

TEST AND EVALUATION OVERVIEW

1. Introduction. During the development of a weapon system, subcontractors, prime contractors and Army agencies perform various tests. In the early stages of development, these tests evaluate design approaches and select design solutions for further development. As the design matures in the life-cycle, the tests become more complex, attempting to provide confidence that the weapon system will satisfactorily perform all intended functions during combat. This period of instruction will involve a discussion of the procedures used to plan, conduct, evaluate and report on test and evaluation (T&E) of materiel systems.

2. Objectives. At the conclusion of this unit of instruction the student should:

- a. Identify and explain why materiel is tested.
- b. Differentiate between the two categories of test and evaluation and describe the major tests conducted in each category.
- c. Know the DoD and Army test and evaluation responsibilities.
- d. Explain how test and evaluation supports the Life Cycle Model.

3. Scope. The Test and evaluation process is a requirement and a tool to help reduce risk. Army testing encompasses a wide variety of tasks, ranging from closely controlled developmental laboratory bench tests through free play operational force-on-force tactical tests. These tests try to answer questions that cannot be answered by other sources such as inspections, studies, simulations, and wargames. The extent of testing depends on the questions to be answered (critical test issues), the measure of satisfactory performance (test criteria), and if the information collected is enough to reduce the risk to a level acceptable to those responsible for the application. But, how much testing is enough? The acquisition community determines how much to test, depending on resource and schedule constraints, environmental and safety concerns, a valid need to deploy new technology, and political and social considerations. The issues to be tested determine the purpose and objectives of each test, which in turn determine the type of test needed. The Army conducts test for several reasons:

- a. To support decision making/risk assessment during the materiel acquisition process.
- b. To verify attainment of technical performance specifications and functional objectives.
- c. To verify a system is operationally effective and suitable for its intended use.
- d. To incorporate solutions for previously identified problems.
- e. Ensure that critical issues to be resolved have been adequately addressed.

- f. To provide sufficient data for modeling and simulation verification.
- g. To verify production.
- h. To meet Congressional requirements.

4. Test and Evaluation Policies.

a. DOD 5000.2-R and AR 73-1, Test and Evaluation Policy, broadly outlines T&E policy. The purpose of T&E is to make a direct contribution to the timely development, production, and fielding of systems that meet the user's need and that are effective, suitable and safe.

b. **Testing demonstrates a system's technical capabilities and its operational effectiveness** (including lethality, survivability, and vulnerability) **and suitability** (including compatibility, inter-operability, reliability and maintainability (R&M), logistic supportability, safety, health, human factors, and trainability). By conducting appropriate T&E, the decision authority and program manager will have critical input for decision reviews that commit significant additional resources to a program. The basis for a decision from the executive making decisions depends on the strengths, the weaknesses, and the amount of acceptable risk. Thus, T&E provides a solid base of information to the decision maker for risk analysis.

c. Test planning begins in the Concept Exploration Phase and T&E consideration continues throughout the system's life cycle. Testers and evaluators/assessors must be involved early to ensure the test program for the most promising alternative materiel solution can support the acquisition strategy or plan. Once a materiel solution alternative has been selected, the user defines Critical System Characteristics (CSC), Critical Operational Issues (COIs), and Key Performance Parameters (KPPs) in the ORD to help the T&E community design evaluations and tests. These CSCs, COIs, and KPPs provide the basis for the Test and Evaluation Master Plan (TEMP). The TEMP is the most important document for test and evaluation planning and resourcing. The TEMP's focus is on the overall structure, major elements, and objectives of the test program consistent with the acquisition strategy. It includes sufficient detail to ensure timely availability of both existing and planned test resources required to support the T&E program. All Army acquisition programs require a TEMP.

d. Tests provide quantitative information and minimize the need for subjective interpretation of system performance. Once the independent evaluator develops the test plan, the testers determine the number of test articles needed, any requirements for special range instrumentation or measurements, the need for targets and simulators or surrogates or threat weapons or devices, and any other test resources.

e. Materiel developers or combat developers structure T&E programs concurrently with the acquisition strategy (AS). Tailored T&E takes advantage of previous evaluations and data from other available data sources. Modeling and simulation may provide early predicted outcomes. Structured test programs provide sufficient data and information to allow evaluation of issues regarding technical requirements, operational effectiveness and suitability. The T&E strategy provides for a determination of performance, and safety and health hazard limi-

tations prior to use by soldiers in test. In addition, the commander of the users must be made aware of and agree to the safety and health hazards that may harm soldiers. Combined as well as Concurrent developmental and operational testing is encouraged, when appropriate. An integrated developmental and operational testing approach will be used to provide the maximum benefits from a complete, unified T&E program by shortening the acquisition time and by using resources efficiently. Combined DT/OT and Concurrent DT and OT are examples of tailoring of the acquisition process. However, evaluation efforts will remain separate.

(1) Combined DT/OT. This combined testing is conducted simultaneously on the same hardware and software. This is encouraged if the independence of evaluations will not be compromised and the integrity of the results is not jeopardized.

(2) Concurrent DT/OT. This testing approach will normally be conducted on separate prototypes but overlapped in time. However, before operational testing, sufficient developmental testing must be done to ensure system readiness and to support a safety release for operational testing. Concurrent testing is authorized if the integrity of results is not jeopardized.

f. T&E sufficient to address all critical issues must be done before each major decision point to reduce acquisition risks and to estimate the capability of the system underdevelopment to meet technical and operational requirements. The independent evaluators, through the Continuous Comprehensive Evaluation (C²E) process, advises the system decision makers of any testing that does not adequately address critical issues and associated acquisition risks.

g. Tests are not repeated if satisfactory results are achieved. However, if test results reflect significant deficiencies in the system support package (SSP), the program will not move into a succeeding phase until all deficiencies have been corrected. The Test Incident Report (TIR) describes the minimum essential data for test incidents. Before testing, members of the acquisition team, including the tester, agree upon Failure Definition and Scoring Criteria (FDSC) IAW DA Pam 73-1. Critical and major TIRs require the MD to prepare a Corrective Action Report (CAR). A corrective action review team reviews all corrective action data to verify if they are appropriate and effective. The review team classifies the CAR either as open, pending, verified, closed, or not required (to close). The tester is NOT a member of the review team, only an advisor. If necessary, the corrections might need to be verified by re-test. A deficiency is considered significant if it makes the system unacceptable for deployment; or involves corrections requiring more than the most routine engineering.

h. Modeling and simulation will be considered to support the developmental and operational T&E of software and systems as they proceed throughout the life cycle. Models and simulations used for T&E must be accredited and validated prior to their use for extrapolation or predicting system performance.

