs

Commercial items or NDIs may be used in the same environment for which the items were designed. Such items normally do not require development testing prior to the Production Qualification Test (PQT) 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.[4] An operational assessment or some more rigorous level of OT&E might be appropriate.

Commercial items or NDIs may be used 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.

Integration of commercial items or NDIs into a new development system may require some regression testing. These efforts require more extensive research, development, and testing to achieve effective operation of 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 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 commercial items and NDI systems and the environment in which those systems will be tested.

23.2 MARKET INVESTIGATION AND PROCUREMENT

A market investigation is the central activity leading to the program initiation decision regarding the use of a commercial item or 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 decision, to finalize the requirements, and to develop an acquisition strategy that is responsive to these requirements. Included in the search would be dual use technologies that meet military needs, yet have sufficient commercial application to support a viable production base. An assessment of Technology Readiness Level will provide the program office with insights into the readiness of the commercial item technologies for military application.[5]

During the market investigation, 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.[6]

It is imperative that technical and operational evaluators become involved during this early stage of any commercial item or NDI procurement and that they perform an early assessment of the initial issues. The evaluator must also relate these issues to Test and Evaluation (T&E) criteria and provide their Independent Evaluation Plans (IEPs) and reports to the decision authorities before the Milestone B decision review.

23.3 COMMERCIAL ITEM AND NDI TESTING

23.3.1 General Considerations

Test planning for commercial and NDIs shall recognize commercial testing and experience, but nonetheless determine the appropriate DT&E, OT&E, and LFT&E [Live Fire Test and Evaluation] needed to ensure effective performance in the intended operational environment.[7] The *Defense Acquisition Guidebook* suggests that "the PM shall develop an appropriate T&E strategy for commercial items to include evaluating commercial items in a system test bed, when practical; focusing test beds on high-risk items; and testing commercial item upgrades for unanticipated side effects in areas such as security, safety, reliability, and performance." T&E must be considered throughout the acquisition of a system that involves commercial items and NDI.

Program Managers (PMs) and their 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 commercial item or NDI and its anticipated use; it should be planned to support the design and decision process. At a minimum, T&E will be conducted to verify integration and interoperability with other system elements. All commercial item and 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. For example, a commercial item or 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. PMs and their staff must remember that these 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 a commercial item or NDI acquisition is assembled, PMs must ensure that it includes the following T&E-related items:

(1) Approved T&E issues and criteria;

(2) A requirement that the contractor provide a description of the testing previously performed on the system, including test procedures followed, data, and results achieved;

(3) PQT and quality conformance requirements;

(4) Acceptance test plans for the system and its components.

23.3.2 Testing Before Program Initiation

Since an important advantage of using a commercial item or NDI acquisition strategy is reduced acquisition time, 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, commercial items or NDI candidates may be bought or leased, and technical and operational tests may be conducted.

23.3.3 Testing After Program Initiation

All testing to be conducted after the program initiation milestone decision to proceed with the commercial item or NDI acquisition should be described in the Acquisition Strategy and the Test and Evaluation Master Plan (TEMP). 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 (IOT&E) agency should concur in any decisions to limit or eliminate operational testing.

T&E 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 (RAM) characteristics; logistics support plans; and training requirements, doctrine, and tactics.

23.4 RESOURCES AND FUNDING

Programming and budgeting for a commercial item or NDI acquisition present a special challenge. Because of the short duration of the acquisition process, the standard lead times required in the normal Planning, Programming, Budgeting and Execution (PPBE) 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 candidate systems/components required for T&E purposes. The RDT&E funds are also used to support T&E activities such as: modification of the test article; purchase of specifications, manufacturer's publications, repair parts, special tools and equipment; transportation of the test article 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 (P&D) costs.

One chief reason for using a commercial item or NDI acquisition strategy is reduced overall TOC. Additional cost savings can be achieved after a contract has been awarded if the PM ensures that incentives are provided to contractors to submit Value Engineering Change Proposals (VECPs) to the government when unnecessary costs are identified.

23.5 SUMMARY

The use of commercial items and NDIs in a system acquisition can provide considerable time and cost savings. The testing approach used must be carefully tailored to the type of system, levels of modifications, technology risk areas, 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.

ENDNOTES

1. Department of Defense Directive (DoDD) 5000.1, *The Defense Acquisition System*, May 12, 2003.

2. DoD Instruction (DoDI) 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003.

3. *Ibid*.

4. *Ibid*.

5. *Defense Acquisition Handbook*, Appendix 6, http://akss.dau.mil/DAG/.

6. Department of the Army Pamphlet (DA PAM) 73-1, *Test and Evaluation in Support of Systems Acquisition*, February 1997.

7. DoDI 5000.2.

24

TESTING THE SPECIAL CASES

24.1 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, Advanced Concept Technology Demonstrations (ACTDs), and systems with high-cost and/or special security considerations. Examples include: chemical and laser weapons; ships; space weapons; and missile systems.

24.2 TESTING WITH LIMITATIONS

Certain types of systems cannot be tested using relatively standard Test and Evaluation (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 [Initial Operational Test and Evaluation] and FOT&E [Follow-on Operational Test and Evaluation])," 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 (COIs) 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; and ships. 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.

The tester of dangerous or lethal systems, like chemical and laser weapons, must consider various safety, health, and medical factors in developing 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 (EISs).

Also, the tester must allow for additional planning time, test funds, and test resources to accommodate such factors.

24.2.1 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. 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 *Index of Test Operations Logistics Support: Developmental Supportability Test and Evaluation Guide*, a bibliography that includes numerous reports on chemical testing issues and procedures, can be consulted for specific documentation on chemical testing.[1]

24.2.2 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 specially configured geographic locations during T&E. For instance, the program seeking to free-play non-eye safe airborne or ground lasers during tests involves significant precautions, such as the airspace must be restricted from overflight during active testing and guards must be posted to prevent anyone (hikers, bikers, off-road vehicles, equestrians) 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/directed energy weapons tester is the accurate determination of energy levels and location on the target. Measurements of the 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 a laser designator did indeed illuminate the target. Such determinations are important when the tester is trying to attribute weapon performance to behavior of the directed energy weapon, behavior of the guidance system, or some other factor.

A bibliography of Army test procedures, Test and Evaluation Command (ATEC) Pamphlet 310-4, lists several documents that cover the special issues associated with laser testing.

24.3 SPACE SYSTEM TESTING

From a historical perspective, space system acquisition has posed several unique problems to the test process (especially the Operational Test (OT) process) that generally fall into four categories: limited quantities/high cost, "incremental upgrade" approach to acquisition, operating environment (peacetime and wartime), and test environment. Expanded Air Force guidance can be found in *Space Systems Test and Evaluation Process*.[2] More generic guidance is in National Security Space (NSS) Acquisition Policy that outlines a streamlined acquisition framework.[3]

(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 "incremental upgrade" approach to acquisition.

(2) **Incremental approach to acquisition:** Due to the "limited buy" and "high per-unit cost" nature of spacecraft acquisition, these systems tend to be procured using an individual increment 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 Operational Test and Evaluation (OT&E)], discrepancies are corrected by "increment 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 wartime 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. But this is no longer valid. 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. This situation is changing with the advent of concepts like the Missile Defense Agency [MDA] system where actual weapons systems (impact anti-satellite and laser) may be in operation, 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 is changing the thrust of how space systems need to be tested in space. The Department of Defense (DoD) should anticipate an increased need for dedicated on-orbit testing on a type of space range where the battlefield environment that 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 warfighting 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.

24.4 OPERATIONS SECURITY AND T&E

Operations Security (OPSEC) issues must be considered in all test planning. Security regulations and contracting documents require 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 *Management of Operational Test and Evaluation*.[4]

24.5 ADVANCED CONCEPT TECHNOLOGY DEMONSTRATIONS

Systems with potential utility for the user and having relatively mature technology may be evaluated by a user in an operational field environment. These programs are not an acquisition program and therefore are not subject to the normal acquisition T&E processes. A favorable evaluation may result in the decision to acquire additional systems for Service use, bypassing a number of the normal acquisition phases. The Services have been using their operational test agencies to assist the field commanders in structuring an evaluation process that would provide the documented data necessary for an informed acquisition decision.

24.6 SUMMARY

All weapon systems tests are limited to some degree, but certain systems face major limitations that could preclude a comprehensive and realistic evaluation. 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.

ENDNOTES

1. U.S. Army Test and Evaluation Command (ATEC) Pamphlet 310-4, *Index of Test Operations Logistics Support: Developmental Supportability Test and Evaluation Guide.*

2. Air Force Manual (AFM) 99-113, *Space Systems Test and Evaluation Process,* May 1, 1996.

3. National Security Space (NSS) Acquisition Policy 03-01, October 3, 2003.

4. Air Force Regulation (AFR) 55-43, *Management of Operational Test and Evaluation.*

25

BEST PRACTICES IN T&E OF WEAPON SYSTEMS

25.1 INTRODUCTION

Numerous pre-millennium 2000 studies were conducted by various agencies that highlighted different perspectives on best practices for Test and Evaluation (T&E). The Office of the Secretary of Defense (OSD) published their study, *Best Practices Applicable to DoD Developmental Test and Evaluation*.[1] The Executive Summary stated "While the study team found no 'Silver Bullets,' it did identify some 20 practices used by commercial enterprises that are relevant to ODTSE&E [Office of the Director, Test Systems Engineering and Evaluation] business practices. These practices have been grouped under the categories 'Policy,' 'Planning,' 'Test Conduct,' and 'Test Analysis'."

Shortly thereafter in September 1999, the Defense Science Board (DSB) Task Force released its report on a broad review of the entire range of activities relating to T&E. Their summary recommendations were: start T&E early—very early; make T&E part of the acquisition process—not adversarial to it; consolidate Development Test (DT) and Operational Test (OT); provide joint test leadership; fund Modeling and Simulation (M&S) support of T&E in program budgets; maintain independence of evaluation process while integrating all other activities; and, establish range ownership and operation structure separate from the Service DT/OT organizations. A follow-on DSB report addressed the value of testing, management of T&E resources, quality of testing, specific T&E investments, and the use of training facilities/exercises for T&E events.[2]

In the same time frame, A. Lee Battershell released a study for the National Defense University (NDU) comparing the acquisition practices of the Boeing 777 and the Air Force C-17. Her most interesting yet not surprising conclusion was that some commercial best practices do not transfer to government.

This was followed by the publication of the General Accounting Office (GAO) Report *Best Practices: A More Constructive Test Approach is Key to Better Weapon System Outcomes.*"[3] After comparing commercial and defense system development practices, the three main findings were: problems found late in development signal weakness in testing and evaluation; testing early to validate product knowledge is a best practice; and, different incentives make testing a more constructive factor in commercial programs than in weapon system programs.

The following information offers guidance to Department of Defense (DoD) personnel who plan, monitor, and execute T&E. Checklists found in the remainder of the chapter were obtained from the *Report of Task Force on Test and Evaluation*.[4] This excellent study is highly regarded in the T&E community but has become somewhat dated; consequently, the Defense Acquisition University (DAU) decided to update the study findings and include those findings and summary checklists in this guide. This discussion follows the system from early concept through the various stages of technology maturation demonstrated in different system/test article configurations.

25.2 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.

25.2.1 Aircraft Systems

25.2.1.1 Concept Assessment

- *Test Program/Total Costs*. Prior to program initiation, 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.* 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 assessment.

- *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.

25.2.1.2 Prototype Development

- *T&E Strategy.* 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 Operational Test and Evaluation (OT&E) should be specified.

- *Test Program Building Blocks.* In testing the prototype, demonstrate that high-risk technology is in hand. In planning the system level 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.

- *Development Test and Evaluation (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 multi-Service testing during OT&E and prior to Low Rate Initial Production (LRIP), the test plan should include the types 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 missions.

