GAO

Report to the Chairman, Senate Committee on Governmental Affairs

December 1986

WEAPON PERFORMANCE

Operational Test and Evaluation Can Contribute More to Decisionmaking

GAO/NSIAD-87-57

GAO

United States
General Accounting Office
Washington, D.C. 20548

National Security and
International Affairs Division

B-222886

December 23, 1986

The Honorable William V. Roth, Jr.
Chairman, Committee on Governmental
Affairs
United States Senate

Dear Mr. Chairman:

This report was prepared at your request and addresses whether operational test and evaluation results have been limited because of longstanding problems in test planning, test conduct, and the reporting of test and evaluation results.

As arranged with your office, unless you publicly announce its contents earlier, we plan no further distribution of this report until 30 days from the date of issue. At that time, we will send copies to cognizant congressional committees and other interested parties and make copies available to others upon request.

Sincerely yours,

Frank C. Conahan
Assistant Comptroller General

Executive Summary

Purpose

Short of war, operational test and evaluation (OT&E) is the primary means of predicting weapon system performance in a combat representative environment. To replicate realistic wartime conditions, military personnel representative of those expected to operate a weapon system participate in the field testing. OT&E results are particularly critical to help decide whether weapon system modifications are needed before production to avoid costly retrofit or improvement programs after systems are fielded.

The Chairman of the Senate Committee on Governmental Affairs requested GAO to (1) summarize historical problems on the adequacy of OT&E reported since 1970 and (2) examine recently deployed major weapon systems and determine whether the usefulness of OT&E has been limited.

Background

Two types of test and evaluation of weapon systems serve distinctly different purposes. Developmental test and evaluation is designed to assist the engineering design and development process and to verify that technical performance specifications are met. Development test and evaluation is normally accomplished or managed by the agency responsible for developing the weapon system.

OT&E is designed to assess a weapon system's operational effectiveness and suitability in a realistic combat environment. OT&E is of particular value to decisionmakers because it is performed by organizations in each military service that are intended to be independent of the developing and using commands.

The Department of Defense (DOD) Directive on Test and Evaluation states that an initial phase of OT&E should be done during the development of a weapon system before the decision to proceed into low-rate initial production—using production representative articles or prototypes. The Directive also states that another phase of OT&E should take place before the full-rate production decision.

Public Law 98-94, enacted in 1983, established the Director of OT&E, to ensure that OT&E policies are given proper recognition and enforcement. The legislation requires the Director, among other things, to report to the Secretary of Defense and the authorizing and appropriating committees on the adequacy of test and evaluation results to confirm that items tested are effective and suitable. Also, the law requires the Director to prescribe the policies and procedures for the conduct of OT&E in DOD.

Results in Brief

The usefulness of OT&E in estimating a weapon system's performance has been limited because of long-standing problems in test planning, test conduct, and the reporting of test and evaluation results. For example

- selection of test sites have not always been representative of operating environments,
- test objectives and evaluation criteria have not always been established,
- test resources have not always been available or adequate, and
- test reports have sometimes lacked complete, current, or accurate data.

Principal Findings

Decisionmakers rely on the results of OT&E to estimate weapon performance. Historical problems involving test planning and conduct, and reporting of results surfaced during the acquisition process for several of the deployed weapon systems GAO reviewed—which can result in the fielding of a weapon system with performance shortcomings affecting mission effectiveness that will require redesign and modification. GAO's analysis was based on over 60 previously reported cases dealing with OT&E concerns. GAO also reviewed six recently deployed weapon systems.

Test Sites Were Not Always Representative of Operational Environment

There were 32 historical cases where testing was unrealistic because it did not adequately replicate the operational environment. Most recently, GAO found that in assessing the F/A-18 aircraft's operational suitability, the Navy used contractor support which limited the realism of the test agency's OT&E results and the evaluation of the aircraft's effectiveness.

Test Objectives and Evaluation Criteria Were Not Always Clearly Defined

GAO identified 25 historical cases where test objectives, criteria, and plans were incomplete, unclear, and/or absent before testing began. GAO recently found that when the Ground Launched Cruise Missile was undergoing operational testing, military personnel were confused on how to evaluate and interpret test results because evaluation criteria was not clearly defined.

Test Resources Were Limited or Not Available

There were 27 cases where test resources were limited or not available. GAO recently found that the nonavailability of key subsystems, on the F/A-18 aircraft, such as radar warning receivers and defensive countermeasures equipment, precluded testing the effect this equipment would have on the aircraft's survivability/vulnerability.

Test Reports Have Sometimes Not Contained the Most Current, Complete, or Accurate Data

GAO also identified 22 cases where test reports did not always contain the most current, complete, or accurate data on the performance of major systems before production. GAO recently found that the Army's test report on the Multiple Launch Rocket System did not adequately describe the significance of a problem with the ammunition resupply trailer—namely, when carrying a full load of ammunition and traveling at normal speeds it tipped over.

OT&E Results Are Important in Making Production Decisions

Waiting for a production prototype and operationally testing it before production is frequently not done because of the lack of production representative items especially for weapon programs that are concurrently developed and procured.

GAO recognizes the need to shorten the acquisition cycle for a major program and begin low-rate initial production without the benefit of full OT&E. However, OT&E results allow decisionmakers to assess whether there is a potential need for costly retrofitting and modifications. Thus, in balancing the need to make informed decisions and to shorten the production cycle, decisionmakers must consider the risk of advancing a weapon system without knowledge of whether it can meet operational requirements.

Recommendations

GAO recommends that the Secretary of Defense improve the reporting of OT&E results by requiring the services' OT&E agencies to: (1) state whether OT&E demonstrated that the system met operational requirements, (2) discuss the operational effect of significant test limitations and adverse test results on system performance, and (3) clearly state whether the system tested is operationally effective and suitable.

Other recommendations for improving the usefulness of OT&E results are contained in the report. (See p. 30.)

Agency Comments

DOD agreed for the most part with GAO's recommendations. The draft of the GAO report proposed that the OT&E agencies should be required to recommend whether a system is ready for production. DOD disagreed that the services' OT&E agencies should be required to recommend whether a weapon system is ready for production. Instead, DOD stated that the OT&E agencies should generally make a recommendation on a weapon system's operational effectiveness and suitability.