5. Test and Evaluation Types. Test and evaluation has two primary categories: Developmental Test and Evaluation (DT&E) and Operational Test and Evaluation (OT&E). DT&E and OT&E should be coordinated so each test cycle precludes unnecessary duplica-

tion, requires minimum resources, and yields maximum data to satisfy the common needs of the materiel developer and operational testers. Developmental Tests (DT) and Operational Tests (OT) should be combined where cost/time benefits result.

a. Developmental Test and Evaluation (DT&E). AR 73-1 and DA Pam 73-4 cover Army Developmental Testing and Evaluation. DT is a generic term that encompasses engineering-type testing conducted in laboratories, contractor facilities, proving ground or developer testbed facilities, or field environments using engineers, technicians, or specially trained soldiers. The purposes of DT&E is to identify potential limitations of alternate designs, support identification and description of design risks and cost-performance trade-offs, support decisions to certify system for production, plus verify contractual requirements. The Army uses DT&E to verify that design risks are minimized, substantiate achievement of contractor technical performance, and certify readiness for operational test and evaluation. DT&E conducted throughout the acquisition process assists in engineering design and development of a system and to verify that technical performance specifications have been met. DT's structure helps estimate the outer limit of the system's operational envelope. The emphasis in DT is on achieving repeatability of test results, thereby requiring strictly controlled test conditions. The materiel developer is currently responsible for DT&E. Two agencies of the U.S. Army Materiel Command currently perform a majority of the DT&E for Army systems. The **Army Test and Evaluation Command (ATEC)** performs **Developmental Testing** for programs in all acquisition categories (ACAT I-IV). TECOM operates a variety of extensive proving grounds, missile ranges, and environmental test centers for the purpose of performing DT. Developmental modeling and simulation evaluations are performed by the Army Materiel Systems Analysis Activity (AMSAA). **Developmental evaluations** are currently performed by the Evaluation Analysis Center (EAC). Premilestone III testing for materiel systems ranges from mission needs determination to the milestone III production decision and covers funding categories 6.1 through 6.4. The categories of ATEC Premilestone III testing are:

(1) Technical Feasibility Test (TFT). A TFT is conducted during concept exploration, post MS 0, pre-MS I or MS I/II to assist in determining safety and establishing system performance specifications and determining feasibility of alternative concepts. Testing during Phase 0 (Concept Exploration & Definition) identifies and reduces risks in subsequent acquisition phases. The TFT provides data for independent developmental evaluation/assessment that supports the MS I or MS I/II decision. The program funding category is 6.3.

(a) While not tied to specific acquisition programs, the following Technology Base demonstrations may be conducted by the Government developmental tester. Technology demonstrations are conducted to assess the military utility or cost reduction potential of innovative Government or commercially developed technology.

1 Advanced Technology Demonstrations (ATDs) are used to expedite technology transition from the laboratory to operational use. They are demonstrations conducted in an operational environment and are primarily funded with 6.3 funds. They may integrate advanced technologies to establish the feasibility of a concept or may utilize prototypes, surrogates, and simulations to show that existing technology can support a concept. ATDs should include provisions for early testability and operational assessments.

2 Proof of Principle (POP) demonstrates, in a nonoperational environment, innovative technologies that will support system upgrades or provide new operational capabilities. POPs are technical demonstrations and troop experimentations conducted with brassboard configurations, subsystems, or surrogate systems.

(2) Engineering Development Test (EDT). An EDT is a DT conducted during Program Definition & Risk Reduction (Phase I) after MS I and pre-MS II in factory, laboratory, and proving ground environments. EDT provides data on safety, the achievability of critical system technical parameters, refinement and ruggedization of hardware configurations, and determination of technical risks. It includes the testing of compatibility and interoperability with existing or planned equipment and systems and the system effects caused by natural and induced environmental conditions. The EDT provides data for independent developmental evaluation/assessment that supports the MS II decision. The program funding category is 6.3.

(3) Production Prove Out Test (PPT). A PPT is a DT conducted post-MS II or post MS I/II prior to production testing, and provides data for the selection of the most appropriate source or design. When MS I and MS II are combined, PPT may also provide data on safety, the achievability of critical system technical parameters, refinement and ruggedization of hardware and software configurations, and determination of technical risks. Program funding categories are 6.4 and 6.5.

(4) Production Qualification Test (PQT). A PQT is a system level DT conducted post MS II or MS I/II (usually just prior to MS III), that demonstrates design integrity over the specified operational and environmental range. This test usually uses prototype or pre-production hardware and software fabricated to the proposed production design specifications and drawings. Such tests include contractual reliability and maintainability demonstration tests required prior to production release. The PQT provides data for independent developmental evaluation/assessment which supports the MS III decision. The program funding category is 6.4.

(a) PQT objectives are to obtain Government confirmation that the design is stable, logistically supportable, capable of being produced efficiently, and will meet the user requirements by assessing the performance envelope, to determine the adequacy of any corrective action indicated by previous tests, and to validate contractor quality control.

(b) PQT may include tests which are not included in the Tech Data Package or contract (i.e., environmental extremes, test-to-failure, etc) when such tests are necessary to obtain engineering data to verify corrective action or other purposes.