- *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 Engineering Development Model (EDM) and production aircraft. Thus, test results that are derived from the prototype in the interim period prior to availability of the EDM 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 test planning.

- *Logistics T&E.* Adequate resources should be scheduled for the aircraft logistics system T&E, and a positive program should exist for the utilization of this 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.

25.2.1.3 Engineering Development Model

- *Test Design.* Test programs should be designed to have a high probability of early identification of major deficiencies during the DT&E and Initial Operational Test and Evaluation (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 Research and Development (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 tests 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 (MOEs) 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 OT plans should be evaluated to determine that the test-imposed conditions on the crew do not invalidate the applicability of the collected data.

- *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.

- *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.

- *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 OT Plans.* Ensure that OT plans are reviewed and updated, as needed, to make them relevant to evolving concepts.

- *Plan OT&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.

25.2.1.4 Production (LRIP and Full Rate), Deployment and Operational Support

- *OT Realism.* Assure IOT&E and Follow-on Test and Evaluation (FOT&E) are conducted under realistic conditions.

- *Design FOT&E for Less-Than-Optimal Conditions*. Structure the FOT&E logistical support for simulated combat conditions.

- *New Threat.* Be alert to the need to extend the IOT&E if a new threat shows up. Address IOT&E limitations in FOT&E.

- *Certification of Ordnance.* Ensure that ordnance to be delivered by an aircraft is certified for the aircraft.

- *Inadvertent Influence of Test.* The IOT&E/FOT&E plans should provide measures for 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 that extensive IOT&E/FOT&E probably would not uncover.

- *Lead the Fleet.* Accelerated Service test of a small quantity of early production aircraft is advisable and periodically during FOT&E thereafter.

25.2.2 Missile Systems

25.2.2.1 Concept Assessment

- *Weapon System Interfaces*. Consider significant weapon system interfaces, their test requirements and probable costs at the outset of the concept assessment. Ensure that the program plan assembled before program start 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.

- *T&E 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 Component Tests.* Results of tests conducted during the concept assessment and the prototype testing, 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 early Test and Evaluation Master Plan (TEMP) should include the type of tests and resources required from other activities and Services.

- *Test Facilities and Instrumentation Requirements.* The test facilities and instrumentation requirements to conduct tests should be generally identified early along with a tentative schedule of test activities.

25.2.2.2 Prototype Testing

- *Establish Test Criteria.* By the end of prototype testing, 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 TEMP 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 (ITP).* Ensure development of an integrated system test plan that preestablishes milestones and goals for easy measurement of program progress at a later time.

- *T&E 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. T&E 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 (PM) 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.

25.2.2.3 Engineering Development Model

- *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 (DTs) should be determined by the primary objectives of that test. Whenever a non-operational configuration is dictated by early test requirements, tests should not be challenged by the fact that configuration is not operational.

- *Testing of Software.* T&E 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 prototype development.

- *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,
  — Safing, 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 OT 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 versus 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 Production Qualification Testing (PQT) 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 production.

- *Flight Tests Representative of Operations.* Ascertain that final DT&E system tests and planned IOT&E flight tests are representative of operational flights. Some ballistic missile R&D programs have shown high success rates in R&D 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.

- *System Interfaces in OT.* Ensure the primary objective of an OT planning 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.

25.2.2.4 Production (LRIP, Full Rate), Deployment and Operational Support

- *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 ASM 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 OT&E for New Threats.* Be alert to the need to extend the IOT&E/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.

25.2.3 Command and Control (C^2) Systems

25.2.3.1 Concept Assessment

- *Concept Test Philosophy.* The T&E planners must understand the nature of Command and Control (C^2) systems early in the concept assessment. In a complex C^2 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 C^2 systems development.

- *Software Test Scheduling – Contractors' Facilities.* Provision should be made for including software T&E during each phase of C^2 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 C^2 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 Program Start.* A Joint-Service OT&E (multi-Service) T&E strategy should be considered for C^2 systems.

25.2.3.2 Prototype Testing

- *Test Prototypes.* In C^2 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 prototype testing should address the functional issues of a C^2 system.

- *Real-Time Software: Demonstration of "Application Patches."* Tests of real-time C^2 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 C^2 systems and to interfaces between subsystems. Particular attention should be devoted to interfaces with other C^2 systems and to the interfaces between sensors (e.g., radar units), communications systems (e.g., modems), and the specific processors (e.g., Central Processing Units (CPUs)). Interface with information processing C^2 systems must also address data-element and code-standardization problems if data is to be processed online.

- *Human Factors.* In a C^2 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 C^2 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 C^2 systems is needed early in the prototype development.

- *Software-Hardware Interfaces.* The software-hardware interfaces, with all operational backup modes to a new C^2 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 are needed during the prototype development for input to cost-effectiveness analyses of C^2 systems.

25.2.3.3 Engineering Development Model

- *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 C^2 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.

- *Displays.* The display subsystems of a C^2 system should provide an essential function to the user. Displays are key subsystems of a C^2 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 C^2 systems.

- *Subsystem Tests.* Every major subsystem of a C^2 system should have a successful DT&E before beginning overall system operational testing.

- *Communications.* The C^2 systems must be tested in the appropriate electromagnetic environment to determine the performance of its communications system.

- *Demonstration of Procedures.* Test plans should include a procedural demonstration whereby the tested C^2 system works in conjunction with other systems.

- *GFE and Facilities*. T&E should be concerned about the availability of GFE as specified in the proposed contract.

- *User Participation in T&E*. The varying needs of the user for a C^2 system make participation in all phases of T&E mandatory.

25.2.3.4 Production (LRIP, Full Rate), Deployment and Operational Support

- *Critical Issues.* IOT&E should be designed during early planning to provide the answers to some critical issues peculiar to C^2 systems. Some of these critical issues that IOT&E of C^2 systems should be planned to 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 location.)
  - — 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?

- *IOT&E Reliability Data.* The IOT&E can provide valuable data on the operational reliability of a C^2 system; this data cannot be obtained through DT&E.

- *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 C^2 system on overall mission effectiveness.

- *Imitative Deception.* The IOT&E should provide for tests to assess the susceptibility of the data links of a C^2 system to imitative deception.

- *First Article Testing (FAT).* The pre-production, FAT 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.

25.2.4 Ship Systems

25.2.4.1 Concept Assessment

- *TEMP.* Prior to program initiation, 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 prototype test and evaluation address the major critical issues, especially technological issues.

- *Test Facilities and Instrumentation Required.* 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 versus Old System.* The procedure for examining the relative performance of new or modified systems versus 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 MOEs.* During the concept assessment 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 MOEs, 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.

25.2.4.2 Prototype Testing

- *Authentication of Human Factors Concepts*. T&E 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 GFE 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.

- *Waivers to T&E of Ship Systems*. Waivers to T&E of pre-production 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.

25.2.4.3 Engineering Development Model

- *Initial or Pilot Phase of IOT&E.* Before any OTs to demonstrate operational suitability and effectiveness are conducted, an initial or pilot test should be conducted (Technical Evaluation (TECHEVAL)).

- *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.* T&E 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.* T&E 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 (FRP) and incorporation of the major weapons subsystems.

25.2.4.4 Production (LRIP, Full Rate), Deployment and Operational Support

- *The OT&E of Shipboard Gun Systems.* The OTs 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.

- *Operational Reliability.* The OT&E should provide valuable data on the operational reliability of ship weapon systems that cannot be obtained through DT&E.

- *Targets for Antiaircraft Warfare (AAW) OT&E.* The OT of shipboard AAW weapons demands the use of targets that realistically simulate the present-day threat.

- *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. This will cause post deployment FOT&E.

- *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 they are completed.

- *Effect of Major Electronic Countermeasures (ECM) 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.* FOT&E 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 over-design as well as minimum design adequacy of the interlock systems.

- *Intra-ship Communication*. In conducting lead ship trials and evaluations, particular attention should be given to the operational impact resulting from absence, by design, of intra-ship communications circuits and stations from important operating locations.

25.2.5 Surface Vehicle Systems

25.2.5.1 Concept Assessment

- *Preparing Test Plans.* It is necessary that detailed evaluation criteria be established that includes all items to be tested.

- *Test Plans.* Prior to program initiation, 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 advanced engineering development 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 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.* Test range and resource requirements to conduct validation tests and a tentative schedule of test activities should be identified.

25.2.5.2 Prototype Testing

- *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,
  — A further refinement of the cost estimates that were initiated during the Concept Evaluation Phase.

- *Prototype Tests.* Prototype tests should show satisfactory meeting of success criteria that are meaningful in terms of operational usage. It is essential in designing contractually required 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.

25.2.5.3 Engineering Development Model

- *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.

25.2.5.4 Production (LRIP, Full Rate), Deployment and Operational Support

- *Planning the IOT&E.* The IOT&E should be planned to be cost-effective and provide meaningful results.

- *Performance and Reliability Testing*. The production FAT 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.

ENDNOTES

1. Office of the Secretary of Defense (OSD), *Best Practices Applicable to DoD [Department of Defense] Developmental Test and Evaluation.* Conducted by the Science Applications International Corporation (SAIC), June 1999.

2. Defense Science Board Report, *Test and Evaluation Capabilities*, December 2000.

3. GAO/NSIAD-00-199, *Best Practices: A More Constructive Test Approach is Key to Better Weapon System Outcomes,* July 2000.

4. Defense Science Board Study, *Report of Task Force on Test and Evaluation,* April 2, 1974.

APPENDICES

APPENDIX A
ACRONYMS AND THEIR MEANINGS

Aa	Achieved availability
AAE	Army Acquisition Executive
AAH	Advanced Attack Helicopter
AAW	Antiaircraft Warfare
ACAT	Acquisition Category
ACE	Army Corps of Engineers
ACTD	Advanced Concept Technology Demonstration
ADM	Acquisition Decision Memorandum
ADP	Automated Data Processing
ADPE	Automated Data Processing Equipment
AEC	Army Evaluation Center
AF	Air Force
AFCEA	Armed Forces Communications and Electronics Association
AFEWES	Air Force Electronic Warfare Evaluation Simulator
AFFTC	Air Force Flight Test Center
AFM	Air Force Manual
AFMC	Air Force Materiel Command
AFOTEC	Air Force Operational Test and Evaluation Center
AFR	Air Force Regulation
AFSC	Air Force Space Command
AF/TE	Chief of Staff, Test and Evaluation (Air Force)
AFV	Armored Family of Vehicles
Ai	Inherent Availability
AIS	Automated Information System
ALCM	Air-Launched Cruise Missile
AMC	Army Materiel Command
AMSDL	Acquisition Management Systems and Data Requirements Control List
Ao	Operational Availability

AoA Analysis of Alternatives

AP Acquisition Plan

APB Acquisition Program Baseline

AR Army Regulation

ARL-HRED Army Research Laboratory-Human Research and Engineering Division/Directorate

ASA(ALT) Assistant Secretary of the Army for Acquisition, Logistics, and Technology

ASD(NII) Assistant Secretary of Defense for Networks and Information Integration

ASM Air-to-Surface Missile

ASN(RD&A) Assistant Secretary of the Navy for Research, Development, and Acquisition

ASN/RE&S Assistant Secretary of the Navy/Research, Engineering, and Science

ASR Alternative System Review

ASW Antisubmarine Warfare

ATD Advanced Technology Demonstration (or Development)