GAO agrees that the OT&E agencies should present their views on a weapon system's operational effectiveness and suitability. GAO recognizes that the final decision for advancing a system to production rests with the Secretary of Defense; accordingly, GAO has modified its recommendation to recognize this point.

Contents

Executive Summary		2
Chapter 1 Introduction		8
	Congressional Concern With Testing and Weapon System Performance	8
	DOD's Policy on Acquisition and Testing	8
	Office of the Director of OT&E	10
	Objectives, Scope, and Methodology	11
Chapter 2 Historical OT&E Concerns		14
	Past Recommendations Need to Be Fully Implemented to Improve OT&E	15
	Actions Taken and Planned to Improve OT&E	16
	Impediments to Useful OT&E	19
	Agency Comments and Our Evaluation	21
Chapter 3 OT&E Needs to Be More Realistic		22
	Maintenance Practices Not Fully Representative of Operational Environment	23
	Clearer Test Objectives and Criteria Are Needed	24
	Lack of Test Resources	25
	Problems Surfaced During OT&E Were Not Adequately Reported to Decisionmakers	26
	Conclusions	30
	Recommendations	30
	Agency Comments and Our Evaluation	31
Chapter 4 Office of the Director of OT&E		32
Appendixes		
	Appendix I: Review of Selected Reports Issued Since 1970 Dealing With OT&E	34
	Appendix II: Description of Weapon Systems Reviewed	37
	Appendix III: Comments From the Director, OT&E	39

Tables

Table 1.1: Weapon Systems Selected for Review 11
Table 2.1: Historical OT&E Concerns and Adverse Effects 15

Abbreviations

DDT&E	Director, Defense Test and Evaluation
DOT&E	Director, Operational Test and Evaluation
DT&E	Development Test and Evaluation
DOD	Department of Defense
GAO	General Accounting Office
GLCM	Ground Launched Cruise Missile
HARM	High Speed Anti-Radiation Missile
MLRS	Multiple Launch Rocket System
OT&E	operational test and evaluation
OTEA	Operational Test and Evaluation Agency
TEMP	Test and Evaluation Master Plan

Introduction

At the request of the Chairman of the Senate Committee on Governmental Affairs, we summarized the conclusions and recommendations contained in reports we have issued since 1970 concerning the operational test and evaluation (OT&E) of major weapon systems. The Chairman also requested that we examine recently deployed weapon systems to identify operational concerns not adequately addressed during OT&E or not adequately addressed after problems were identified by OT&E.

Congressional Concern With Testing and Weapon System Performance

The Congress has consistently shown a long-standing interest in the performance of major weapon systems and the adequacy of OT&E, planned, completed, and reported. Annual hearings on weapon systems have continually highlighted weapon system performance issues and the need for better OT&E. As early as 1971, the Congress passed Public Law 92-156 requiring the Department of Defense (DOD) to provide the Congress with data on the OT&E results of major weapon systems before committing major production dollars.

DOD's Policy on Acquisition and Testing

DOD's policy[1] on the acquisition and testing of major weapon systems emphasizes the need for early and timely testing to identify and reduce the acquisition risk of costly redesign and modification after deployment.

The Acquisition Process

Improved readiness and sustainability are primary objectives of the acquisition process. To achieve these objectives and acquire weapon systems that meet operational needs of the armed forces, DOD has prescribed procedures, which among other matters, require:

- Realistic cost estimating and budgeting, and adequate funding of procurements, including testing.
- Reduction in the time it takes to acquire and deploy a weapon system. This includes encouraging the use of concurrency in development and production, combined development and operational testing, preplanned product improvements, and combining or omitting phases of the acquisition process when applicable.

[1]DOD Directive 5000.1, Major System Acquisition; DOD Instruction 5000.2, Major System Acquisition Procedures; and DOD Directive 5000.3, Test and Evaluation.

- Delegation of authority to the lowest level at which a comprehensive program view exists; this includes delegating the decision to enter production.

The acquisition of a major system includes three distinct milestones or decision points:

- Milestone I. The decision by the Secretary of Defense approving a specific concept to be pursued into the demonstration and validation phase. This decision establishes thresholds and objectives to be met and reviewed at the next milestone.
- Milestone II. The decision, also by the Secretary of Defense, authorizing program go-ahead—approval to proceed with the full-scale development phase.
- Milestone III. At this point, the decision is made to enter production. This decision is made by the Secretary of Defense or may be delegated to the service secretary, provided thresholds established at milestone II are met.

The Test and Evaluation Function

DOD requires that test and evaluation begin as early as possible and continue through the acquisition process to assess and reduce development risks and to estimate the operational effectiveness and suitability[2] of the system being developed. Meaningful critical issues, test objectives, and evaluation criteria related to the mission are to be established before testing begins.

The successful accomplishment of test and evaluation objectives is a key requirement for committing significant additional resources or to advancing a system from one acquisition phase to another.

Types of Testing

DOD defines test and evaluation as being essentially of two different types—development test and evaluation (DT&E) and OT&E.

Development test and evaluation is part of the engineering design and development process. Its purpose is to verify attainment of technical

[2]Operational effectiveness can be defined as the ability of a system to accomplish its mission when placed in use in the planned operational environment. Operational suitability is the degree to which a system can be placed satisfactorily in field use considering, among other factors, availability, maintainability, and logistic support.

performance specifications and objectives. Development test and evaluation is normally performed or managed by the service activity responsible for developing the weapon system.

OT&E is done to estimate a system's operational effectiveness and operational suitability in its intended environment when operated, maintained, and supported by personnel having the same qualifications as those who would operate the system in the field. OT&E can often identify needed modifications and provide information on tactics, doctrine, organization, and personnel requirements. OT&E is performed by service organizations which are intended to be independent of the organizations responsible for system development. To be most effective, OT&E requires a realistic operational environment.

OT&E done before the low-rate initial production decision is called initial OT&E; OT&E done after production has started is called follow-on OT&E. DOD's policy does not require, however, that OT&E separate from development test and evaluation be accomplished just before starting initial production. In fact, DOD encourages combining developmental and operational testing to reduce cost and acquisition time.