(5) **Live Fire Test (LFT).** Live Fire Tests are DTs that demonstrate the ability of the system to provide battle resilient survivability, or the munitions to provide lethality (see DA Pam 73-6). Vulnerability testing must address crew, hardware, and system (crew and hardware) vulnerability for threats and impact conditions that the system is designed to protect against and could encounter on the battlefield. Lethality testing must address lethality against the threat systems by firing the munitions or missile at appropriate targets configured for combat or targets comparable with such targets. For programs designated for Live Fire Test and Evaluation (LFT&E), legislation requires that LFT&E be conducted and reported before the program can proceed beyond Low Rate Initial Production. LFT&E uses either 6.4 or procurement funding.

(6) **Logistic Demonstration (LD).** An LD is a system-level test to provide data for the evaluation of the supportability of the materiel design and the system support package. It includes a nondestructive disassembly and reassembly of equipment, conducted on a dedicated engineering prototype or limited production item prior to MS III. The LD evaluates the achievement of maintainability goals, the adequacy and sustainability of tools, built-in-test equipment, selected test program sets, technical publications, maintenance instructions, trouble-shooting procedures, personnel skill requirements, the selection and allocation of spares and repair parts, tools, test equipment, and tasks to appropriate maintenance levels, and the adequacy of maintenance time standards. The LD is ideally conducted at least 6 months prior to the IOT&E to allow time to make corrections, if required. It is often convenient to conduct an LD in conjunction with the PQT. The LD may use selected analysis, evaluations, demonstrations, and testing tailored to each acquisition program to demonstrate adequacy of the proposed support concept and programmed support resources. Program funding is 6.4.

Pre Milestone III Information Systems and Software Testing.

Premilestone III testing for information systems ranges from mission needs determination to the milestone III production decision but is typically conducted during the development phase. These tests include the **Software Development Tests (SDT), Software Qualification Tests (SQT), and C3I Interoperability Certification Tests. This material is covered in a separate Unit of Instruction, Software Acquisition Management (ALM 73-7787).** A brief description of these three tests are listed below.

(1) **C3I Interoperability Certification Test.** This test applies to C3I systems having interfaces or interoperability requirements with other systems. This test may consist of simple demonstrations using message analysis or parsing software with limited interface connectivity, or extend to full-scale scenario-driven exercises with all interfaces connected. The U.S. Army Communications-Electronics Command (CECOM) serves as the Army Participating Test Unit (APTU), which supports these types of tests conducted by the Defense Information Systems Agency (DISA) for system certification and recertification. (See Joint Interoperability and Engineering Organization/Joint Interoperability Test Command (JIEO/JITC) Circular 9002, Requirements Assessment and Interoperability Certification of C4I and Automated Systems (AIS) Equipment and Systems.)

(2) **Software Development Test.** The SDT consists of program or module and cycle or system levels of testing. The software developer conducts the SDT, and the independent quality control element of the software development center also participates.

(3) **Software Qualification Test.** This is a system test conducted by the developmental tester using live-data files supplemented with user prepared data and executed on target hardware. The objectives are to obtain Government confirmation that the design will meet performance and operational requirements and to determine the adequacy and timeliness of any corrective action indicated by previous testing.

Post Milestone III Developmental Testing.

Production and Post-production Testing is covered in a separate Unit of Instruction (31-5813), and a brief discussion follows.

(1) **Production Testing.** Testing during the production phase of the materiel life cycle determines the contractor capability to produce an item. It includes the testing necessary to verify that requirements specified in technical data packages and production contracts for hardware and software are met. This testing also confirms corrections of deficiencies reported in earlier developmental and/or operational test reports. Production testing also provides a baseline for post production testing. **Production testing includes the following developmental tests: Production Verification Tests that includes First Article Tests, Follow-on PVT, Comparison Tests, Tests in support of post-deployment software support (PDSS), C3I Interoperability Certification Test, and Quality Conformance (Acceptance) Inspections.**

(2) **Post Production Testing.** Post production testing is the testing of Army materiel during the Deployment and Operational Phases of the materiel life cycle. Testing during this phase includes surveillance and reconditioning tests required to measure the ability of materiel in the field, in storage, and after maintenance actions to meet prescribed specifications and requirements. **Post-production testing includes the following tests: Surveillance Tests and Reconditioning Tests.**

b. Operational Test and Evaluation (OT&E). AR 73-1 and DA Pam 73-5 cover Army Operational Testing and Evaluation. OT is a generic term that encompasses testing and experimentation in a realistic operational environment. OT requires that the test or experiment be conducted with users that are representative of those expected to operate and maintain the system when fielded or deployed. OT&E is planned and coordinated with developmental testing but reported and evaluated independently from DT. OT&E and DT&E may be combined or done concurrently when clearly identified significant cost/time benefits would result or when separation would cause delay involving an unacceptable military risk or unacceptable increase in acquisition cost. The Army Test and Evaluation Command (ATEC) performs OT&E for all ACATs. All OT&E related to acquisition of a system will be documented in part IV of the TEMP. The Premilestone III TEMP incorporates as required testing and experimentation.

(1) As required testing refers to testing that is addressed by the T&E IPT (i.e., the T&E strategy) and approved in the TEMP. The intent is to be flexible and to tailor the T&E approach and conduct only the testing that is required. These tests can include Force Development Test and Experimentation (FDTE), Supplemental Site Test (for information systems), or Customer Tests (CT). CT's are conducted by ATEC for a requesting agency external to OPTEC and is not directly responsive to Army program objectives and is not scheduled or approved by the TSARC unless external sources are required for test support.

(a) FDTE is a generic term encompassing a range of tests and experiments conducted with troops under field conditions to support both materiel system acquisition and the development of doctrine, training, organizations, leadership, materiel, and soldiers (DTOLMS) concepts/requirements. FDTEs support combat development, training development, and materiel acquisition processes. FDTE develops the concept of employment, determines operational feasibility, estimates the potential operational advantage of a proposed system, and can assist the combat and materiel developer in developing the Operational Requirements Document (ORD). FDTE also supports the development and approval of DTLOMS not specifically tied to a materiel system acquisition. Proposals for FDTE initiated by any command or agency must have DCSOPS approval. Once approved, DCSOPS, in coordination with ATEC, will designate the DA Staff proponent and test organization. FDTE uses OMA funds.