ATE Automatic Test Equipment

ATEC Army Test and Evaluation Command

ATIRS Army Test Incident Reporting System

ATP Automated Test Plan

ATPS Automated Test Planning System

ATR Acceptance Test Report

AWACS Airborne Warning and Control System

BA Budget Activity; Budget Authority

BIT Built-in Test

BITE Built-in Test Equipment

BLRIP Beyond Low Rate Initial Production

BoD Board of Directors

BoOD Board of Operating Directors

C^2 Command and Control

C^3 Command, Control, and Communications

C^3I	Command, Control, Communications, and Intelligence
C^4	Command, Control, Communications, and Computers
C^4I	Command, Control, Communications, Computers, and Intelligence
C^4ISR	Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance
CAD	Computer-Aided Design
CAE	Component Acquisition Executive
CAIV	Cost as an Independent Variable
CAM	Computer-Aided Manufacturing
CARS	Consolidated Acquisition Reporting System
CDD	Capability Development Document
CDR	Critical Design Review
CDRL	Contract Data Requirements List
CE	Continuous Education; Concept Exploration
CEP	Circle Error Probability; Concept Experimentation Program
CERDEC	Communications-Electronics Research Development and Engineering Center
CFR	Code of Federal Regulations
CG MCSC	Commanding General, Marine Corps Systems Command
CI	Configuration Item
CII	Compatibility, Interoperability, and Integration
CIO	Chief Information Officer
CJCSI	Chairman of the Joint Chiefs of Staff Instruction
CJCSM	Chairman of the Joint Chiefs of Staff Manual
CLIN	Contract Line Item Number
CNO	Chief of Naval Operations
CNP	Candidate Nomination Proposal
COCOM	Combatant Commander
COEA	Cost and Operational Effectiveness Analysis
COI	Critical Operational Issue
COIC	Critical Operational Issues and Criteria

COMOPTEVFOR Commander, Operational Test and Evaluation Force (Navy)
COO Concept of Operations
CPD Capability Production Document
CPS Competitive Prototyping Strategy
CPU Central Processing Unit
CR Concept Refinement
CSCI Computer Software Configuration Item
CTEIP Central Test and Evaluation Investment Program
CTO Comparative Test Office
DA Department of the Army
DAB Defense Acquisition Board
DAE Defense Acquisition Executive
DAES Defense Acquisition Executive Summary
DAG Data Authentication Group
DAS Director of the Army Staff
DAU Defense Acquisition University
DBDD Data Base Design Document
DCMA Defense Contract Management Agency
DCS Deputy Chief of Staff
DCS/R&D Deputy Chief of Staff for Research and Development
DCSOPS Deputy Chief of Staff for Operations
DDR&E Director of Defense Research and Engineering
DIA Defense Intelligence Agency
DID Data Item Description
DISA Defense Information Systems Agency
DLT Design Limit Test
DMSO Defense Modeling and Simulation Office
DNA Defense Nuclear Agency
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

DOTLPF Doctrine, Organization, Training, Leadership, Personnel, and Facilities

DPRO Defense Plant Representative Office

DRR Design Readiness Review

DS Defense Systems

DSARC Defense Systems Acquisition Review Council (now the Defense Acquisition Board (DAB))

DSB Defense Science Board

DSMC Defense Systems Management College

DT Development Test

DT&E Development Test and Evaluation

DT&E/DS Development Test and Evaluation/Defense Systems

DTC Developmental Test Command (Army)

DTIC Defense Technical Information Center

DTRMC Defense Test Resource Management Center

DTSE&E Director, Test Systems Engineering and Evaluation

DTTSG Defense Test and Training Steering Group

DUSA(OR) Deputy Under Secretary of the Army (Operations Research)

DUSD(A&T) Deputy Under Secretary of Defense for Acquisition and Technology

DVAL Data Link Vulnerability Analysis

EA Evolutionary Acquisition

EC Electronic Combat

ECCM Electronic Counter-Countermeasures

ECM Electronic Countermeasures

ECP Engineering Change Proposal

ECR Engineering Change Review

EDM Engineering Development Model

EDP Event Design Plan (Army)

EDT Engineering Design Test

EFA Engineering Flight Activity

EIS	Environmental Impact Statement
EMCS	Electromagnetic Control Study
EMI	Electromagnetic Interference
EMP	Electromagnetic Pulse
EOA	Early Operational Assessment
ERAM	Extended Range Anti-armor Munitions
ESM	Electronic Support Measures
ESOH	Environmental Safety and Occupational Health
ESS	Environmental Stress Screening
EW	Electronic Warfare
FA/A	Functional Analysis/Allocation
FACA	Federal Advisory Committee Act
FAR	Federal Acquisition Regulation
FAT	First Article Testing
FCA	Functional Configuration Audit
FCT	Foreign Comparative Testing
FDTE	Force Development Tests and Experimentation
FFBD	Functional Flow Block Diagram
FMECA	Failure Mode, Effects, and Criticality Analysis
FOC	Full Operational Capability
FORSCOM	Forces Command (Army)
FOT&E	Follow-on Operational Test and Evaluation
FQR	Formal Qualification Review
FQT	Formal Qualification Test/Testing
FRP	Full Rate Production
FRPDR	Full Rate Production Decision Review
FWE	Foreign Weapons Evaluation
FY	Fiscal Year
FYDP	Future Years Defense Program
FYTP	Future Years Test Program; Five Year Test Program (Army)
GAO	General Accounting Office (now Government Accountability Office)

GFE	Government Furnished Equipment
HELSTF	High Energy Laser System Test Facility
HQDA	Headquarters, Department of the Army
HSI	Human Systems Integration
HW	Hardware
HWCI	Hardware Configuration Item
HWIL	Hardware-in-the-Loop
ICBM	Intercontinental Ballistic Missile
ICD	Initial Capabilities Document
ICE	Independent Cost Estimate
IDD	Interface Decision Document
IEP	Independent Evaluation Plan
IER	Independent Evaluation Report
IFPP	Information for Proposal Preparation
ILS	Integrated Logistics Support
INSCOM	Intelligence and Security Command
IOA	Independent Operational Assessment
IOC	Initial Operating Capability
IOT&E	Initial Operational Test and Evaluation
IPPD	Integrated Product and Process Development
IPT	Integrated Product Team
IR&D	International Research and Development
IRA	Industrial Resource Analysis
IRS	Interface Requirements Specification
IT	Information Technology
ITEA	International Test and Evaluation Association
ITP	Integrated Test Plan
ITOP	International Test Operating Procedure
ITSCAP	Information Technology Security Certification and Accreditation Program
ITTS	Instrumentation, Targets, and Threat Simulators

IV&V Independent Verification and Validation

JCIDS Joint Capabilities Integration and Development System

JITC Joint Interoperability Test Command

JLF Joint Live Fire

JROC Joint Requirements Oversight Council

JTCG Joint Technical Coordinating Group

JT&E Joint Test and Evaluation

JTTR Joint Training and Test Range Roadmap

KPP Key Performance Parameter

Kr Contractor

LCC Life Cycle Cost

LCCE Life Cycle Cost Estimate

LFT Live Fire Testing

LFT&E Live Fire Test and Evaluation

LRIP Low Rate Initial Production

LS Logistics Support

LSA Logistics Support Analysis

LSP Logistics Support Plan

LTBT Limited Test Ban Treaty

M&S Modeling and Simulation

MAIS Major Automated Information System

MAJCOM Major Commands

MANPRINT Manpower and Personnel Integration

MAV Minimum Acceptable Value

MCOTEA Marine Corps Operational Test and Evaluation Activity

MCP Military Construction Program

MCSC Marine Corps Systems Command

MDA Milestone Decision Authority; Missile Defense Agency

MDAP Major Defense Acquisition Program

MEDCOM Medical Command (Army)

MIL-HDBK Military Handbook

MIL-SPEC	Military Specification
MIL-STD	Military Standard
MOA	Memorandum of Agreement
MOE	Measure of Effectiveness
MOP	Measure of Performance
MOS	Military Occupational Specialty
MOT&E	Multi-Service Operational Test and Evaluation
MOU	Memorandum of Understanding
MPE	Military Preliminary Evaluation
MRTFB	Major Range and Test Facility Base
MS	Milestone
MSIAC	Modeling and Simulation Information Analysis Center
MTBF	Mean Time Between Failure
MTTR	Mean Time To Repair
NAPMA	North Atlantic Program Management Agency
NATO	North Atlantic Treaty Organization
NAVAIR	Naval Air Systems Command
NAVSEA	Naval Sea Systems Command
NAWC	Naval Air Warfare Center
NBC	Nuclear, Biological, and Chemical
NBCC	Nuclear, Biological, and Chemical Contamination
NDAA	National Defense Authorization Act
NDCP	Navy Decision Coordinating Paper
NDI	Nondevelopment(al) Item
NDU	National Defense University
NEPA	National Environmental Policy Act
NH&S	Nuclear Hardness and Survivability
NSIAD	National Security and International Affairs Division
NSS	National Security Systems
O&M	Operations and Maintenance
O&S	Operations and Support

OA Operational Assessment

OCD Operational Concept Document

OIPT Overarching Integrated Product Team

OMB Office of Management and Budget

OPEVAL Operational Evaluation

OPNAV Operational Navy

OPNAVIST Operational Navy Instruction

OPSEC Operations Security

OPTEVFOR Operational Test and Evaluation Force (Navy)

ORMAS/TE Operational Resource Management Assessment Systems for Test and Evaluation

OSD Office of the Secretary of Defense

OT Operational Test

OT&E Operational Test and Evaluation

OTA Operational Test Agency

OTC Operational Test Command (Army)

OTD Operational Test Director

OTRR Operational Test Readiness Review

OTP Outline Test Plan

P^3I Preplanned Product Improvements

P&D Production and Deployment

PAT&E Production Acceptance Test and Evaluation

PCA Physical Configuration Audit

PCO Primary Contracting Officer

PDR Preliminary Design Review

PDRR Program Definition and Risk Reduction

PDUSD(A&T) Principal Deputy Under Secretary of Defense for Acquisition and Technology

PE Program Element

PEO Program Executive Officer

PEP Production Engineering and Planning

PF/DOS Production, Fielding/Deployment, and Operational Support

PI	Product Improvement
Pk	Probability of kill
Pk/h	Probability of kill given a hit
PM	Program Manager; Product Manager
PMO	Program Management Office
PO	Program Office, Purchase Order
POM	Program Objectives Memorandum
PPBE	Planning, Programming, Budgeting and Execution
PQT	Production Qualification Test
PRAT	Production Reliability Acceptance Test
PRESINSURV	President of the Board of Inspection and Survey
PRR	Production Readiness Review
PSA	Principal Staff Assistant
QA	Quality Assurance
QOT&E	Qualification Operational Test and Evaluation
R&D	Research and Development
R&E	Research and Engineering
R&M	Reliability and Maintainability
RAM	Reliability, Availability, and Maintainability
RAS	Requirements Allocation Sheet
RCC	Range Commanders Council
RCS	Radar Cross Section
RD&A	Research, Development, and Acquisition
RDECOM	Research, Development and Engineering Command (Army)
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
RSI	Rationalization, Standardization, and Interoperability

RSM Radar Spectrum Management

S&TS Strategic and Tactical Systems

SAF/AQ Secretary of the Air Force for Acquisition

SAF/USA Director of Space Acquisition

SAR Selected Acquisition Report

SAT Ship Acceptance Test

SDD Software Design Document; System Development and Demonstration

SECARMY Secretary of the Army

SECDEF Secretary of the Defense

SECNAV Secretary of the Navy

SEP Systems Engineering Process; System Engineering Plan; System Evaluation Plan (Army)

SFR System Functional Review

SI Systems Integration

SIL Software Integration Laboratory

SMDC Space and Missile Defense Command

SOCOM Special Operations Command

SOO Statement of Objectives

SOW Statement of Work

SPAWAR Space and Naval Warfare Systems Command

SPEC Specification

SPO System Program Office

SRR Systems Requirements Review

SRS Software Requirements Specification

SSR Software Specification Review

STA System Threat Assessment

STEP Simulation, Test and Evaluation Process

STP Software Test Plan

STRI Simulation, Training, and Instrumentation (Army)

SQA Software Quality Assurance

SVR System Verification Review

SW Software

SWIL Software in the Loop

T&E Test and Evaluation

TAAF Test, Analyze, and Fix

TADS Target Acquisition Designation System; Theater Air Defense System

TAFT Test, Analyze, Fix, and Test

TD Technology Development

TDS Technology Development Strategy

TDY Temporary Duty

TEAM Test, Evaluation, Analysis, and Modeling

TECHEVAL Technical Evaluation (Navy Term)

TEMA Test and Evaluation Management Agency

TEMP Test and Evaluation Master Plan

TERC Test and Evaluation Resources Committee

TFRD Test Facility Requirements Document

TIRIC Training Instrumentation Resource Investment Committee

TLS Time Line Sheet

TM Technical Manual, Test Manager

TMC Test Management Council

TMP Technical Management Plan

TPD Test Program Documentation

TPO Test Program Outline

TPM Technical Performance Measurement

TPWG Test Planning Working Group

TRADOC Training and Doctrine Command

TRIMS Test Resource Information Management System

TRP Test Resources Plan

TRR Test Readiness Review

TRS Test Requirements Sheet

TSARC Test Schedule and Review Committee

UNK Unknown

U.S.	United States
USA	United States Army
USAF	United States Air Force
USAFE/DOQ	U.S. Air Forces in Europe/Directorate of Operations-Operations
USAKA	United States Army Kwajalein Atoll
USD(AT&L)	Under Secretary of Defense for Acquisition, Technology, and Logistics
USMC	United States Marine Corps
USN	United States Navy
U.S.C.	United States Code
VCJCS	Vice Chairman, Joint Chiefs of Staff
VCSA	Vice Chief of Staff, Army
VECP	Value Engineering Change Proposal
VV&A	Verification, Validation and Accreditation
WBS	Work Breakdown Structure
WIPT	Working-level Integrated Product Team
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.