Importance of OT&E

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

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

Office of the Director of OT&E

In 1984, pursuant to Public Law 98-94, which was enacted on September 24, 1983, DOD established, within the Office of the Secretary of Defense, the Office of the Director of OT&E. The Director is a civilian

and principal advisor to the Secretary of Defense on OT&E matters. (See p. 17 for a description of the Director's responsibilities.)

In April 1985, a Director was appointed by the President and confirmed by the Senate. Before the establishment of the Director of OT&E, the Director, Defense Test and Evaluation (DDT&E) was responsible for reviewing, coordinating, and monitoring both DT&E and OT&E. The weapon systems discussed in this report had completed a substantial portion of the acquisition process before the Director of OT&E was established. Therefore, the situation described in this report can be viewed as an agenda of items to be considered by the Director of OT&E.

Objectives, Scope, and Methodology

Our review objectives were to (1) summarize historical problems that have been reported since 1970 on the adequacy of OT&E, (2) identify major improvements DOD made in the area of operational testing, and (3) examine recently deployed weapon systems to determine whether the usefulness of OT&E has been limited.

We reviewed 50 of our reports discussing assessments of OT&E in general or OT&E completed on selected weapons. We also reviewed other studies and reports—dealing with OT&E—issued by other organizations such as the DOD Inspector General. (App. I lists all reports reviewed.)

Our report includes an evaluation of six deployed major weapon systems. These particular weapon systems were selected because they (1) were recently deployed, (2) were of congressional interest in the past, and (3) represented major acquisitions from all three military services.

Table 1.1: Weapon Systems Selected for Review

Air Force Systems	Army Systems	Navy Systems
Ground Launched Cruise Missile (GLCM)	Bradley Fighting Vehicle	F/A-18 Aircraft
EF-111A Tactical Jamming System	Multiple Launch Rocket System (MLRS)	AGM-88A High Speed Anti-Radiation Missile (HARM)

(A description of the weapon systems is contained in app II.)

Before selecting the weapon systems for review, we discussed our objectives with officials of the Office of the Director of OT&E. These officials provided some initial candidate weapon systems from which we made our final selection. We did not evaluate the need for the weapon systems nor the rationale for the production decisions. We did, however, review

the assessment of OT&E by DDT&E—a function that is now performed by the Office of the Director of OT&E.

For the selected weapon systems, we reviewed test plans to identify test criteria and test limitations and compared this with the test results reported. We also discussed test results with knowledgeable officials from the services' OT&E agencies. We interviewed officials from the program offices and deployed units to obtain their views of operational successes or concerns with the deployed systems.

We also attempted to obtain the cost to correct known shortcomings after deployment of the weapon systems. However, available financial data did not clearly distinguish between operational shortcomings and capability enhancement programs. Accordingly, financial data obtained could not be used to identify the cost of only fixing weapon system shortcomings.

Our review was performed in accordance with generally accepted government auditing standards.

Historical OT&E Concerns

Since 1970, reports have stressed, among other things, the importance of effective and timely OT&E, including the accurate and timely reporting of test results. However, problems in test planning, conducting, and reporting have limited OT&E as an effective tool in decisionmaking. We analyzed 63 reports issued over the last 16 years and identified:

- Twenty-five cases where test objectives, criteria, and plans were incomplete, unclear, and/or absent.
- Twenty-seven cases where test resources were limited or not available for testing.
- Thirty-two cases where testing was unrealistic because it did not adequately replicate the operational environment.
- Twenty-two cases where test reports did not always contain the most current, complete, or accurate data on the performance experienced by major weapon systems before production. Also, the operational effects of test limitations or results were not disclosed.
- Forty-one cases where production was approved before OT&E was started or completed.

These historical concerns are shown in table 2.1.

Table 2.1: Historical OT&E Concerns and Adverse Effects

		OT&E Problem Areas		Adverse Effects	
		GAO	DOD & others	GAO	DOD & others
A	Inadequate OT&E plans, objectives, and/or criteria	18	7	•	•
B	Test resources have limitations or not available	16	11	•	•
C	Inadequate test realism —personnel —environment —equipment	22	10	•	•
D	Production decision made before OT&E started or completed —concurrency —accelerated acquisitions	31	10	•	•
E	Inadequate reporting of OT&E results	14	8	•	•
F	Decreased effectiveness/ suitability	•	•	15	4
G	"Get Well" Program required —additional costs —deployment delays	•	•	11	5
H	Reported OT&E results inconclusive, questionable, optimistic, or misleading	•	•	31	5

Note The problem areas described above are a summation of major recurring OT&E concerns identified in the reports we reviewed (See app I)

An example of the adverse effect of a long-standing concern with OT&E is illustrated by the Army's Sergeant York weapon program, namely that OT&E results were not available before the decision to begin limited production. The recent cancellation of the program emphasizes the adverse effect of a long-standing concern with OT&E—over a billion dollars was spent on an unproven system that ultimately had to be terminated.

Past Recommendations Need to Be Fully Implemented to Improve OT&E

Our recommendations to improve the planning, conducting, and reporting of OT&E results have called for such efforts as:

1. Better test planning. It was recommended that DDT&E enforce existing requirements for preparing and approving a weapon system Test and Evaluation Master Plan (TEMP)[1] before milestone I (the decision to proceed into the concept demonstration and validation phase) and subsequent decision milestones. Further, it was recommended in another

[1]The TEMP is the primary document used to assess the adequacy of planned test and evaluation. It should describe technical and operational issues, availability of test resources, and key OT&E events

report that operational tests critical to determining system effectiveness and suitability be done before initial production decisions.

2. Better identification of needed test resources, including key subsystems before the start of operational testing. It was recommended that OT&E agencies state in the initial TEMPs their ability to adequately assess effectiveness and suitability, given available resources, or to identify the operational areas that may not be adequately assessed due to the lack of test resources or key subsystems.

3. More realism in testing. It was recommended that (a) personnel doing the test be more representative of the type and qualifications of the typical personnel who will operate and maintain the equipment when deployed, (b) testing be done in more representative geographic settings and weather conditions, and (c) equipment be used which is representative of the threat.