(b) CEP (Concept Experimentation Program) is a TRADOC program (see TRADOC Pam 71-9) which provides TRADOC commanders, Battle Laboratories, and combat and training developers, with a quick reaction and innovative evaluation to resolve doctrinal, combat and training development issues. The primary focus is on developing a materiel requirement. The CEP is RDTE funded, and ATEC maintains these funds for disbursement at the direction of TRADOC. TRADOC conducts a CEP Schedule and Review Committee (CEPSARC), similar to the General Officer Test Schedule and Review Committee (TSARC), to approve and prioritize CEP requirements. The CEP is a form of FDTE, and as with FDTE, planning and execution is patterned after OT&E of materiel systems, with as much scientific rigor as practical. Separate, dedicated tests may be necessary to provide data to support the CEP evaluation. These are called CEP tests to distinguish them from other tests conducted for customers. The CEP can be used at any time as part of the continuous evaluation process. The CEP alone is NOT USED to support MS decisions.

(c) Limited Objective Experiment (LOE). (formerly called Battle Laboratory Warfighter Experiments–BLWE). The TRADOC Battle Laboratories conduct these experiments (see TRADOC Pam 71-9) around single events or progressive, iterative simulations with primary relevance to a single issue. LOEs allow the proponent and Battle Lab to conduct low-cost, quick analysis of an issue. LOEs are funded by sources other than the CEP.

(d) Army Warfighting Experiments (AWEs). AWEs are the culminating efforts in evaluating major increases to warfighting capability. They cross TRADOC domains of DTLOMS. They impact most, if not all, of the battlefield dynamics and battlefield operat-

ing systems. They are sponsored by the CG, TRADOC and approved and resourced by the CSA. They have extensive involvement by HQDA, FORSCOM, AMC, TRAC, and ATEC. These experiments, using progressive and iterative mixes of high-fidelity constructive, virtual and live simulations provide Army leaders with future operational capability insights. OPTEC will be involved in the AWE planning and execution to collect data and provide an analysis and evaluate the performance of the Experimental Force (EXFOR). In essence, the AWE is performed in a training environment where both training and analysis occur. ATEC will also have the opportunity to review and staff the findings and recommendations for the Final AWE Report. Approved findings and recommendations do not constitute approved requirements. These requirements must still be formally generated IAW TRADOC PAM 71-9 and approved by the CG, TRADOC. The AWE results help direct efforts to further refine requirements and document feasible solutions and their contributions to force effectiveness and efficiency. The early interaction between OPTEC and the Combat Development community is extremely important and cannot be overemphasized.

(e) Advanced Concepts and Technology Program (ACT II). The ACT II TRADOC program (see TRADOC Pam 71-9) provides industry a vehicle to demonstrate its 6.3 independent research and development products to TRADOC and provides TRADOC a means to examine potential technology solutions to FOCs. This unique program is executed in a partnership between TRADOC and AMC's Army Research Office (ARO). Individual contracts are limited to 12 months or less and have a ceiling of \$1.5M. At the conclusion, the contractor provides a demonstration to the Battle Lab. The goal of the program is to find and acquire new and innovative hardware and software that can satisfy capability requirements. However, success can also be a determination that a particular technology is not the solution to a requirement. ARO sends out a Broad Area Announcement (BAA) and industry responds to it. Once the company has been selected, the appropriate contract is negotiated to include Statement of Work, Milestone Schedule, Surveillance Plan, Budget Status Reports, Government Furnished Equipment, Deliverables, and Evaluation Plan. The Evaluation Plan is the interface point between ATEC and all the other players regarding T&E. The Evaluation Plan is a description of how the Battle Lab will assess the system's future value to the warfighter. The Test and Evaluation Coordination Office (TECO), an ATEC office assigned to selected TRADOC installations to provide liaison, coordination, and support to TRADOC elements, provides Battle Labs with extensive knowledge and/or experience in formulating the Evaluation Plan. Those Battle Labs that don't have a TECO do have a collocated ATEC Test Directorate. The TECO can assist in developing guidelines for this plan which the contractor should provide. As a minimum, the deliverable must be checked against the performance requirements. The contractor has agreed to deliver an item capable of performing to a desired level, and the Battle Lab Project Officer and COR must evaluate that item and judge it as being acceptable or not. The Evaluation Plan should consider not only the final project, but also any intermediate milestones from which an evaluation would provide indications that the item is not performing to standard. The contractor should staff the Evaluation Plan through the project officer before contract start to ensure methodology, measure of performance and evaluation, and deliverable are acceptable to the government.

(2) Premilestone III tests for materiel systems ranging normally from MS I to MS III.

(a) Early User Test and Experimentation (EUTE). EUTE is a generic term encompassing a system level test employing representative user troops during the Program Definition & Risk Reduction Phase. EUTEs purpose is to test a materiel concept, support planning for training and logistics, identify interoperability problems, and identify future testing requirements. EUTE estimates: the potential of the new materiel system in relation to existing capabilities; the relative merits of competing prototypes; the adequacy of system concepts for employment, supportability, trainability, organizational, doctrinal and tactical requirements. This testing also provides pertinent information to aid in design of the Initial Operational Test and Evaluation (IOT&E). EUTE provides data to support Milestone II decisions. FDTE and/or CEP may comprise all or part of EUTE. The line between EUTE and FDTE conducted prior to MS II is indistinct. Funding is the best discriminator. EUTE is RDTE funded.

1 Early User Test (EUT). A test prior to MS II conducted with RDTE funds, which uses procedures described for OT, modified as necessary by maturity and availability of test systems and support packages. EUT seeks answers to known issues which must be addressed prior to MS II.

2 Early User Experiment (EUE). A field experiment conducted to generate data, which is subsequently used to identify potential system related solutions, or to define issues to be addressed at MS II and beyond.