ACQUISITION — The conceptualization, initiation, design, development, test, contracting, production, deployment, Logistics Support (LS), modification, and disposal of weapons and other systems, supplies, or services (including construction) to satisfy Department of Defense (DoD) needs, intended for use in or in support of military missions.

ACQUISITION CATEGORY (ACAT) — ACAT I programs are Major Defense Acquisition Programs (MDAPs). An MDAP is defined as a program that is not highly classified and is designated by the Under Secretary of Defense (Acquisition, Technology, and Logistics) (USD(AT&L)) as an MDAP or is estimated by USD(AT&L) to require eventual total expenditure for research, development, test and evaluation of more than $365 million Fiscal Year (FY) 2000 constant dollars, or for procurement of more than $2.190 billion in FY2000 constant dollars.

*1. ACAT ID for which the Milestone Decision Authority (MDA) is USD(AT&L). The "D" refers to the Defense Acquisition Board (DAB), which advises the USD(AT&L) at major decision points.

2. ACAT IC for which the MDA is the DoD Component Head or, if delegated, the DoD Component Acquisition Executive (CAE). The "C" refers to Component.

The USD(AT&L) designates programs as ACAT ID or ACAT IC.

ACAT IA programs are Major Automated Information Systems (MAISs). A MAIS is any program designated by the Assistant Secretary of Defense for Command, Control, Communications, and Intelligence (ASD(C^3I)) as an MAIS or is estimated to require program costs for any single year in excess of $32 million in Fiscal Year (FY) 2000 constant dollars, total program in excess of $126 million in FY2000 constant dollars, or total life cycle costs in excess of $378 million FY2000 constant dollars.

MAIS do not include highly sensitive classified programs or tactical communications systems.

ACAT II program is a major system that is a combination of elements that shall function together to produce the capabilities required to fulfill a mission need, excluding construction or other improvements to real property. It is estimated by the DoD Component head to require eventual total expenditure for Research, Development, Test, and Evaluation (RDT&E) of more than $140 million in FY2000 constant dollars, or procurement of more than $660 million in FY2000 constant dollars, or is designated as major by the DoD Component head.

ACAT III programs are defined as those acquisition programs that do not meet the criteria for an ACAT I, an ACAT IA, or an ACAT II. The MDA is designated by the CAE and shall be at the lowest appropriate level. This category includes less-than-major AISs.

ACQUISITION DECISION MEMORANDUM (ADM) — A memorandum signed by the Milestone Decision Authority (MDA) that documents decisions made as the result of a milestone decision review or in-process review.

ACQUISITION LIFE CYCLE — The life of an acquisition program consists of phases, each preceded by a milestone or other decision point, during which a system goes through research, development, test and evaluation, and production. Currently, the four phases are: (1) Concept Exploration (CE) (Phase 0); Program Definition and Risk Reduction (PDRR) (Phase I); (3) Engineering and Manufacturing Development (EMD) (Phase II); and (4) Production, Fielding/Deployment, and Operational Support (PF/DOS) (Phase III).

ACQUISITION PHASE — All the tasks and activities needed to bring a program to the next major milestone occur during an acquisition phase. Phases provide a logical means of progressively translating broadly stated mission needs into well-defined system-specific requirements and ultimately into operationally effective, suitable, and survivable systems.

ACQUISITION PROGRAM BASELINE (APB) — A document that contains the most important cost, schedule, and performance parameters (both objectives and thresholds) for the program. It is approved by the Milestone Decision Authority (MDA), and signed by the Program Manager (PM) and their direct chain of supervision, e.g., for Acquisition Category (ACAT) ID programs it is signed by the PM, Program Executive Officer (PEO), Component Acquisition Executive (CAE), and Defense Acquisition Executive (DAE).

ACQUISITION STRATEGY — A business and technical management approach designed to achieve program objectives within the resource constraints imposed. It is the framework for planning, directing, contracting for, and managing a program. It provides a master schedule for research, development, test, production, fielding, modification, postproduction management, and other activities essential for program success. Acquisition strategy is the basis for formulating functional plans and strategies (e.g., Test And Evaluation Master Plan (TEMP), Acquisition Plan (AP), competition, prototyping, etc.).

ACQUISITION RISK — See Risk.

ADVANCED CONCEPT TECHNOLOGY DEMONSTRATION (ACTD) — Shall be used to determine military utility of proven technology and to develop the concept of operations that will optimize effectiveness. ACTDs themselves are not acquisition programs, although they are designed to provide a residual, usable capability upon completion. Funding is programmed to support 2 years in the field. ACTDs are funded with 6.3a (Advanced Technology Development (or Demonstration) (ATD)) funds.

ADVANCED TECHNOLOGY DEVELOPMENT (OR DEMONSTRATION) (ATD) (Budget Activity 6.3) — Projects within the 6.3a (ATD) program that are used to demonstrate the maturity and potential of advanced technologies for enhanced military operational capability or cost effectiveness. It is intended to reduce technical risks and uncertainties at the relatively low costs of informal processes.

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.

ANALYSIS OF ALTERNATIVES (AoA) — 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. Formerly known as Cost and Operational Effectiveness Analysis (COEA).

AUTOMATED INFORMATION SYSTEM (AIS) — A combination of computer hardware and software, data, or telecommunications that performs functions such as collecting, processing, transmitting, and displaying information. Excluded are computer resources, both hardware and software, that are physically part of, dedicated to, or essential in real time to the mission performance of weapon systems.

AUTOMATIC TEST EQUIPMENT (ATE) — Equipment that is designed to automatically conduct analysis of functional or static parameters and to evaluate the degree of performance degradation and perform fault isolation of unit malfunctions.

BASELINE — Defined quantity or quality used as starting point for subsequent efforts and progress measurement that can be a technical cost or schedule baseline.

BRASSBOARD CONFIGURATION — An experimental device (or group of devices) used to determine feasibility and to develop technical and operational data. It will normally 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 to develop technical data. It will normally 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.

CAPSTONE TEST AND EVALUATION MASTER PLAN (TEMP) — A TEMP which addresses the testing and evaluation of a program consisting of a collection of individual systems (family of systems, system of systems) which function collectively. Individual system-unique content requirements are addressed in an annex to the basic Capstone TEMP.

CERTIFICATION FOR INITIAL OPERATIONAL TEST AND EVALUATION (IOT&E) — A service process undertaken in the advanced stages of system development, normally during Low Rate Initial Production (LRIP), resulting in the announcement of a system's readiness to undergo IOT&E. The process varies with each Service.

COMPETITIVE PROTOTYPING STRATEGY (CPS) — Prototype competition between two or more contractors in a comparative side-by-side test.

CONCEPT EXPERIMENTATION PROGRAM (CEP) — The CEP is an annual program executed by the Training and Doctrine Command (TRADOC) for commanders who have requirements determination and force development missions for specific warfighting experiments. Experiments are discrete events investigating either a material concept or a warfighting idea. The CEP procedures are provided in TRADOC Pamphlet 71-9.

CONCURRENCY — Part of an acquisition strategy that would combine or overlap life cycle phases (such as Engineering and Manufacturing Development (EMD), and production), or activities (such as development and operational testing).

CONTINGENCY TESTING — Additional testing required to support a decision to commit added resources to a program, when significant test objectives have not been met during planned tests.

CONTINUOUS EVALUATION (CE) — 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 — The technical and administrative direction and surveillance actions taken to identify and document the functional and physical characteristics of a Configuration Item (CI), to control changes to a CI and its characteristics, and to record and report change processing and implementation status. It provides a complete audit trail of decisions and design modifications.

CONTRACT — An agreement between two or more legally competent parties, in the proper form, on a legal subject matter or purpose and for legal consideration.

CONTRACTOR LOGISTICS SUPPORT — The performance of maintenance and/or material management functions for a Department of Defense (DoD) system by a commercial activity. Historically done on an interim basis until systems support could be transitioned to a DoD organic capability. Current policy now allows for the provision of system support by contractors on a long-term basis. Also called Long-Term Contractor Logistics Support.

COOPERATIVE PROGRAMS — Programs that comprise one or more specific cooperative projects whose arrangements are defined in a written agreement between the parties and which are conducted in the following general areas:

1. Research, Development, Test, and Evaluation (RDT&E) of defense articles (including cooperative upgrade or other modification of a U.S.-developed system), joint production (including follow-on support) of a defense article that was developed by one or more of the participants, and procurement by the United States of a foreign defense article (including software), technology (including manufacturing rights), or service (including logistics support) that are implemented under Title 22 U.S.C. §2767, Reference (c), to promote the Rationalization, Standardization, and Interoperability (RSI) of North Atlantic Treaty Organization (NATO) armed forces or to enhance the ongoing efforts of non-NATO countries to improve their conventional defense capabilities.

2. Cooperative Research and Development (R&D) program with NATO and major non-NATO allies implemented under Title 10 U.S.C. §2350a, to improve the conventional defense capabilities of NATO and enhance RSI.

3. Data, information, and personnel exchange activities conducted under approved DoD programs.

4. Test and Evaluation (T&E) of conventional defense equipment, munitions, and technologies developed by allied and friendly nations to meet valid existing U.S. military requirements.

COST AS AN INDEPENDENT VARIABLE (CAIV) — Methodologies used to acquire and operate affordable DoD systems by setting aggressive, achievable Life Cycle Cost (LCC) objectives, and managing achievement of these objectives by trading off performance and schedule, as necessary. Cost objectives balance mission needs with projected out-year resources, taking into account anticipated process improvements in both DoD and industry. CAIV has brought attention to the government's responsibilities for setting/adjusting LCC objectives and for evaluating requirements in terms of overall cost consequences.

CRITICAL ISSUES — Those aspects of a system's capability, either operational, technical, or other, that must be questioned before a system's overall suitability can be known. Critical issues are of primary importance to the decision authority in reaching a decision to allow the system to advance into the next phase of development.