4. More comprehensive disclosure of the limitations to test results. It was recommended that test reports should better describe the operational risk/effect of (a) not testing against an operational requirement, (b) test limitations due to lack of test resources or key subsystems, or (c) adverse test results. That is, the effect of incomplete/inadequate testing should be clearly spelled out. The need for clear and comprehensive reporting of test results and operational implications is particularly important since test reports are the primary source of information for the DDT&E and the Director of OT&E in making their assessments.

In addition to our office, others have commented on the importance of and need to improve operational testing. For example, in June 1986, the President's Commission on Defense Management emphasized the importance of performing operational testing early in advanced development and continuing through full-scale development, using prototype hardware. The Commission reported that the proper use of operational testing is critical to improving the operational performance of new weapons.

Actions Taken and Planned to Improve OT&E

DOD has implemented or is planning to implement many of the recommendations made over the years. Some of the more significant actions taken by DOD to improve overall OT&E are described in the following sections.

has assessed programs that are being considered for procurement at rates above low-rate initial production.

Impediments to Useful OT&E

A major obstacle in accomplishing sufficient OT&E to fully support the production of major weapon systems is reflected in a statement by a former Deputy Under Secretary of Defense:

> "Many people in the acquisition process take for granted that weapon systems will work as advertised and regard test and evaluation as a wicket to be passed rather than a tool in the acquisition process."

Factors hindering the planning, conducting, and reporting of OT&E are:

- the belief that a weapon system must go into production and be deployed regardless of its readiness for production (production imperative),
- compressed acquisition cycles limiting the time available to perform planned OT&E before production, and
- the diffusion of the responsibility for correction of deficiencies identified during OT&E.

Production Imperative

In hearings before the Senate Committee on Governmental Affairs in June 1983, it was stated that weapon systems create a momentum of their own. All too often there is an overpowering tendency to "build now and fix problems later" and to ensure that rigorous operational testing is not allowed to stand in the way of a decision to produce a weapon. As stated in the hearings, this tendency is the result of (1) many persons believing that testing of weapon systems causes unnecessary delays and increases costs, (2) military and civilians within DOD having their names and reputations at stake, and (3) contractor and its personnel wanting to move forward with production because their profits and jobs depend on it.

Role of OT&E in Current Acquisition Environment

DOD's acquisition policy encourages concurrent development and production of major weapon systems. In a 1985 report,[3] we stated that in all of the concurrently developed and produced weapon systems reviewed, DOD did not obtain OT&E results critical to assessing mission performance before production start-up, even though DOD had initially planned to

[3]Production of Some Major Weapon Systems Began With Only Limited Operational Test and Evaluation Results (GAO/NSIAD-86-68, June 19, 1986).

have these results available before making such decisions. We recognize that DOD encourages the concurrent development and production of weapon systems because of the importance of shortening the acquisition cycle.

However, DOD's policy does not address how a shortened acquisition cycle can provide for planned OT&E results before production start-up.

In our earlier report, we concluded that concurrency[4] can be an effective technique to expedite development and production of weapon systems, provided the practice is well planned and controlled. This requires that safeguards be built into a program to minimize the risk associated with concurrency. At the very least, the safeguards should provide for performance of at least one separate phase of OT&E and the completion of any planned OT&E before production start-up. DOD's policy specifies that the degree of concurrency will be based on the savings in acquisition time balanced against cost, risk, and urgency of the mission need in each acquisition program. Weapon system programs with planned concurrency should allow for special attention to OT&E so that performance risks resulting from a shortened acquisition time do not affect the planned deployment date.

Diffusion of Management Responsibility

Other factors hampering the conduct and usefulness of OT&E include the diffusion of (1) control and responsibility for test resources and (2) the responsibility for correction of deficiencies identified during OT&E. Without sufficient management emphasis and accountability to ensure that test resources are sufficient to meet test goals and objectives and deficiencies identified during OT&E are corrected before production and deployment, OT&E results will continue to be of limited usefulness.

Performance shortcomings identified during OT&E are frequently not corrected before production and deployment. While it could be argued that the failure to correct a deficiency before production is due to an urgent military need to deploy a system (as in the case of GLCM) or the lack of funds, it may, nonetheless, be due to ineffective planning to ensure that deficiencies are corrected before or after production. In April 1984, the Air Force emphasized that an improved process for correcting deficiencies was needed throughout the acquisition cycle. Further, review of deficiencies was cited as an area needing to be institutionalized for major acquisition programs at the Secretary of the Air Force level. To

[4]Ibid

Office of the Director of OT&E

DOD was required by law to establish an OT&E office with civilian leadership within the Office of the Secretary of Defense. This Office was established in 1984.

Among other things, the OT&E Director's responsibilities and functions would include:

- Reviewing new major system requirement documents, concept papers, test plans, and other documents for OT&E implications, including approval of the OT&E sections of TEMPs for major defense acquisition programs.
- Examining budget submissions to determine the adequacy of OT&E funding, including test facilities and equipment.
- Analyzing the results of OT&E done for each major defense acquisition and reporting to the Secretary of Defense and the Committees on Armed Services and Appropriations on the adequacy of test and evaluation results to confirm that items tested are effective and suitable for combat before a decision is made to proceed beyond low-rate initial production.

Independent Operational Test Agencies

Another recommendation essentially suggested that OT&E be accomplished by independent organizations reporting directly to the chiefs of staff, service secretaries, or both.

In February 1971, the Deputy Secretary of Defense directed the military departments to designate field commands, independent of the weapon system developers, to be responsible for OT&E. This direction was implemented and, since then, OT&E on major systems has generally been done by the Army's Operational Test and Evaluation Agency (OTEA), the Navy's Operational Test and Evaluation Force, the Air Force's Operational Test and Evaluation Center, and the Marine Corps' Operational Test and Evaluation Activity. These testing agencies are independent of the acquisition/development command, as well as the using command, and report directly to their chiefs of staff.

Historically, the services' test agencies report on a weapon system's operational performance capabilities before major decisions are made to proceed with the full production of the weapon program.

DOD Emphasis on OT&E

Another very pertinent but somewhat general recommendation pertains to the Secretary of Defense's communication to the military departments

concerning the importance assigned to accomplishing of productive, objective, and timely OT&E.