3 Limited User Test (LUT). A generic term encompassing all RDTE funded testing, normally conducted between MS II and MS III, that is not a part of IOT. LUT addresses limited operational issues and is used to accomplish the following: (a) testing necessary to supplement DT before a decision to purchase long-lead items or an LRIP release decision for IOT; (b) verify a fix to a problem discovered in IOT before the production decision (so important that the verification of fix cannot be deferred to FOT); (c) support NDI or modifications that may not require a dedicated phase of IOT before a production decision. A LUT will not circumvent requirements for IOT, nor be used to piece-meal IOT through a series of limited objective tests.

(b) Initial Operational Test and Evaluation (IOTE). The IOTE is a field test, under realistic operational conditions, of a production-representative system (or key component of such a system) to determine operational effectiveness and suitability for soldiers in combat or when otherwise deployed. The IOTE is conducted to support a MS III decision in a traditional acquisition program, and MS III or MS II/III in a streamlined acquisition program. The OT environment is equivalent to that expected during the initial operating capability (IOC) and as operationally realistic as possible, to include projected realistic threats. Organizational units, Tables of Organization and Equipment (TOE) units, provisional units, or elements typical of those that will employ and support the system will be used. Typical soldiers operate and maintain the system under realistic field conditions simulating actual deployment. The IOTE determines if the minimum acceptable operational performance requirements, specified in the Operational Requirements Document, have been satisfied. *Title 10, US Code, Section 2399*, states that "... A Major Defense Acquisition Program many not

proceed beyond Low-Rate Initial Production until Initial Operational Test and Evaluation of the program is completed.” *US Code 10, Section 138* defines the IOT&E as “The field test under realistic combat conditions, of any item of (or key component of) weapons, equipment, or munitions for the purpose of determining the effectiveness and suitability of the weapons, equipment, or evaluation of the results of such test.”

(3) The Production and Deployment Phase has one required operational test, **Follow-On Operational Test and Evaluation (FOTE)**. The objective of FOTE is to fill data voids from the IOTE or to verify correction of deficiencies in materiel, training, or concepts. When IOTE uncovers problems that dictate a low-rate (rather than full-rate) production decision, then an FOTE prior to the full-scale production decision at MS III becomes a requirement. Otherwise, the FOTE may not be necessary. FOTEs normally use OMA funds.

(4) And finally, across the spectrum of tests is the **Joint or Multi-Service Test and Evaluation (JT&E)**. This is a special category of test and evaluation known as Joint Test and Evaluation (JT&E). The Secretary of Defense develops and administers testing programs requiring multi-service participation in planning, conducting, or supporting OSD requirements. These tests are not normally acquisition oriented, but oriented towards interoperability, tactics, or doctrine. Multi-Service T&E is when two or more services want to acquire a “like” system. OSD designates a lead organization to coordinate all testing involving more than one Military Department or Defense Agency. This lead organization will prepare a single Test and Evaluation Master Plan and a single test and evaluation report on the operational effectiveness and suitability of the system for each participating organization. OPTEC is the focal point within the Army for coordinating all aspects of JT&E and Multi-Service T&E.

6. OSD Oversight of Test and Evaluation

a. DoD T&E Policy Documents

o DoDD 5000.1 Defense Acquisition
o DoD 5000.2 - R Mandatory Procedures for Major Defense Acquisition Programs(MDAPS) and Major Automated Information System (MAISs) Acquisition Programs
o DODD 3200.11 Major Range and Test Facility Base
o DoD 5000.3-M-2 DoD Foreign Comparative Testing (FCT) Program.
o DoD 5000.3-M-4 Joint T&E Procedures

b. Service/component Documents. Key service/component guidance is found in the documents listed in figure 2.

Regulation	Title
AR 73-1	Test and Evaluation Policy
DA PAM 73-1	Test and Evaluation
SECNAVINST 5000.2B	Acquisition Policy

OPNAVINST 5000.42D	Research, Development and Acquisition Procedures
AFR 800-2	Acquisition Program Management
AF Policy Directive 99-1	Test and Evaluation
AFI 99-101	DT&E Policy
AFI 99-102	OT&E Policy
AFI 99-103	Test Procedures (5 manuals)

c. DoD T&E OVERSIGHT

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

d. The Congress.

(1) The DoD is required to provide to the Congress the following reports on test and evaluation:

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

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

(c) Annual Operational Test and Evaluation Reports*. The DOT&E prepares an annual OT&E and live fire test and evaluation activities report, in both classified and unclassified form, summarizing all operational test and evaluation and live fire test and evaluation activities within the DoD during the preceding fiscal year. Each such report is submitted concurrently to the Secretary of Defense, the USD (A&T), and Congress not later than 10 days after transmission of the President's Budget for the next fiscal year to Congress (* Not applicable to ACAT 1A programs.)

(d) Beyond Low-Rate Initial Production Report*. Before an ACAT I or DOT&E-designated program can proceed beyond low-rate initial production, the DOT&E, submits a written report to the Secretary of Defense and Congress. This report assesses:

- 1 The adequacy of conducted operational test and evaluation, and

2 Whether the test and evaluation results confirm that the items or components tested are operationally effective and suitable for use in combat.

*Not applicable to ACAT IA programs.

(e) Live Fire Test and Evaluation Report*. An independent OSD Live Fire Test and Evaluation Report is prepared by the DOT&E within 45 days after receipt of the DoD Component's Live Fire Test Report. The Secretary of Defense (or the DOT&E if so delegated) approves the OSD Live Fire Test and Evaluation Report and submits the report to Congress prior to the decision to proceed beyond low-rate initial production. The report shall address survivability or lethality testing in the following cases:

1 Realistic survivability testing of ACAT I and II covered system programs or covered product improvement programs.

2 Realistic lethality testing of ACAT I and II major munitions programs, missile programs, or major munitions or missile covered product improvement programs.