CRITICAL OPERATIONAL ISSUE (COI) — Operational effectiveness and operational suitability issues (not parameters, objectives, or thresholds) that must be examined in Operational Test and Evaluation (OT&E) to determine the system's capability to perform its mission. A COI is normally phrased as a question that must be answered in order to properly evaluate operational effectiveness (e.g., "Will the system detect the threat in a combat environment at adequate range to allow successful engagement?") or operational suitability (e.g., "Will the system be safe to operate in a combat environment?").

DATA SYSTEM — Combinations of personnel efforts, forms, formats, instructions, procedures, data elements and related data codes, communications facilities, and Automated Data Processing Equipment (ADPE) 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 individual responsible for all acquisition matters within the DoD. (See DoD Directive (DoDD) 5000.1.)

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 DoD. The system encompasses establishing and enforcing policies and practices that govern acquisitions, to include documenting mission needs and establishing performance goals and baselines; determining and prioritizing resource requirements for acquisition programs; planning and executing acquisition programs; directing and controlling the acquisition review process; developing and assessing logistics implications; contracting; monitoring the execution status of approved programs; and reporting to the Congress.

DESIGNATED ACQUISITION PROGRAM — Program designated by the Director, Operational Test and Evaluation (DOT&E) or the Deputy Director, Developmental Test and Evaluation (S&TS) for OSD oversight of test and evaluation.

DEVELOPING AGENCY (DA) — The Systems Command or Chief of Naval Operations (CNO)-designated project manager assigned responsibility for the development, test and evaluation of a weapon system, subsystem, or item of equipment.

DEVELOPMENT TEST AND EVALUATION (DT&E) — T&E conducted throughout the life cycle to identify potential operational and technological capabilities and limitations of the alternative concepts and design options being pursued; support the identification of cost-performance tradeoffs by providing analyses of the capabilities and limitations of alternatives; support the identification and description of design technical risks; assess progress toward meeting Critical Operational Issues (COIs), mitigation of acquisition technical risk, achievement of manufacturing process requirements and system maturity; assess validity of assumptions and conclusions from the Analysis of Alternatives (AoAs); provide data and analysis in support of the decision to certify the system ready for Operational Test and Evaluation (OT&E); and in the case of Automated Information Systems (AISs), support an information systems security certification prior to processing classified or sensitive data and ensure a standards conformance certification.

DEVELOPMENTAL TESTER — The command or agency that plans, conducts, and reports the results of developmental testing. Associated contractors may perform developmental testing on behalf of the command or agency.

EARLY OPERATIONAL ASSESSMENT (EOA) — An operational assessment conducted prior to or in support of prototype testing.

EFFECTIVENESS — The extent to which the goals of the system are attained, or the degree to which a system can be elected to achieve a set of specific mission requirements.

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) — A proposal to the responsible authority recommending that a change to an original item of equipment be considered, and the design or engineering change be incorporated into the article to modify, add to, delete, or supersede original parts.

ENGINEERING DEVELOPMENT — The RDT&E 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.

FIRST ARTICLE — First article includes pre-production models, initial production samples, test samples, first lots, pilot models, and pilot lots; and approval involves testing and evaluating the first article for conformance with specified contract requirements before or in the initial stage of production under a contract.

FIRST ARTICLE TESTING (FAT) — Production testing that is planned, conducted, and monitored by the materiel developer. FAT includes pre-production and initial production testing conducted to ensure that the contractor can furnish a product that meets the established technical criteria.

FOLLOW-ON OPERATIONAL TEST AND EVALUATION (FOT&E) — That test and evaluation that may be necessary after system deployment to refine the estimates made during operational test and evaluation, to evaluate changes, and to re-evaluate the system to ensure that 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.

FOREIGN COMPARATIVE TESTING (FCT) — A DoD T&E program that is prescribed in Title 10 U.S.C. §2350a(g), and is centrally managed by the Director, Foreign Comparative Test (Strategic and Tactical Systems (S&TS)). It provides funding for U.S. T&E of selected equipment items and technologies developed by allied countries when such items and technologies are identified as having good potential to satisfy valid DoD requirements.

FUTURE-YEAR DEFENSE PROGRAM (FYDP) —(Formerly the Five Year Defense Program). The official DoD document that summarizes forces and resources associated with programs approved by the Secretary of Defense (SECDEF). Its three parts are the organizations affected, appropriations accounts (Research, Development, Test, and Evaluation (RDT&E), Operations and Maintenance (O&M), etc.), and the 11 major force programs (strategic forces, airlift/sealift, Research and Development (R&D), etc.). R&D is Program 06. Under the current Planning, Programming, Budgeting and Execution (PPBE) cycle, the FYDP is updated when the Services submit their Program Objective Memorandums (POMs) to the Office of the Secretary of Defense (OSD) (May/June), when the Services submit their budgets to OSD (September), and when the President submits the national budget to the Congress (February). The primary data element in the FYDP is the Program Element (PE).

HARMONIZATION — Refers to the process, or results, of adjusting differences or inconsistencies in the qualitative basic military requirements of the United States, its allies, and other friendly countries. It implies that significant features will be brought into line so as to make possible substantial gains in terms of the overall objectives of cooperation (e.g., enhanced utilization of resources, standardization, and compatibility of equipment). It implies especially that comparatively minor differences in "requirements" should not be permitted to serve as a basis for the support of slightly different duplicative programs and projects.

HUMAN SYSTEMS INTEGRATION (HSI) — A disciplined, unified, and interactive approach to integrate human considerations into system design to improve total system performance and reduce costs of ownership. The major categories of human considerations are manpower, personnel, training, human factors engineering, safety, and health.

INDEPENDENT EVALUATION REPORT (IER) — A report that provides an assessment of item or system operational effectiveness and operational suitability versus 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 Test and Evaluation Command (ATEC), the Navy Operational Test and Evaluation Force (OPTEVFOR), the Air Force Operational Test and Evaluation Center (AFOTEC), the Marine Corps Operational Test and Evaluation Activity (MCOTEA), and for the Defense Information Systems Agency (DISA) – the Joint Interoperability Test Command (JITC).

INDEPENDENT VERIFICATION AND VALIDATION (IV&V) — An independent review of the software product for functional effectiveness and technical sufficiency.

INITIAL OPERATIONAL CAPABILITY (IOC) — The first attainment of the capability to employ effectively a weapon, item of equipment, or system of approved specific characteristics with the appropriate number, type, and mix of trained and equipped personnel necessary to operate, maintain, and support the system. It is normally defined in the Operational Requirements Document (ORD).

INITIAL OPERATIONAL TEST AND EVALUATION (IOT&E) — Operational test and evaluation conducted on production, or production representative articles, to determine whether systems are operationally effective and suitable for intended use by representative users to support the decision to proceed beyond Low Rate Initial Production (LRIP). For the Navy, see Operational Evaluation (OPEVAL).

IN-PROCESS REVIEW — Review of a project or program at critical points to evaluate status and make recommendations to the decision authority.

INSPECTION — Visual examination of the item (hardware and software) and associated descriptive documentation, which compares appropriate characteristics with predetermined standards to determine conformance to requirements without the use of special laboratory equipment or procedures.

INTEGRATED PRODUCT AND PROCESS DEVELOPMENT (IPPD) — A management technique that simultaneously integrates all essential acquisition activities through the use of multidisciplinary teams to optimize the design, manufacturing, and supportability processes. IPPD facilitates meeting cost and performance objectives from product concept through production, including field support. One of the key IPPD tenets is multidisciplinary teamwork through Integrated Product Teams (IPTs).

INTEGRATED PRODUCT TEAM (IPT) — Team composed of representatives from all appropriate functional disciplines working together to build successful programs, identify and resolve issues, and make sound and timely recommendations to facilitate decisionmaking. There are three types of IPTs: Overarching IPTs (OIPTs) focus on strategic guidance, program assessment, and issue resolution; Working-level IPTs (WIPTs) identify and resolve program issues, determine program status, and seek opportunities for acquisition reform; and program-level IPTs focus on program execution and may include representatives from both government and, after contract award, industry.

INTEROPERABILITY — The ability of systems, units, or forces to provide services to or accept services from other systems, units, or forces and to use the services so exchanged to operate effectively together. The conditions achieved among communications-electronics systems or items of communications-electronics equipment when information or services can be exchanged directly and satisfactorily between them and/or their users. Designated an element of the Net-Ready Key Performance Parameter (KPP) by Joint Chiefs of Staff (JCS).

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.

KEY PERFORMANCE PARAMETERS (KPPs) — Those minimum attributes or characteristics considered most essential for an effective military capability. The Capability Development Document (CDD) and the Capability Production Document (CPD) KPPs are included verbatim in the Acquisition Program Baseline (APB).

LIVE FIRE TEST AND EVALUATION (LFT&E) — A test process that is defined in Title 10 U.S.C. §2366, that must be conducted on a covered system, major munition program, missile program, or Product Improvement (PI) to a covered system, major munition program, or missile program before it can proceed beyond Low Rate Initial Production (LRIP). A covered system is a major system (Acquisition Category (ACAT) I, II) that is: 1) user-occupied and designed to provide some degree of protection to its occupants in combat, or ; 2) a conventional munitions program or missile program, or; 3) a conventional munitions program for which more than 1,000,000,000 rounds are planned to be acquired, or; 4) a modification to a covered system that is likely to affect significantly the survivability or lethality of such a system.

LIVE FIRE TEST AND EVALUATION (LFT&E) REPORT — Report prepared by the Director, Operational Test and Evaluation (DOT&E) on survivability and lethality testing. Submitted to the Congress for covered systems prior to the decision to proceed beyond Low Rate Initial Production (LRIP).

LETHALITY — The probability that weapon effects will destroy the target or render it neutral.

LIFE CYCLE COST (LCC) — The total cost to the government for the development, acquisition, operation, and logistics support of a system or set of forces over a defined life span.

LOGISTICS SUPPORTABILITY — The degree of ease to which system design characteristics and planned logistics resources (including the Logistics Support (LS) elements) allow for the meeting of system availability and wartime usage requirements.

LONG LEAD ITEMS — Those 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.

LOT ACCEPTANCE — This test is based on a sampling procedure to ensure that the product retains its quality. No acceptance or installation should be permitted until this test for the lot has been successfully completed.

LOW RATE INITIAL PRODUCTION (LRIP) — The minimum number of systems (other than ships and satellites) to provide production representative articles for Operational Test and Evaluation (OT&E), to establish an initial production base, and to permit an orderly increase in the production rate sufficient to lead to Full Rate Production (FRP) upon successful completion of operational testing. For Major Defense Acquisition Programs (MDAPs), LRIP quantities in excess of 10 percent of the acquisition objective must be reported in the Selected Acquisition Report (SAR). For ships and satellites LRIP is the minimum quantity and rate that preserves mobilization.

MAINTAINABILITY — The ability of an item to be retained in, or restored to, a specified condition when maintenance is performed by personnel having specified skill levels, using prescribed procedures and resources, at each prescribed level of maintenance and repair. (See Mean Time To Repair (MTTR))

MAJOR DEFENSE ACQUISITION PROGRAM (MDAP) — See “Acquisition Category.”

MAJOR SYSTEM (DoD) — See "Acquisition Category."

MAJOR RANGE AND TEST FACILITY BASE (MRTFB) — The complex of major DoD ranges and test facilities managed according to DoD 3200.11 by the Director, Test Resource Management Center reporting to the USD(AT&L).

MEAN TIME BETWEEN FAILURE (MTBF) — For a particular interval, the total functional life of a population of an item divided by the total number of failures within the population. The definition holds for time, rounds, miles, events, or other measures of life unit. A basic technical measure of reliability.

MEAN TIME TO REPAIR (MTTR) — The total elapsed time (clock hours) for corrective maintenance divided by the total number of corrective maintenance actions during a given period of time. A basic technical measure of maintainability.