Starting in 1970, the Secretary of Defense had emphasized the importance of adequate OT&E in various policies and practices to improve the acquisition of weapon systems. Among other matters, these directives required (1) the Defense Systems Acquisition Review Council to formally review the progress of weapon systems at three predetermined milestone points, including milestone III, when the first major production decision is to be made and (2) that an initial phase of operational testing be accomplished for all new weapon systems before the first production decision. Although there have been several later changes and revised directives to the acquisition policies,[2] the thrust of DOD's policy on OT&E has remained the same. Specifically, the current policy requires that (1) before a production decision is made, adequate test and evaluation is to be done to provide a valid estimate of a system's operational effectiveness and suitability and (2) items to be tested adequately represent the expected production items to ensure that a valid assessment can be made of the system expected to be produced.

DOD has recently revised its directive governing test and evaluation to help strengthen OT&E. The current directive contains subtle but important changes which should, if properly implemented, improve the usefulness of OT&E. For example, DOD's earlier policy merely required that OT&E provide a valid estimate of a system's operational effectiveness. The current directive requires OT&E to estimate a system's effectiveness and suitability to determine whether it meets required operational thresholds. Further, the current directive (1) requires that a critical analysis of test resource shortfalls and plans to correct the limitations be addressed in TEMPs and (2) sets out the roles and responsibilities of the Office of the Director of OT&E.

Another indicator of the importance being given by DOD to operational testing is contained in the Secretary of Defense's fiscal year 1985 Annual Report to the Congress. In this report, the Secretary of Defense said the Office of the Director of OT&E has reviewed the adequacy of test plans on eight major programs before starting operational testing, and

[2]For example, the production decision on major system acquisitions, which formerly was reserved for the Secretary of Defense, is now often delegated to the lowest level of the military department at which a comprehensive view of the program rests, provided the Office of the Secretary of Defense thresholds established at milestone II are met. On June 3, 1986, the Joint Requirements and Management Board was established to replace the former Defense Systems Acquisition Review Council. Its responsibilities, among other things, are to recommend full-scale development and high-rate production for major programs.

ensure the new system works, the Air Force recommended increased emphasis on discipline and consistency in deficiency tracking and reporting.

Agency Comments and Our Evaluation

DOD agrees that long-standing OT&E problems have limited the quality and usefulness of OT&E results to decisionmakers. The military services and the Director of OT&E have initiated a number of efforts to improve operational testing. The Air Force is instituting the "Baseline Correlation Matrix" that aligns requirements, specifications, and test criteria to improve test planning. The Army has established its "Continuous Comprehensive Evaluation" program to assess a weapon system's operational effectiveness and suitability throughout the acquisition process. The Navy promulgated new instructions setting forth requirements designed to improve the realism of OT&E.

DOD also agreed that there are three primary factors that hinder OT&E and stated that these problems are being addressed under recent initiatives by the Office of the Secretary of Defense and the military services. For example, DOD stated its newly revised Directive on "Test and Evaluation" emphasizes: (1) early OT&E involvement, (2) the requirement for OT&E before production decisions, and (3) long-range test resource planning.

OT&E Needs to Be More Realistic

Test realism is a critical factor in the performance of OT&E and in the quality of test results achieved. If a test that is supposed to replicate the operational environment lacks realism—for example, varying weather and terrain conditions, sufficient quantity or quality of test resources, and so forth—the test results may be invalid and decisionmakers may be deprived of important data needed to assess weapon system performance.

In the six weapon systems we examined,[1] we found that the usefulness of OT&E in estimating a weapon system's performance has been limited and continues to reflect the long-standing problems in test planning, test conduct, and the reporting of test and evaluation results. For example,

- test sites were not representative of operational environments,
- test objectives and evaluation criteria were not clearly defined, and
- test resources were either not available or not fully representative of threats to be encountered.

We recognize that operational testing cannot always be done in an environment that totally represents all operational conditions. For example, it is not practical or possible to use aircraft and missiles as targets that totally represent potential adversaries. Nevertheless, operational testing is important in the weapon systems acquisition process, and if properly planned, conducted, and reported, it can be crucial to identifying risk before a weapon system is produced and deployed.

In the following sections, we highlight the historical OT&E concerns and relate them to testing done on the selected weapon systems, and how they have impaired the operational effectiveness and suitability of the systems reviewed.

[1]The weapon systems included in this review began to be deployed in late 1983. The Director of OT&E was appointed in April 1985. Therefore, the situation described in this chapter did not occur under the auspices of the new Director. The situation described in this chapter can be viewed as a further expansion of the historical perspective noted in chapter 2 but can also be viewed as an agenda of items to be reviewed by the Director.

Maintenance Practices Not Fully Representative of Operational Environment

Lack of realism can be a severe shortcoming in the conduct of OT&E. As a result of this test limitation, OT&E cannot fully identify operational problems that will be encountered once weapons are fielded and exposed to their actual operational environment.

F/A-18 Aircraft

In assessing the F/A-18's operational suitability, the Navy used contractor support which hindered the realism of the test agency's OT&E results.

Using contractor support for the F/A-18, maintenance and logistic support was a test limitation in assessing its operational suitability. The contractor used equipment, procedures, transportation, stocking, and response systems not fully representative of fleet level maintenance and supply responsiveness. These maintenance and support limitations precluded a full evaluation of F/A-18 availability and limited the evaluation of logistic supportability to merely an assessment of trends.

More than 2 years after the activation of the three Marine F/A-18 squadrons, contractor support for maintaining the system is still being used extensively by Navy and Marine field units. Field officials said the need for contractor support has resulted from inadequate in-house capability which was attributed primarily to delivery delays of automatic test equipment and associated test program sets. One of the recommendations made in the OT&E report was to expedite the development of automatic equipment to allow speedy repair of avionics.

Intermediate maintenance is normally a base level responsibility and is done by designated activities in direct support of user organizations. Assigned work includes calibrating, repairing, or replacing damaged or unserviceable parts, components, or assemblies; modifying material; and providing technical assistance to user organizations.