3 Realistic lethality testing of a major munitions program for which more than 1 million rounds (which may be less than an ACAT II program) are planned to be acquired.

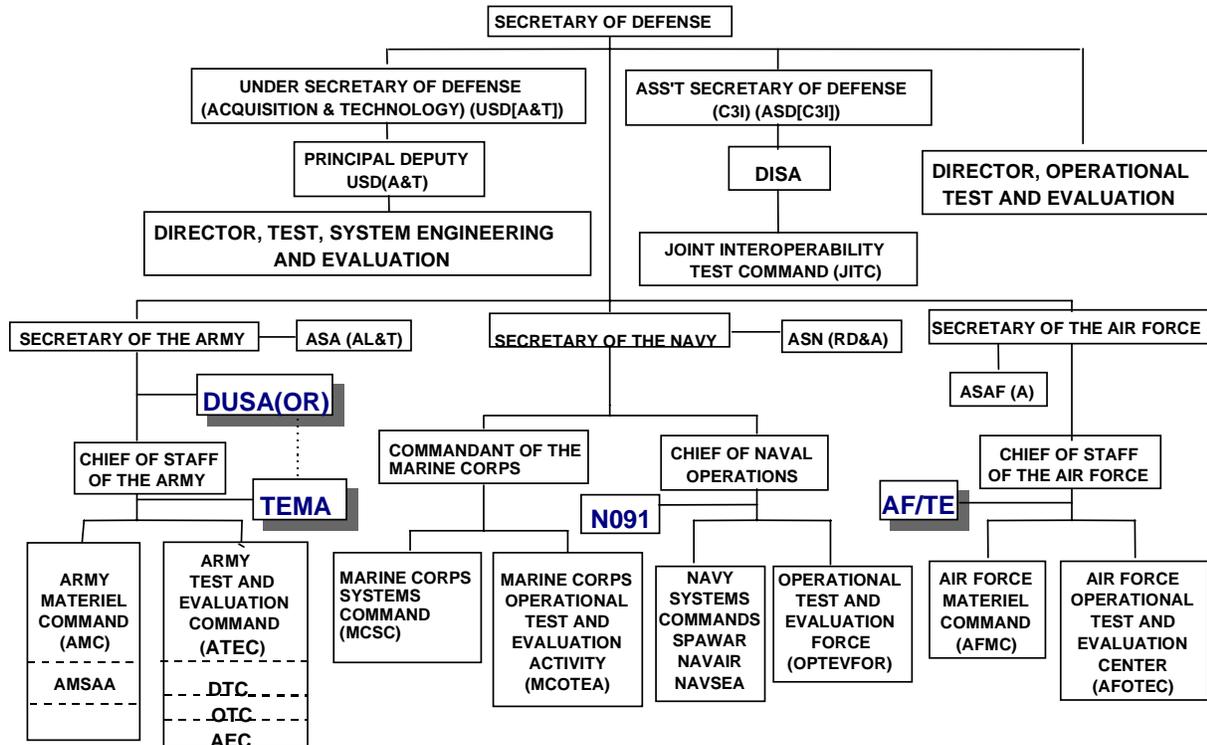
* Not applicable to ACAT IA programs.

(f) Electronic Warfare (EW) Test and Evaluation Reports. Designated Electronic Warfare programs report annually their compliance with the EW Test and Evaluation Process. Report applicability, submission format, and date are contained in the DoD EW plan.

e. OSD Oversight Structure. The DoD organization for the oversight of T&E is illustrated in figure 3. In the OSD, T&E oversight is performed by two primary offices: the Director, Test, Systems Engineering, and Evaluation (DTSE&E) and the Director Operational Test and Evaluation (DOT&E). The management of acquisition programs in OSD is performed by the UnderSecretary of Defense (Acquisition and Technology) (USD[A&T]), who uses the Defense Acquisition Board (DAB) and the Overarching Integrated Product Team (OIPT) to process information for the senior-level decision process in the acquisition of weapon systems. Both DOT&E and DTSE&E organizations are active participants in the program integrated product teams (IPT).

Figure 3 (DoD T&E ORGANIZATION)

DOD TEST AND EVALUATION ORGANIZATION



NOTE: Other Defense Components (e.g. DFAS, DLA, DISA, etc..) are also subject to rules and regulations governing Test & Evaluation

The Army is complying with OSD guidance by combining DT&E and OT&E as much as possible to streamline the process, eliminate redundancies, and reduce overall cost of testing. The Army solution has been to plan DT done by one command and OT done by another as closely together as possible, and to consolidate all evaluation under one command. The objective organization was achieved 1 October 1999 as the Army Test and Evaluation (ATEC) conducting all developmental and operational testing and evaluation and making one consolidated report.

f. Director, Test, Systems Engineering and Evaluation (DTSE&E). The DTSE&E serves as the principal staff assistant and advisor to the USD(A&T) for T&E matters. He or she has authority and responsibility for all DT&E conducted on designated major programs and OSD Oversight programs.

(1) Duties of the DTSE&E

(a) Serves as the focal point for coordination of all major and OSD oversight program test and evaluation master plans (TEMPs). Signs for approval of DT&E portion of TEMPs;

- (b) Reviews major defense acquisition program documentation for DT&E implications and resource requirements to provide comments to the USD(A&T), DAB, OIPT, or MDA;
- (c) Observes DT&E to ensure adequacy of testing and to assess test results;
- (d) Provides the DAE and OIPT with a technical assessment of developmental T&E conducted on a weapon system;
- (e) Provides advice and makes recommendations to the USD(A&T) and issues guidance to the component acquisition executives with respect to DT&E;
- (f) Ensures that threat target and simulator acquisitions meet developmental and operational test and evaluation requirements;
- (g) Performs independent oversight of Component model and simulation validation processes;
- (h) Performs the administrative processing of nominations and charters for joint developmental test and evaluation (JT&E) programs;
- (i) Provides management oversight of the Major Range and Test Facility Base (MRTFB) Figure 4 lists the MRTFB locations for all the services;
- (j) Administers the Foreign Weapons Evaluation Program;
- (k) Confirms, with advice from the Assistant to the Secretary of Defense (Atomic Energy), that nuclear survivability and hardness objectives have been addressed during DT&E;