MEASURE OF EFFECTIVENESS (MOE) — A measure of operational performance success that must be closely related to the objective of the mission or operation being evaluated. For example, kills per shot, probability of kill, effective range, etc. Linkage shall exist among the various MOEs used in the Analysis of Alternatives (AoAs), Operations Requirements Document (ORD) and Test and Evaluation (T&E); in particular, the MOEs, Measures of Performance (MOPs), criteria in the ORD, the AoAs, the Test and Evaluation Master Plan (TEMP), and the Acquisition Program Baseline (APB) shall be consistent. A meaningful MOE must be quantifiable and a measure of what degree the real objective is achieved.

MEASURE OF PERFORMANCE (MOP) — Measure of a lower level of performance representing subsets of Measures of Effectiveness (MOEs). Examples are speed, payload, range, time on station, frequency, or other distinctly quantifiable performance features.

MILESTONE — The point when a recommendation is made and approval sought regarding starting or continuing (proceeding to next phase) an acquisition program.

MILESTONE DECISION AUTHORITY (MDA) — The individual designated in accordance with criteria established by the Under Secretary of Defense (Acquisition, Technology, and Logistics) (USD(AT&L)), or by the Assistant Secretary of Defense (Command, Control, Communications, and Intelligence) (ASD(C^3I)) for Automated Information System (AIS) acquisition programs (DoD 5000.2-R (Reference C)), to approve entry of an acquisition program into the next acquisition phase.

MILITARY OPERATIONAL REQUIREMENT — The formal expression of a military need, responses to which result in development or acquisition of item(s), equipment, or systems.

MISSION RELIABILITY — The probability that a system will perform mission essential functions for a given period of time under conditions stated in the mission profile.

MODEL — A model is a representation of an actual or conceptual system that involves 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 TEST AND EVALUATION — T&E conducted by two or more DoD Components for systems to be acquired by more than one DoD Component (Joint acquisition program), or for a DoD Component's systems that have interfaces with equipment of another DoD Component. May be developmental testing or operational testing (MOT&E).

NON-DEVELOPMENTAL ITEM (NDI) — A NDI is any previously developed item of supply used exclusively for government purposes by a federal agency, a state or local government, or a foreign government with which the United States has a mutual defense cooperation agreement; any item described above that requires only minor modifications or modifications of the type customarily available in the commercial marketplace in order to meet the requirements of the processing department or agency.

NONMAJOR DEFENSE ACQUISITION PROGRAM — A program other than a Major Defense Acquisition Program (MDAP) Acquisition Category (ACAT) I or a highly sensitive classified program: i.e., ACAT II and ACAT III programs.

NUCLEAR HARDNESS — A quantitative description of the resistance of a system or component to malfunction (temporary and permanent) and/or degraded performance induced by a nuclear weapon environment. Measured by resistance to physical quantities such as overpressure, peak velocities, energy absorbed, and electrical stress. Hardness is achieved through adhering to appropriate design specifications and is verified by one or more test and analysis techniques.

OBJECTIVE — The performance value that is desired by the user and which the Program Manager (PM) is attempting to obtain. The objective value represents an operationally meaningful, time-critical, and cost-effective increment above the performance threshold for each program parameter.

OPEN SYSTEMS — Acquisition of Weapons Systems—An integrated technical and business strategy that defines key interfaces for a system (or a piece of equipment under development) in accordance with those adopted by formal consensus bodies (recognized industry standards' bodies) as specifications and standards, or commonly accepted (de facto) standards (both company proprietary and non-proprietary) if they facilitate utilization of multiple suppliers.

OPERATIONAL ASSESSMENT (OA) — 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 OA 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 (OT). OA may be made at any time using technology demonstrators, prototypes, mock-ups, Engineering Development Models (EMDs), or simulations but will not substitute for the independent Operational Test and Evaluation (OT&E) necessary to support full production decisions.

OPERATIONAL AVAILABILITY (Ao) — The degree (expressed in terms of 1.0 or 100 percent as the highest) to which one can expect equipment or weapon systems to work properly when required. The equation is uptime over uptime plus downtime, expressed as Ao. It is the quantitative link between readiness objectives and supportability.

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 (NBCC) 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 REQUIREMENTS — User- or user representative-generated validated needs developed to address mission area deficiencies, evolving threats, emerging technologies, or weapon system cost improvements. Operational requirements form the foundation for weapon system unique specifications and contract requirements.

OPERATIONAL REQUIREMENTS DOCUMENT (ORD) — Documents the user's objectives and minimum acceptable requirements for operational performance of a proposed concept or system. Format is contained in the Chairman of the Joint Chiefs of Staff Instruction (CJCSI) 3170.01B, *Requirements Generation System (RGS)*, April 15, 2001.

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

OPERATIONAL TEST AND EVALUATION (OT&E) — The field test, under realistic 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 tests. (Title 10 U.S.C. 2399)

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 Operational Test and Evaluation (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, Headquarters, Department of the Army (HQDA) staff elements, and others as necessary.

PARAMETER — A determining factor or characteristic. Usually related to performance in developing a system.

PERFORMANCE — Those operational and support characteristics of the system that allow it to effectively and efficiently perform its assigned mission over time. The support characteristics of the system include both supportability aspects of the design and the support elements necessary for system operation.

PILOT PRODUCTION — Production line normally established during first production phase to test new manufacturing methods and procedures. Normally funded by Research, Development, Test, and Evaluation (RDT&E) until the line is proven.

POSTPRODUCTION TESTING — Testing conducted to assure 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 improvements planned for ongoing systems that go beyond the current performance envelope to achieve a needed operational capability.

PREPRODUCTION PROTOTYPE — An article in final form employing standard parts, representative of articles to be produced subsequently in a production line.

PREPRODUCTION QUALIFICATION TEST — The formal contractual tests that ensure design integrity over the specified operational and environmental range. These tests usually use prototype or pre-production hardware fabricated to the proposed production design specifications and drawings. Such tests include contractual Reliability and Maintainability (R&M) demonstration tests required prior to production release.

PROBABILITY OF KILL (Pk) — The lethality of a weapon system. Generally refers to armaments. (e.g., missiles, ordnance, etc.). Usually the statistical probability that the weapon will detonate close enough to the target with enough effectiveness to disable the target.

PRODUCT IMPROVEMENT (PI) — Effort to incorporate a configuration change 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 useful military life. Usually results from feedback from the users.

PRODUCTION ACCEPTANCE TEST AND EVALUATION (PAT&E) — T&E of production items to demonstrate that items procured fulfill requirements and specifications of the procuring contract or agreements.

PRODUCTION ARTICLE — The end item under initial or Full Rate Production (FRP).

PRODUCTION PROVE OUT — A technical test conducted prior to production testing with prototype hardware to determine the most appropriate design alternative. This testing may also provide data on safety, the achievability of critical system technical characteristics, refinement and ruggedization of hardware configurations, and determination of technical risks.

PRODUCTION QUALIFICATION TEST (PQT) — A technical test completed prior to the Full Rate Production (FRP) decision to ensure the effectiveness of the manufacturing process, equipment, and procedures. This testing also serves the purpose of providing data for the independent evaluation required for materiel release so that 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 online.

PROGRAM MANAGER (PM) — A military or civilian official who is responsible for managing, through Integrated Product Teams (IPTs), an acquisition program.

PROTOTYPE — An original or model on which a later system/item is formed or based. Early prototypes may be built during early design stages and tested prior to advancing to advanced engineering. Selected prototyping may evolve into an Engineering Development Model (EDM), as required to identify and resolve specific design and manufacturing risks in that phase or in support of Preplanned Product Improvement (P^3I) or Evolutionary Acquisition (EA).

QUALIFICATIONS TESTING — Simulates defined operational environmental conditions with a predetermined safety factor, the results indicating whether a given design can perform its function within the simulated operational environment of a system.

QUALITY ASSURANCE (QA) — A planned and systematic pattern of all actions necessary to provide confidence that adequate technical requirements are established, that products and services conform to established technical requirements, and that satisfactory performance is achieved.

REALISTIC TEST ENVIRONMENT — The conditions under which the 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 ability of a system and its parts to perform its mission without failure, degradation, or demand on the support system. (See Mean Time Between Failure (MTBF).)

REQUIRED OPERATIONAL CHARACTERISTICS — System parameters that are primary indicators of the system's capability to be employed to perform the required mission functions, and to be supported.

REQUIRED TECHNICAL CHARACTERISTICS — System parameters selected as primary indicators of achievement of engineering goals. These need not be direct measures of, but should always relate to the system's capability to perform the required mission functions, and to be supported.

RESEARCH — 1. Systematic inquiry into a subject in order to discover or revise facts, theories, etc., to investigate. 2. Means of developing new technology for potential use in defense systems.

RISK — A measure of the inability to achieve program objectives within defined cost and schedule constraints. Risk is associated with all aspects of the program, e.g., threat, technology, design processes, Work Breakdown Structure (WBS) elements, etc. It has two components:

1. The probability of failing to achieve a particular outcome; and
2. The consequences of failing to achieve that outcome.

RISK ASSESSMENT — The process of identifying program risks within risk areas and critical technical processes, analyzing them for their consequences and probabilities of occurrence, and prioritizing them for handling.

RISK MONITORING — A process that systematically tracks and evaluates the performance of risk items against established metrics throughout the acquisition process and develops further risk reduction handling options as appropriate.

SAFETY — The application of engineering and management principles, criteria, and techniques to optimize safety within the constraints of operational effectiveness, time, and cost throughout all phases of the system life cycle.

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 (DT) and user or Operational Test (OT) 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 (SAR) — Standard, comprehensive, summary status reports on Major Defense Acquisition Programs (MDAPs) (Acquisition Category (ACAT) I) required for periodic submission to the Congress. They include key cost, schedule, and technical information.

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 of evaluating various strategies for the operation of the system within the 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.

SIMULATOR — A generic term used to describe equipment used to represent weapon systems in development testing, operational testing, and training, e.g., a threat simulator has one or more characteristics which, when detected by human senses or man-made sensors, provide the appearance of an actual threat weapon system with a prescribed degree of fidelity.

SPECIFICATION — A document used in development and procurement that describes the technical requirements for items, materials, and services, including the procedures by which it will be determined that the requirements have been met. Specifications may be unique to a specific program (program-peculiar) or they may be common to several applications (general in nature).

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).

SURVIVABILITY — Survivability is the capability of a system and its crew to avoid or withstand a man-made hostile environment 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. Susceptibility is considered a subset of survivability.

SYSTEM — 1. The organization of hardware, software, material, facilities, personnel, data, and services needed to perform a designated function with specified results, such as the gathering of specified data, its processing, and delivery to users. 2. A combination of two or more interrelated equipment (sets) arranged in a functional package to perform an operational function or to satisfy a requirement.

SYSTEM ENGINEERING, DEFENSE — A comprehensive, iterative technical management process that includes translating operational requirements into configured systems, integrating the technical inputs of the entire design team, managing interfaces, characterizing and managing technical risk, transitioning technology from the technology base into program-specific efforts, and verifying that designs meet operational needs. It is a life cycle activity that demands a concurrent approach to both product and process development.

SYSTEMS ENGINEERING PROCESS (SEP) — A logical sequence of activities and decisions transforming an operational need into a description of system performance parameters and a preferred system configuration.

SYSTEM SPECIFICATION — States all necessary functional requirements of a system in terms of technical performance and mission requirements, including test provisions to assure that all requirements are achieved. Essential physical constraints are included. System specifications state the technical and mission requirements of the system as an entity.

SYSTEM THREAT ASSESSMENT (STA) — Describes the threat to be countered and the projected threat environment. The threat information must be validated by the Defense Intelligence Agency (DIA) programs reviewed by the Defense Acquisition Board (DAB).

TECHNICAL EVALUATION (TECHEVAL) — The study, investigations, or Test and Evaluation (T&E) by a developing agency to determine the technical suitability of materiel, equipment, or a system, for use in the military Services. (See Development Test and Evaluation (DT&E), Navy.)