As of October 1985, field units at the Navy and Marine Corps Air Stations were without adequate in-house repair capability, which degraded the mission readiness of the F/A-18. To minimize this potential operational effect, contractor operated avionics repair facilities were established at the Naval Air Stations at Lemoore, California, and Cecil Field, Florida. Another adverse effect of the limitations in assessing F/A-18

logistic support is that the inadequate in-house capability has resulted in increased spare parts usage.

In commenting on our draft report, DOD stated that the lack of maintenance support did indeed limit the operational evaluation of the F/A-18. DOD stated that it was not financially feasible to have a complete maintenance organization in place to support testing. Furthermore, DOD noted that the F/A-18 composite fleet squadron operational readiness exceeds established goals.

Clearer Test Objectives and Criteria Are Needed

The DOD policy on test and evaluation requires that

- meaningful critical issues, test objectives, and mission related evaluation criteria need to be established before tests begin and
- dependence on subjective judgment concerning system performance be minimized during testing.

Examples which illustrate these problems are presented below.

GLCM

Misunderstandings on the meaning of requirements and criteria have been common among users, developers, and testers. These misunderstandings are due, in part, to the lack of specificity and requirements and, in part, to the different perspectives and understandings of the users, developers, and testers. As a result, test personnel and decisionmakers have been confused in the evaluation and interpretation of test results.

The Air Force Operational Test and Evaluation Center has done a study to determine how to provide better linkage between user needs, requirements, specifications, and test objectives, as well as to provide a common base for terms, definitions, and assessment methodologies and criteria. This study showed that for GLCM, every major document such as the required operational capability, program management directive, and the TEMP required changes to achieve a common understanding. This study resulted in a process called the baseline correlation matrix, which compares the requirements, specifications, and test criteria to ensure these factors correlate and that a common understanding exists among users, developers, and testers.

F/A-18

The need for clearer criteria was cited by the Chief of Naval Operations in the fiscal year 1986 military posture and budget statement to the Congress. In this statement, the Chief of Naval Operations said that too frequently operational requirements and performance thresholds critical in doing test and evaluation are poorly defined and need improvement. Without clear objectives and criteria for evaluating mission capabilities, weapon systems may not be adequately tested or assessed. For example, criteria to evaluate the operational range of the F/A-18 in the attack configuration were not established before the start of OT&E. Specifically, the Navy's "requirements people" did not establish standards for combat radius (operational range). The Navy's evaluation process would have gone considerably more smoothly if clear operational thresholds had been specified before starting OT&E.

In a 1983 report on the operational evaluation of the F/A-18 weapon system, the Navy's Operational Test and Evaluation Force stated that the lack of specificity in defining operational thresholds affected the Navy's ability to evaluate the test results. Due to the lack of test criteria, information of this nature, normally available before production and deployment, was not available to decisionmakers.

In commenting on our draft report, DOD stated that it had issued instructions clarifying the performance thresholds to use in determining operational performance.

Lack of Test Resources

Test and evaluation policy requires that to the extent permitted by available resources and the need for realistic test environments, appropriate test instrumentation be used to provide quantitative data for system evaluation. The policy also requires the items to be sufficiently representative of the expected production items to ensure that valid assessments can be made of systems expected to be produced.

F/A-18

In performing OT&E to determine the operational effectiveness and suitability of the Navy's F/A-18 aircraft and its readiness for full fleet introduction, there were a number of significant test limitations pertaining to the availability of test resources which precluded a complete evaluation of the system. For example, the lack of clearance to fly and test certain items, and the nonavailability of key subsystems, such as radar warning receivers and defensive electronic countermeasures equipment, precluded testing the effect this equipment would have on the aircraft's survivability/vulnerability. Officials involved with the introduction of

the F/A-18 into Marine Corps air operations said the electronic warfare equipment is extremely important to the aircraft survivability and mission capability. Without this equipment, the F/A-18 cannot be deployed effectively.

Problems Surfaced During OT&E Were Not Adequately Reported to Decisionmakers

Our review of test reports on the deployed systems identified cases where the operational effects of adverse test results were not clearly communicated to decisionmakers. Further, DDT&E assessments of OT&E results did not always adequately present to decisionmakers, at critical milestones, system deficiencies and their operational effects.

The DOD policy requires that each service's operational test agency report directly to the military service chief:

- the results of its independent evaluation of a weapon system and
- the issues which have an adverse effect on the accomplishment of adequate OT&E.

Specific guidance on preparing test reports is included in the operating instructions of the service's test agencies and DDT&E. For example, instructions of the Navy OT&E agency for reporting test results state the writer should assume the report recipient has only general familiarity with the warfare area being discussed. With this assumption, "jargon" and the use of too many acronyms is to be avoided. Also, the instruction states that limitations are to be expressed so their significance is readily understood. Similarly, DDT&E reporting instructions state that its assessments of OT&E should, among other matters, discuss the effect of insufficient or unavailable test resources, and the effect of system thresholds not achieved—a function that is now performed by the Office of the Director of OT&E.

These policies require that report presentation be clear and in sufficient detail to help the reader understand, and include an assessment of significant test limitations and results which may adversely affect the operational capabilities of weapon systems. Comments from our report[2] on the Army's weapon system testing underscore the need for clarity and comprehensive reporting, noting that most Army decisionmakers expressed dissatisfaction with the reporting and scope of the evaluations. Many decisionmakers believed that information presented was

[2]The Army Needs More Comprehensive Evaluations to Make Effective Use of Its Weapon System Testing (GAO/NSIAD-84-40, Feb 24, 1984).

fragmented, sometimes contradictory, and usually not complete or convincing. The decisionmakers said they need a more thorough interpretation or evaluation of test results. Without comprehensive reporting of test results, the decisionmakers have had to rely extensively on data provided by system project managers and other system advocates. In responding to our report, DOD agreed that better and more thorough reporting was needed.

Examples which illustrate these problems are presented below.

MLRS

Logistic supportability, particularly ammunition resupply, was a critical issue in assessing MLRS operational effectiveness and suitability. OTEA's test report stated that the trailer did not perform satisfactorily and that it was incapable of operating on unimproved roads when carrying a full load of ammunition. OTEA's test report did not adequately describe the significance or effect of the problem with the ammunition resupply trailer—namely that it tipped over when carrying a full load of MLRS rocket pods at normal speeds. OTEA did not address this situation as an area of concern until after the weapon system was fielded.