Figure 4 - DoD Major Range and Test Facility Base [MRTFB] Facilities

TEST CENTER	LOCATION
1. Kwajelein Missile Range	Kwajelein Atoll
2. 30 th Space Wing	Vandenberg AFB, CA
3. Naval Air Warfare Ctr, Weapons Div	Point Mugu, CA
4. Naval Air Warfare Ctr, Weapons Div	China Lake, CA
5. Air Force Flight Test Ctr	Edwards AFB, NV
6. Air Warfare Ctr	Nellis AFB, NV
7. Yuma Proving Ground	Yuma, AZ
8. Dugway Proving Ground	Dugway, UT
9. Utah Test and Training Range	Hill AFB, UT
10. U.S. Army Electronic Proving Ground	Fort Huachuca, AZ
11. Joint Interoperability Test Command	Fort Huachuca, AZ
12. White Sands Missile Range	White Sands, NM
13. 46 th Test Group	Holloman AFB, NM
14. Arnold Engineering Development Ctr	Tullahoma, TN
15. Air Force Development Test Ctr	Eglin AFB, FL
16. Aberdeen Test Ctr	Aberdeen, MD
17. Naval Air Warfare Ctr, Aircraft Div	Patuxent River, MD
18. 45 th Space Wing	Patrick AFB, FL
19. Naval Air Warfare Ctr, Aircraft Div	Trenton, NJ
20. Atlantic Undersea Test & Evaluation Ctr	Andros Island, Bahamas
21. Atlantic Fleet Weapons Training Facility	Roosevelt Roads, Puerto Rico

g. DTSE&E and Service Reports. During the testing of major and designated weapon systems, the DTSE&E and Services interaction includes the following reporting requirements:

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

(2) Prior to a milestone decision or the final decision to proceed beyond LRIP, T&E results with conclusion and recommendations must be submitted to the DTSE&E.

h. Director Operational Test and Evaluation (DOT&E). As illustrated in figure 3, the director reports directly to the SECDEF and has special reporting requirements to the Congress. The DOT&E's responsibility to the Congress is to provide an independent assessment of the operational effectiveness and suitability of new weapon systems.

(1) **Duties and Functions of the DOT&E.** The specific duties of DOT&E are outlined in DoD Directive 5141.2. The functions of the office include:

- (a) 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);
- (b) Signing the TEMPs for approval of OT&E and LFT&E and approving the OT&E funding for major systems acquisition;
- (c) Approving test plans on all major systems prior to system starting OT&E and LFT&E (approval in writing required before operational testing may begin);
- (d) Providing observers during preparation and conduct of OT&E and LFT&E;
- (e) Analyzing results of OT&E and LFT&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 test and evaluation performed;
- (f) A final decision to proceed with a major program beyond LRIP cannot be made until DOT&E has reported (Beyond LRIP and LFT&E Report(s)) to the SECDEF and to congressional Committees on Armed Services and Appropriations.

(2) DOT&E and Service Interactions. For OSD oversight designated acquisition programs, the Service provides the DOT&E the following:

- (a) A draft copy of the Operational Test Plan for review;
- (b) The final IOT&E Test Plan for approval.
- (c) Significant Test Plan changes;
- (d) The final Service IOT&E report must be submitted to DOT&E before the DAB Milestone C review.

7. Testing to Support the Life Cycle Model.

a. Pre-Milestone A.

- (1) Demonstration and experimentation before MS A:
- (2) Provides fundamental knowledge for solutions of identified military problems;
- (3) Develops and assesses feasibility and practicability of proposed solutions;
- (4) Determines their parameters;
- (d) And addresses technological options or uncertainties.

(2) During the initial planning phase, the combat developer ensures special effort to characterize the realistic environment of the system. This includes organizational structures, skill levels, manpower requirements, threat, mobility and deployability requirements, climatic extremes, electromagnetic environmental effects, and concepts of operation and maintenance. The CD investigates DT and OT options to support evaluations of the system in its intended environment. This early involvement will facilitate early development of the MNS, ORD, draft TEMP and the T&E portions of the acquisition strategy.

b. Concept and Technology Development.

(1) Test and evaluation in Concept and Technology Development Phase:

- (a) identifies and assesses high risk areas, critical components and subsystems;
- (b) establishes safety for operational testing;
- (c) provides data for component development of a potential requirement, tactics, doctrine, organization, training, transportability, and logistic support for the overall system;
- (d) assists in selecting preferred alternative system concept;
- (e) and assesses the operational impact of candidate technical approaches.

(2) Before the (MS) A decision, appropriate T&E must be documented in test and evaluation reports and the TEMP. This aids in selecting preferred alternative system concepts, technologies, and designs. In particular, modeling and simulation is helpful and encouraged in this phase to aid in the assessment of alternatives at MS A.

c. System Development and Demonstration Phase.

(1) System T&E conducted in this phase:

- (a) supports the hardware and associated software design through a test-analyze-fix-test program approach, or a Model-Test-Analyze-Fix-Model approach (part of the Simulation, Test and Evaluation Process [STEP]) performed at the component, subsystem, and system level;
- (b) identifies the preferred technical approach, including the identification of technical risks and feasible solutions;
- (c) examines the operational aspects of support requirements and the selected alternative technical approaches;
- (d) estimates the potential operational effectiveness and suitability of systems;
- (e) supports the materiel change process;

- (f) identifies design risks;
- (g) establishes contractual compliance, including component qualification;
- (h) makes preliminary assessment of MANPRINT requirements;
- (i) and supports the evaluation of the critical developmental characteristics and operational issues.

(2) T&E also addresses doctrine, training, organization, and logistics support aspects of the system using surrogate systems if necessary.