TECHNICAL FEASIBILITY TEST — A technical feasibility test is a developmental test typically conducted during Concept Refinement (CR) and Technology Development (TD) to provide data to assist in determining safety and health hazards, and establishing system performance specifications and feasibility.

TECHNICAL PERFORMANCE MEASUREMENT (TPM) — Describes all the activities undertaken by the government to obtain design status beyond those treating schedule and cost. A TPM manager is defined as the product design assessment, which estimates through tests the values of essential performance parameters of the current design of Work Breakdown Structure (WBS) product elements. It forecasts the values to be achieved through the planned technical program effort, measures differences between achieved values and those allocated to the product element by the System Engineering Process (SEP), and determines the impact of those differences on system effectiveness.

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.

TEST — Any program or procedure that is designed to obtain, verify, or provide data for the evaluation of Research and Development (R&D) (other than laboratory experiments); progress in accomplishing development objectives; or performance and operational capability of systems, subsystems, components, and equipment items.

TEST AND EVALUATION (T&E) — Process by which a system or components provide information regarding risk and risk mitigation and empirical data to validate models and simulations. T&E permits, as assessment of the attainment of technical performance, specifications, and system maturity to determine whether systems are operationally effective, suitable, and survivable for intended use. There are two general types of T&E—Developmental (DT&E) and Operational (OT&E). (See Operational Test and Evaluation (OT&E), Initial Operational Test and Evaluation (IOT&E), and Developmental Test and Evaluation (DT&E).)

TEST AND EVALUATION MASTER PLAN (TEMP) — Documents the overall structure and objectives of the Test and Evaluation (T&E) program. It provides a framework within which to generate detailed T&E plans, and it documents schedule and resource implications associated with the T&E program. The TEMP identifies the necessary Developmental Test and Evaluation (DT&E), Operational Test and Evaluation (OT&E), and Live Fire Test and Evaluation (LFT&E) activities. It relates program schedule, test management strategy and structure, and required resources to: Critical Operational Issues (COIs); critical technical parameters; objectives and thresholds documented in the Operational Requirements Document (ORD); evaluation criteria; and milestone decision points. For multi-Service or joint programs, a single integrated TEMP is required. Component-unique content requirements, particularly evaluation criteria associated with COIs, can be addressed in a component-prepared annex to the basic TEMP. (See Capstone TEMP).)

TEST BED — A system representation consisting of actual hardware and/or software and computer models or prototype hardware and/or software

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 (ADPE), 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. Audio-visual 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 Temporary Duty (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 strengths, capabilities, and strategic objectives of any adversary that can limit or negate U.S. mission accomplishment or reduce force, system, or equipment effectiveness.

THRESHOLDS — The Minimum Acceptable Value (MAV) which, in the user's judgment, is necessary to satisfy the need. If threshold values are not achieved, program performance is seriously degraded, the program may be too costly, or the program may no longer be viable.

TRANSPORTABILITY — The capability of materiel to be moved by towing, self-propulsion, or carrier through any means, such as railways, highways, waterways, pipelines, oceans, and airways. (Full consideration of available and projected transportation assets, mobility plans, and schedules, and the impact of system equipment and support items on the strategic mobility of operating military forces is required to achieve this capability.)

UNKNOWN-UNKNOWNS (UNK(s)) — Future situation impossible to plan, predict, or even know what to look for.

USER — An operational command or agency that receives or will receive benefit from the acquired system. Combatant Commanders (COCOMs) and the Services are the users. There may be more than one user for a system. The Services are seen as users for systems required to organize, equip, and train forces for the COCOMs of the unified command.

USER FRIENDLY — Primarily a term used in Automated Data Processing (ADP), it connotes a machine (hardware) or program (software) that is compatible with a person's ability to operate each successfully and easily.

USER REPRESENTATIVES — A command or agency that has been formally designated by proper authority to represent single or multiple users in the requirements and acquisition process. The Services and the Service components of the Combatant Commanders (COCOMs) are normally the user representatives. There should be only one user representative for a system.

VALIDATION — 1. The process by which the contractor (or as otherwise directed by the DoD component procuring activity) tests a publication/Technical Manual (TM) for technical accuracy and adequacy. 2. The procedure of comparing input and output against an edited file and evaluating the result of the comparison by means of a decision table established as a standard.

VARIANCE (Statistical) — A measure of the degree of spread among a set of values; a measure of the tendency of individual values to vary from the mean value. It is computed by subtracting the mean value from each value, squaring each of these differences, summing these results, and dividing this sum by the number of values in order to obtain the arithmetic mean of these squares.

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 (manmade) hostile environment. Vulnerability is considered a subset of survivability.

WORK BREAKDOWN STRUCTURE (WBS) — An organized method to break down a project into logical subdivisions or subprojects at lower and lower levels of details. It is very useful in organizing a project.

WORKING-LEVEL INTEGRATED PRODUCT TEAM (WIPT) — Team of representatives from all appropriate functional disciplines working together to build successful and balanced programs, identify and resolve issues, and make sound and timely decisions. WIPTs may include members from both government and industry, including program contractors and sub-contractors. A committee, which includes non-government representatives to provide an industry view, would be an advisory committee covered by the Federal Advisory Committee Act (FACA) and must follow the procedures of that Act.

APPENDIX C
TEST-RELATED DATA ITEM DESCRIPTIONS

extracted from DoD 5010.12-L,
Acquisition Management System and
Data Requirements 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 (Non-standard) 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	
Development Test and Evaluation Plan	DI-MCCR-80309
Integration and Test Plan	DI-MCCR-80308
Sound Test Failure Notification and Recommendation	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
Structure-borne Vibration Acceleration 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

APPENDIX D
BIBLIOGRAPHY

Acquisition Desk Book, http://akss.dau.mil/DAG.

AFLC Pamphlet 800-34, *ACQ Logistics Management*.

AFM [Air Force Manual] 55-43, *Management of Operational Test and Evaluation*.

AFM 99-113, *Space Systems Test and Evaluation Process,* May 1, 1996.

AFOTEC Regulation 23-1, *Organization and Functions Air Force Operational Test and Evaluation Center (AFOTEC)*.

AFOTEC, *Electronic Warfare Operational Test and Evaluation*.

AFOTEC, *Nuclear Survivability Handbook for Air Force OT&E*.

AFOTEC 99-103, *Capabilities Test and Evaluation*, October 22, 2001.

AFR [Air Force Regulation] 23-36, *Air Force Operational Test and Evaluation Center (AFOTEC)*.

AFR 55-30, *Operations Security*.

AFR 55-43, *Management of Operational Test and Evaluation,* June 1979.

AFR 800-2, *Acquisition Program Management*, July 21, 1994.

AFR 800-8, *Integrated Logistics Support (ILS) Program*.

AFR 80-14, *Test and Evaluation*. **(Cancelled)**

AFR 80-38, *Management of the Air Force Systems Survivability Program*, May 27, 1997.

AFR 800-2, *Acquisition Program Management*, July 21, 1994.

AFSCR [Air Force System Command Regulation] 550-13, *Producibility*.

AFSCR 550-15, *Statistical Process Control*.

AMC [Army Materiel Command] Plan 70-2, *AMC-TRADOC Materiel Acquisition Handbook*.

Anderson, J.E., "Towards the Goal of Improving MIL-STD Testing," *Defense Systems Management Journal*, Vol. 12, No. 2, pp. 30-34, April 1976.

AR [Army Regulation] 10-11, *U.S. Army Materiel Development and Readiness Command*.

AR 10-4, *U.S. Army Operational Test and Evaluation Agency*. **(Cancelled)**

AR 10-41, *U.S. Army Training and Doctrine Command*.

AR 10-42, *U.S. Armed Forces Command*.

AR 700-127, *Integrated Logistics Support*, November 10, 1999.

AR 70-24, *Special Procedures Pertaining to Nuclear Weapons System Development*.

AR 70-60, *Nuclear Survivability of Army Materiel*.

AR 70-71, *Nuclear, Biological, and Chemical Contamination Survivability of Army Materiel*.

AR 71-1, *Systems Acquisition Policy and Procedures* (replaced by AR 70-1), December 31, 2003.

AR 71-3, *Force Development User Testing*.

AR 73-1, *Test and Evaluation Policy*, January 7, 2002.

ATEC PAM [Army Test and Evaluation Command Pamphlet] 310-4, *Index of Test Operations Procedures and International Test Operations Procedures*, December 31, 1988.

BDM Corporation, *Application of Simulations to the OT&E Process*, May 1974.

BDM Corporation, *Functional Description of the Acquisition Test and Evaluation Process*, July 8, 1983.

Bolino, John V., *Software T&E in the Department of Defense*.

CJCSI [Chairman of the Joint Chiefs of Staff Instruction] 3170.013, *Joint Capabilities Integration and Development System* (JCIDS), March 12, 2004.

CJCSM [Chairman of the Joint Chiefs of Staff Manual] 3170.01A, *Operation of the Joint Capabilities Integration and Development System*, March 12, 2004.

COMOPTEVFOR [Commander, Operational Test and Evaluation Force], AD Number AD124168, *Operational Test Director Guide.*

COMOPTEVFOR, *Project Analysis Guide*.

COMOPTEVFORINST 3960.1D, *Operational Test Director's Guide*.

DA PAM [Department of the Army Pamphlet] 700-50, *Integrated Logistics Support: Developmental Supportability Test and Evaluation Guide*. **(Cancelled)**

DA PAM 73-1, *Test and Evaluation in Support of Systems Acquisition*, May 30, 2003.

Data Link Vulnerability Analysis Joint Test Force, *DVAL Methodology – Methodology Overview/ Executive Summary,* November 1984.

Defense Acquisition University Press. *The Glossary of Defense Acquisition Acronyms and Terms,* Vol. 11, September 2003.

Defense Science Board Task Force Report, *Solving the Risk Equation in Transitioning from Development to Production*, May 25, 1983 (later published as DoD Manual 4245.7).

Defense Science Board, *Report of Task Force on Test and Evaluation*, April 1, 1974.

Defense Science Board Report, *Test and Evaluation Capabilities*, December 2000.

Defense Systems Management College, *Acquisition Logistics Guide,* December 1997.

Defense Systems Management College, *Joint Logistics Commanders Guidance for the Use of an Evolutionary Acquisition (EA) Strategy in Acquiring Command and Control (C^2) Systems*, March 1987.

Defense Systems Management College, "Risk Management as a Means of Direction and Control," *Program Manager's Notebook*, Fact Sheet Number 4.5, June 1992.

Defense Systems Management College, Defense Manufacturing Management Course.

Defense Systems Management College, *Integrated Logistics Support Guide*, First Edition, May 1986.

Defense Systems Management College, *Joint Logistics Commanders Guide for the Management of Joint Service Programs*, 1987.

Defense Systems Management College, *Joint Logistics Commanders Guide for the Management of Multinational Programs,* 1987.

Defense Systems Management College, *Systems Engineering Fundamentals Guide*, January 2001.

Defense Systems Management College, *Systems Engineering Management Guide*, Second Edition, 1990.

Director, Defense Test and Evaluation, Memorandum of Response of Admiral Isham Linder, Director, Test and Evaluation to Supplemental Questions posed by the Senate Governmental Affairs Committee, August 2, 1983.

DoD 3200.11-D, *Major Range and Test Facility Base Summary of Capabilities,* June 1, 1983.

DoD 3235.1-H, *Test and Evaluation of System Reliability, Availability, and Maintainability: A Primer*, March 1982.

DoD 4245.7-M, *Transition from Development to Production,* September 1985.

DoD 5010.12-L, *Acquisition Management Systems and Data Requirements Control List (AMSDL)*, April 2001.

DoD 5000.2-R, *Mandatory Procedures for Major Defense Acquisition Programs (MDAPs) and Major Automated Information System (MAIS) Acquisition Programs*, June 2001.