MLRS units are unable to maintain the required amount of ammunition because the resupply vehicle, the heavy expanded mobility ammunition trailer, tips over when carrying the required load of four rocket launch pod containers. As an interim solution to the problem, the Army imposed an operational restriction allowing only half loads to be carried, hence reducing the overall MLRS ammunition resupply capability 50 percent below the design and operational requirement.

In 1985, the Army decided to suspend purchase of the trailer, creating a shortfall for support of future units. Army officials told us that there will be a gap in fielding units with trailers while awaiting the development, production, and delivery of a new trailer.

Bradley Fighting Vehicle

In December 1979, when the decision was made to begin full-scale production of the Bradley Fighting Vehicle, the limited test results indicated the vehicle's armor did not meet the ballistic protection requirements and therefore, was vulnerable to hostile fire. This information was not reported to key decisionmakers. Even though the system has been deployed, the vehicle's vulnerability is still a major concern as demonstrated by recent test results.

In February 1980,[3] we issued a report expressing concern about the vulnerability of the Bradley. We stated that the Army expected the Bradley to move more rapidly in a hostile environment with better armor and greater firepower than was available in the then current M113 weapon system. Our report emphasized that one of the intended advantages of the Bradley over the M113 was that it was to have greater survivability on the battlefield because of its supposedly superior armor protection.

We also stated that initial operational testing showed the armor to be "somewhat under the Army's stated ballistic protection requirements." We also pointed out that testing of the vehicle's armor had been very limited, and therefore, the test results were far from being conclusive—raising concern about the vehicle's vulnerability.

Over 6 years after production began on the Bradley, its vulnerability is still a concern to the Army. The system is undergoing vulnerability testing at an Army test facility. Testing was not completed until June 1986. However, we reported in February 1986[4] that in evaluating the results of the phase I tests, it should be recognized that (1) certain shots, which could have caused severe damage to the vehicle and crew, were avoided, (2) some of the most currant threat simulators were not used, and (3) the infantry version, which is more susceptible to greater numbers of casualties, was not tested.

DOD noted that the Bradley vulnerability testing is a developmental test issue and not under the purview of OTEA. Notwithstanding the nature of this testing, the Bradley has been in production since 1979 and its vulnerability continues to be a major operational concern. Therefore, OTEA should continue to monitor the testing of Bradley and should include a discussion of Bradley's vulnerability in its continuing assessment reports on the Bradley.

Bradley's Night Sight

Information provided to the Congress and key DOD decisionmakers did not disclose that the Bradley's ability to fight was degraded because of the ineffectiveness of its night sight. Using organization officials said the Bradley has not been effective in night operations because of problems with the integrated sight unit subsystem which was designed to

[3]Concerns About the Army's Infantry Fighting Vehicle (GAO/PSAD-80-27, Feb 5, 1980)

[4]Concerns about the Army's Vulnerability Testing (GAO/NSIAD-86-67, Feb 14, 1986)

provide a capability to operate at night or in adverse weather for continuous combat operations. In December 1979, before the beginning of full-scale production decision, OTEA reported that the Bradley was not "suitable for issuance" unless the problems with the integrated sight unit (i.e., picture degradation and image flutter) were corrected and corrections verified. However, in 1980 when the full-scale production was made, the information provided to key decisionmakers did not indicate that the system would be seriously degraded because of the ineffectiveness of the night sight.

The Bradley's predecessor system, the M113, also had limited night operation capability. This operational shortcoming with the M113 was used by the Army to justify the need for the Bradley. The Army stated that the M113 had limited night operation capability

> "whereas the enemy threat vehicles and combined arms teams do. Due to the propensity of the enemy to operate at night, it is necessary that the [Bradley] have the capability for full operation at night "

As part of the decisionmaking process, DDT&E provided a system assessment to DOD decisionmakers. However, DDT&E's 1980 assessment of the Bradley did not address the significance of the problem associated with the integrated sight unit. The assessment merely stated that during OTEA's testing "most all maintenance on the integrated sight [unit], the most complex equipment in the vehicle, was done by the contractor."

A modification program has been instituted to improve the integrated sight unit. In 1983, several years after full-scale production began, DOD noted that OTEA identified a problem with the Bradley's night sight. DOD noted that the Bradley's night sight was caused by "out-of-tolerance-input voltages" from the power control unit. Improved power control units were recently introduced on production models. DOD stated that fielded systems demonstrate acceptable effectiveness. However, several of the fielded unit officials continued to express concern about the sight unit's effectiveness during our field visits.

F/A-18, EF-111A, and HARM

We also found that DDT&E did not identify the operational effect of adverse test results on the F/A-18 and EF-111A aircraft and the HARM missile. DDT&E's system assessments, while reporting system shortcomings and deficiencies, did not explain or quantitatively identify the potential adverse operational impact of reported deficiencies. We noted, however, that their operating instruction requires DDT&E to provide a

detailed discussion of OT&E results and the effect of insufficient or unavailable test resources.

Conclusions

DOD's operational test policy calls for realism and it should continue to strive for that goal to help minimize acquisition risks and to increase the likelihood that weapon systems will work as intended when deployed. DOD can accomplish this by (1) requiring the establishment of adequate test objectives and evaluation criteria, (2) making representative test resources available, and (3) using test scenarios representative of geographic and weather environments. Where such circumstances cannot be prevented or controlled, information about their potential effects on operational testing needs to be highlighted for decisionmakers.

DOD decisionmakers rely on the results of OT&E to estimate weapon performance, but in several of the weapon systems we reviewed, problems surfaced during OT&E that were not adequately reported to them. We believe that increased emphasis should be placed in ensuring that (1) test reports contain current, complete, and accurate data and (2) potential operational effects of test limitations be fully disclosed.

Historical problems involving test planning and conduct surfaced during the acquisition process for several of the deployed weapon systems we reviewed—which can result in the fielding of a weapon system with performance shortcomings affecting mission effectiveness that will require redesign and modification.

Recommendations

We recommend that the Secretary of Defense enforce the requirement that operational testing:

- Be done in as operationally realistic an environment as possible.
- Not be performed unless there are clearly stated test objectives and criteria.