(3) Before the MS B decision, adequate and documented T&E helps to identify the preferred technical approach, including the identification of technical risks and feasible solutions. Test data establishes realistic program performance and suitability thresholds. In all cases, appropriate and adequate T&E (including contractual testing) must precede the MS B decision, reducing risk and uncertainty before committing more resources to the Production and Deployment Phase.

d. Production and Deployment Phase.

- (1) In Production and Deployment, the T&E emphasis is:
 - (a) prototype and pre-production systems to ensure the design is sufficiently mature;
 - (b) an SSP for testing;
 - (c) and ensuring the system tested will be representative of the production hardware/software to validate the system can be the system expected to be produced.
- (2) Test and evaluation conducted during this phase supports maturation of the engineering development prototype hardware and associated software, the SSP, plus provides a valid estimate of the system's safety, operational effectiveness and suitability. It ascertains whether engineering is complete (including design and maintenance engineering), identifies design problems and ascertains that solutions to these problems are in hand. It supports the materiel change process, reduces design risks, establishes contractual compliance, provides information for type classification determination, and validates general and detailed specifications, standards, and drawings for use in production. It supports the evaluation of the critical developmental characteristics and operational issues.
- (3) Before the MS C full production decision, production representative article testing results must confirm that all deficiencies identified have solutions available; and that items or components actually tested are effective and suitable for their intended use. Testing during this phase includes all developmental testing required to verify the production process and determines the adequacy and validity of component changes in the system.

(4) IOTE provides the necessary data for the independent evaluation of the system's operational effectiveness and suitability. The system tested must be sufficiently representative of the expected production system to ensure T&E validity supports the production decision. The Director, Operational Test and Evaluation (DOT&E) approves adequacy of IOT test plans for the OSD oversight systems prior to the test. Written Operational Test Readiness Statements (OTRS) are prepared by the MATDEV, CBTDEV, and TNGDEV prior to the start of any OT. The materiel developer formally certifies via their OTRS that the system is production representative and ready for IOTE; the CBTDEV's OTRS verifies that the doctrine, organization, threat, logistics concept, crew drill, and SOPs in the CBTDEV's support packages are complete, represent planned employment, and are approved for use for IOTE; and the TNGDEV's OTRS verifies that the training concepts and materiel and crew drills included in the training support package are complete, representative of the training package to be used at fielding, and approved by TRADOC for use during IOTE. To support a LRIP decision, the required operational assessment may be based on a combination of DT and OT data and does not necessarily require an IOTE.

(5) Test and evaluation is an integral part of the acceptance and introduction of improvements to the system. T&E in this phase also provides a way to react to new threats and reduce life-cycle costs.

(6) During initial deployment of the system, FOTE continues as necessary to refine estimates made in IOTE, provide data to evaluate changes, verify corrected deficiencies in materiel, training, or concept, plus determine overall readiness of the system. All tests focus on confirming fixes to problems identified in earlier tests.

(7) Testing during this phase includes testing the complete system necessary to verify that requirements specified in technical data packages and production contracts for hardware or software are met. Production testing also provides a baseline for follow-on post production testing.

(8) Post-production testing is a follow-on to production testing and includes surveillance and reconditioning tests required to measure the ability of materiel to meet user requirements.

7. Evaluation.

a. Broadly defined, evaluation is the process of determining the worth or value of a course of action through comparison. In materiel acquisition, evaluation is a subjective determination of a system's or a concept's military value and potential based on results from operational or developmental testing; reliability analysis; human factors studies; or simulations and war games.

b. The evaluation process provides an objective assessment. Objectivity and independence in the evaluation process supports objective decision making. Test data comes from an entire spectrum of tests: operational, developmental, concept evaluation, innovative, technical feasibility, joint, production, product improvement, nondevelopmental item (NDI), on-site

user tests, and other analyses (i.e., studies and computer modeling). Although all significant data is of interest, the evaluation plan should not require duplicative testing. T&E's process is to identify, assess, and reduce acquisition risks, then evaluate technical performance, system operational effectiveness and suitability. The issues identified in the Test and Evaluation Master Plan (TEMP) start the assessment process. Sufficient test and evaluation before each major decision point insures that the major objectives of one phase are met and issues are addressed before the next phase begins. Quantitative data must be used to show that the major objectives have been met and to support subjective judgments relative to system performance before progressing to the next acquisition phase.

8. Continuous Comprehensive Evaluation (C²E). Beginning in 1985, the Army initiated a review to determine how to make testing a more effective contributor to the development process. The result was C²E. The basic intent of C²E is to move from a series of discrete event tests to a continuous process linked throughout the entire development cycle. The independent tester becomes involved in planning during the Concept Exploration and Definition Phase. The data generated in early development phases becomes visible and continues as the system moves into formal DT and OT. Testing can then easily verify that deficiencies found in the development process are, in fact, corrected. The most current information becomes available for subsequent testing. In this manner, testing not only fulfills the function of verifying specifications, it also contributes to system modifications to optimize the effectiveness of the product. This program includes that testing related to evaluating the supportability of the system.

9. Summary. We have numerous reasons to test, the most important being verification of effectiveness and suitability for the soldier. The iterative process of testing allows tailoring, provides a common data base, has an independent evaluation and aids the decision makers doing risk trade-offs. Developmental testing differs from operational testing. DT has a controlled climate, contractor involvement, looks at prototypes, uses technical personnel, and provides data for measures of technical performance evaluations. OT uses soldiers in a realistic combat environment using representative production items to determine operational effectiveness and suitability. AMC currently handles DT, while OPTEC handles OT. Although all DT and OT will be combined under one command effective 1 October 1999, there will still be distinctive differences and functions between Developmental and Operational Testing. This streamlining should only serve to make more efficient use of our valuable test and evaluation resources. As a system moves through the LCM, tests supply data for evaluations needed for MS decisions. While testing is only one small part of the total plan to get a system to the soldiers, it is a very important step to ensure the process in the LCM works. **Fewer dollars will likely be spent if you do it right the first time.**

THIS PAGE INTENTIONALLY LEFT BLANK