DoD 5000.3-M-1, *Test and Evaluation Master Plan (TEMP) Guideline*s, October 1986. **(Cancelled)**

DoD 5000.3-M-2, *Foreign Weapons Evaluation Program and NATO Comparative Test Program Procedures Manual,* January 1994.

DoD 5000.3-M-3, *Software Test and Evaluation Manual*. **(Cancelled)**

DoDD [Department of Defense Directive] 3200.11, *Major Range and Test Facility Base (MRTFB)*, May 1, 2002.

DoDD 4245.6, *Defense Production Management*. **(Cancelled)**

DoDD 4245.7, *Transition from Development to Production,* with Change 1, September 1985. **(Cancelled)**

DoDD 5000.1, *The Defense Acquisition System,* May 12, 2003.

DoDD 5000.3, *Test and Evaluation*, December 26, 1979. **(Cancelled)**

DoDD 5000.37, *Acquisition and Distribution of Commercial Products (ADCOP)*, September 29, 1978. **(Cancelled)**

DoDD 5000.59, *DoD Modeling and Simulation Management*, January 20, 1998.

DoDD 5000.38, *Production Readiness Reviews*. **(Cancelled)**

DoDD 5000.39, *Acquisition and Management of Integrated Logistics Support for Systems and Equipment*, November 17, 1983. **(Cancelled)**

DoDD 5010.8, *Value Engineering Program, 1968.*

DoDD 5010.41, *Joint Test and Evaluation (JT&E) Program*, February 23, 1998.

DoDD 5105.31, *Defense Nuclear Agency*, June 14, 1985. **(Cancelled)**

DoDD 5134.1, *Under Secretary of Defense for Acquisition, Technology, and Logistics (USD(AT&L))*, April 21, 2000.

DoDD 5141.2, *Director of Operational Test and Evaluation (DOT&E)*, May 25, 2000.

DoDD 5160.5, *Responsibilities for Research, Development, and Acquisition of Chemical Weapons and Chemical and Biological Defense*, May 1, 1985.

DoDI [Department of Defense Instruction] 4245.4, *Acquisition of Nuclear-Survivable Systems*, July 25, 1988. **(Cancelled)**

DoDI 5000.2, *Operation of the Defense Acquisition System*, May 12, 2003.

DoDI 5000.61, *DoD Modeling and Simulation (M&S) Verification, Validation and Accreditation (VV&A)*, May 13, 2003.

DoDI 5030.55, *Joint AEC-DoD Nuclear Weapons Development Procedures*, January 4, 1974.

DoDI 7000.10, *Contractor Cost Performance, Funds Status, and Cost/Schedule Reports*. **(Cancelled)**

DoD-STD [Department of Defense Standard]-2167, *Defense System Software Development,* January 1, 1987.

DoD-STD-2168, *Defense System Software Quality Program,* May 21, 1996.

DoD FMR 7000.14-R, Vol. 02B, *Budget Formulation and Presentation*, Chapter 5, "Research, Development, Test and Evaluation Appropriations," June 2004.

FY88 Report of the Secretary of Defense to the Congress, 1988.

GAO/NSIAD-00-199, *Best Practices: A More Constructive Test Approach is Key to Better Weapon System Outcome,* July 2000.

GAO/NSIAD-87-57, *Operational Test and Evaluation Can Contribute More to Decision-Making,* February 4, 1987.

GAO/PEMD-86-12BR, *Bigeye Bomb: An Evaluation of DoD's Chemical and Developmental Tests*, June 10, 1986.

GAO/PEMD-87-17, *Live Fire Testing: Evaluating DoD's Programs,* June 12, 1997.

GAO/PSAD-78-102, *Operational Testing of Air Force Systems Requires Several Improvements*.

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,* June 25, 1979.

Georgia Institute of Technology, *The Software Test and Evaluation Project*.

Hearing before the Senate Committee on Governmental Affairs, June 23, 1983.

Hoivik, Thomas H., *The Navy Test and Evaluation Process in Major Systems Acquisition*, DTIC Accession No. AD A035935, November 1976.

International Test and Evaluation Association (ITEA), *Test and Evaluation of Command and Control Systems-Developed Evolutionary Acquisition*, May 1987.

JCHEM Joint Test Force, *Concept and Approach for a Joint Test and Evaluation of U.S. Chemical Warfare-Retaliatory Capabilities*, AD #B088340, February 1984.

Johnson, Larry H., *Contractor Testing and the Army Test and Evaluation Master Plan for Full Engineering Development*, Defense Systems Management College, 1983.

Joint Test Director's Handbook (Draft), military handbook.

Kausal, Tony, *Comparison of the Defense Acquisition Systems of France, Great Britain, Germany, and the United States*, Defense Acquisition University: Ft. Belvoir, VA, September 1999.

Kausal, Tony, and Stefan Markowski, *Comparison of the Defense Acquisition Systems of Australia, Japan, South Korea, Singapore, and the United States*, Defense Acquisition University: Ft. Belvoir, VA, July 2000.

Mann, Greg A., LTC, *The Role of Simulation in Operational Test and Evaluation*, Report AU-ARI-83-10, Air University Press, Maxwell Air Force Base, AL: August 1983.

MCO [Marine Corps Order] 3960.2, *Marine Corps Operational Test and Evaluation Activity (MCOTEA)*; establishment of, with changes, October 14, 1994.

MCO 5000.11B, *Testing and Evaluation of Systems and Equipment for the Marine Corps*.

Memorandum of Agreement of Multi-Service OT&E, with changes, August 2003.

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.

MIL-HDBK[Military Handbook]-881, *Work Breakdown Structure*, January 2, 1998.

MIL-STD[Military Standard]-1388-1A, *Logistics Support Analysis*, April 11, 1983.

MIL-STD-1512A, *Technical Reviews and Audits for Systems, Equipment, and Computer Programs*, September 30, 1997.

MIL-STD-1528, *Manufacturing Master Plan*.

MIL-STD-480, *Configuration Control – Engineering Changes, Deviations, and Waivers*, July 15, 1988.

MIL-STD-483 (USAF), *Configuration Management Practices for Systems, Equipment, Munitions and Computer Software*.

MIL-STD-490, *Specification Practices*, June 4, 1985, modified March 1997.

MIL-STD-499A (USAF), *Engineering Management Practices*, May 1, 1974.

MIL-STD-961D, *DoD Standard Practice for Defense Specifications*, March 22, 1995.

MITRE Corporation, *Software Reporting Metrics*, November 1985.

NAVMATINST [Naval Material Command Instruction] 3960, *Test and Evaluation*.

NAVSO-P-2457, *RDT&E/Acquisition Management Guide*, 10th Edition.

O'Bryon, James F., Assistant Deputy Under Secretary of Defense (Live Fire Test). Statement before the Acquisition Policy Panel of the House Armed Services Committee, September 10, 1987.

Office of the Deputy Director, Test and Evaluation (Live Fire Test), *Live Fire Test and Evaluation Planning Guide*, June 1989.

Office of the Deputy Under Secretary for Research and Engineering (Test and Evaluation), *Joint Test and Evaluation Procedures Manual* (Draft), October 1986.

Office of the Under Secretary of Defense (Acquisition, Technology and Logistics), Director, Strategic & Tactical Systems, *The Foreign Comparative Testing (FCT) Handbook,* March 2001, http://www.acq.osd.mil/cto.

Office of the Under Secretary of Defense for Research and Engineering (Director, Defense Test and Evaluation), *Test and Evaluation in the United States Department of Defense*, August 1983.

OPNAVINST [Office of the Chief of Naval Operations Instruction] 3070.1A, *Operations Security*.

OPNAVINST 3960.10C, *Test and Evaluation* (Draft). **(Cancelled)**

OPNAVINST 5000.49, *ILS in the Acquisition Process*.

OPNAVINST 5440.47F, *Mission and Functions of Operational Test and Evaluation Force (OPTEVFOR)*.

Public Law (P.L.) 92-156, *National Defense Authorization Act for Fiscal Year 1972.*

Public Law 98-94, *National Defense Authorization Act for Fiscal Year 1987*.

Public Law 99-661, N*ational Defense Authorization Act for Fiscal Year 1987.*

Range Commanders Council (RCC), *Universal Documentation System Handbook*, Document 501-84, Vol. 3, August 1989.

Rhoads, D.I., *Systems Engineering Guide for Management of Software Acquisition*, Defense Systems Management College, 1983, pp. IV-1 to IV-23.

Rome Air Development Center, *Standard Procedures for Air Force Operational Test and Evaluation*.

SECNAVINST [Secretary of the Navy Instruction] 5000.1B, *Systems Acquisition*.

SECNAVINST 5000.39, *USN Acquisition and Management of ILS Systems and Equipment*.

SECNAVINST 5000.42C, *RDT&E Acquisition Procedures*.

SECNAVINST 5430.67A, *Assignment of Responsibilities for Research, Development, Test and Evaluation*.

Stevens, Roger T., *Operational Test and Evaluation Process: A Systems Engineering Process*, John Wiley & Sons, New York: 1979, pp. 49-56.

Thomas, E.F., *Reliability Testing Pitfalls*, Reliability and Maintainability Symposium, Vol. 7, No. 2, 1974, pp. 78-85.

USAOTEA Operational Test and Evaluation Guide, Special Supplement, Continuous Comprehensive Evaluation Planning for Nuclear Hardness and Survivability Test and Evaluation (Draft). **(Cancelled)**

Watt, Charles K., *The Role of Test Beds in* C^3I, AFCEA 35th Annual Convention, June 15-17, 1982.

Willoughby, Willis J., "Reliability by Design, Not Chance," *Defense Systems Management Journal*, Vol. 12, No. 2, pp. 12-18.

APPENDIX E
INDEX

APPENDIX F
POINTS OF CONTACT
SERVICE TEST AND EVALUATION COURSES

ARMY

Commander, U.S. Army Developmental Test Command
ATTN: CSTE-DTC
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 Test Command
ATTN: CSTE-TC
4501 Ford Avenue
Alexandria, VA 22302

NAVY

Commander, Space & Naval Warfare Systems Command
ATTN: Management Operations
National Center, Building 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

CUSTOMER FEEDBACK

Dear Valued Customer:

The Defense Acquisition University Press publishes this reference tool for students of the Defense Acquisition University and other members of the defense acquisition workforce. We continually seek customer inputs to maintain the integrity and usefulness of this publication.

We ask that you spend a few minutes evaluating this publication. Although we have identified several topic areas to evaluate, please feel free to add any additional thoughts as they apply to your particular situation. This information will assist us in improving future editions.

To assist you in responding, this pre-addressed form requires no postage. Simply tear-out this page, fold, and tape closed as indicated, then place in the mail. Responses may also be faxed to the DAU Press at (703) 805-2917 or DSN 655-2917.

TEST AND EVALUATION MANAGEMENT GUIDE (January 2005)

1. Do you find this publication a useful reference tool?
YES____
NO ____ How can we improve the level of usefulness?
____
____

2. Is the material presented in this publication current?
YES____ NO_____

3. Using the rating scale below, please indicate your satisfaction level for each topic area. Place your response on the line beside the topic area.

1 = Very satisfied 2 = Satisfied 3 = No opinion 4 = Dissatisfied 5 = Very dissatisfied

_____Readability
_____Scope of coverage
_____Contribution to your knowledge of the subject
_____Contribution to your job effectiveness
_____Contribution to process improvement

4. Are there other subject areas you would like to have included in this publication?
YES___ (Please identify)____
NO ___

5. If you recommend additions, deletions, or changes to this publication, please identify them below.
____
____
____

6. Additional comments
____
____
____

(THIS SECTION IS OPTIONAL)

NAME/TITLE____
ADDRESS____
CITY/STATE/ZIP____
Commercial (____)______________ DSN______________E-Mail____________________________