We also recommend that the Secretary of Defense improve the usefulness of OT&E results by the services' OT&E agencies by requiring the agencies to

- state whether OT&E demonstrated that the system met operational requirements,
- discuss the operational effect of significant test limitations and adverse test results on system performance, and

- clearly state whether a system is operationally suitable and effective.

Agency Comments and Our Evaluation

DOD agreed that the Secretary of Defense should enforce the requirement that operational testing be done in as realistic combat environment as possible. DOD stated that the Director of OT&E, working with the military services and the Deputy Under Secretary of Defense for Research and Engineering (Test and Evaluation), is pursuing a number of initiatives to enhance realism.

DOD concurred with our recommendation that the Secretary of Defense enforce the requirement that operational testing not be performed unless test objectives and criteria are clearly stated. It noted that several actions were being taken and cited that the new TEMP Guidelines manual, which is planned to be published during fiscal year 1987, sets forth requirements for test objectives and criteria.

The draft of our report proposed that the OT&E agencies should be required to recommend whether a system is ready for production. DOD disagreed that the Secretary of Defense should require the OT&E agencies to recommend whether systems are ready for production. DOD stated that the military services OT&E agencies should generally make a recommendation on a weapon system's operational effectiveness and suitability. We recognize that the final decision for advancing a system to production rests with the Secretary of Defense, and have modified our recommendation accordingly to recognize this point.

Office of the Director of OT&E

The usefulness of OT&E in estimating the performance of a weapon system before the start of low-rate initial production may often not be fully realized. Continual trade-offs are made between the extent of operational testing to be performed to identify potential operational shortcomings versus the delays that may occur in fielding a system. In a recent attempt to improve the usefulness of OT&E in the acquisition of major weapon systems, DOD created the Office of the Director of OT&E in 1984, pursuant to Public Law 98-94.

The Director's most recent annual report to the Congress states that it had a significant positive effect on the acquisition process. Specific accomplishments identified by the Director included:

- Reviewing the military services' test and evaluation master plans to ensure that operational test concerns are incorporated early in the life of all new programs.
- Doing an ongoing detailed survey of test resources and capabilities.

We have not determined whether the Director's actions to ensure early planning for OT&E, as well as recognizing a problem associated with the lack of test resources, will be successful in eliminating the long-standing problems described elsewhere in this report. Several members of the Congress recently requested us to evaluate the effectiveness of the Office of the Director of OT&E.

Waiting for a production representative item and operationally testing it just before production is frequently not done because of the lack of production representative items, especially for highly concurrent programs. It has been alleged that doing an operational test on production representative items (i.e., prototypes) will exclude new technology and extend the acquisition process—and may ultimately result in what the Director calls "government-certified antiques."

We recognize, as stated in our 1985 report,[1] that there may be instances where there is a need to begin initial production without the benefit of OT&E. However, when OT&E is done before initial production, information is available on potential shortcomings that would not be forseen through developmental testing. Further, OT&E results permit decisionmakers to assess whether potentially costly modifications are needed. In some

[1] Production of Some Major Weapon Systems Began With Only Limited Operational Test and Evaluation Results (GAO/NSIAD-85-68, June 19, 1985)

cases, conducting timely and effective OT&E can lead to program cancellations. Thus, in balancing the need to make informed decisions and to shorten the acquisition cycle, decisionmakers must consider if the U.S. government should incur significant production expenditures without knowledge of whether a weapon system can effectively perform its mission.

The President's Blue Ribbon Commission on Defense Management has also emphasized the value of developing and using more prototype hardware and testing it under operational conditions before final design approval or authorization for production.

We plan to develop this issue during our current review of the effectiveness of the Office of the Director of OT&E to determine the extent to which a phase of OT&E before the start of low-rate initial production is being planned.

Review of Selected Reports Issued Since 1970 Dealing With OT&E

Our Reports	Problem Areas Discussed in Reports[a]
Generic	
Adverse Effects of Large-scale Production of Major Weapons Before Completion of Development and Testing (B-163058, Nov 19, 1970)	B,D,E,F,G
The Importance of Testing and Evaluation in the Acquisition Process for Major Weapon Systems (B-163058, Aug 7, 1972)	A,C,D,E,F,G,H
Review of Testing and Evaluation Policies and Procedures (B-163058, Apr 18, 1974)	E
Review of the Adequacy of Department of Defense Test Resources (GAO/PSAD-75-84, Apr 30, 1975)	B,C,D,H
Effectiveness of Testing of Selected Major Weapon Systems (GAO/PSAD-75-74, June 4, 1975)	A,B,C,D,E,H
Navy Operational Test and Evaluation—A Valuable Tool Not Fully Utilized (GAO/PSAD-78-77, Mar 29, 1978)	B,C,D,F,G,H
Operational Testing of Air Force Systems Requires Several Improvements (GAO/PSAD-78-102, June 2, 1978)	A,B,C,D,H
Department of Defense's Conduct of Operational Test and Evaluation (OT&E) of Foreign Built Weapon Systems (B-163058, GAO/PSAD-78-131, July 25, 1978)	A,D,F,G
Practices and Procedures for Follow-On Operational Testing and Evaluation of Weapon Systems by the Military Services (B-163058, GAO/PSAD-79-1, Oct 19, 1978)	A,D,H
Army Operational Test and Evaluation Needs Improvements (GAO/C-PSAD-80-2, Nov 13, 1979)	A,B,C,D,E,F,H
Effectiveness of U S Forces Can Be Increased Through Improved Weapon System Design (GAO/PSAD-81-17, Jan 29, 1981)	C,F
Be ter Planning and Management of Threat Simulators and Aerial Targets Is Crucial to Effective Weapon Systems Performance (GAO/MASAD-83-27, June 23, 1983)	A,B,C,D,F,G,H
How Well Do the Military Services Perform Jointly in Combat? DOD's Joint Test-and-Evaluation Program Provides Few Credible Answers (GAO/PEMD-84-3, Feb 22, 1984)	A,C,E,H
The Army Needs More Comprehensive Evaluations to Make Effective Use of Its Weapon System Testing (GAO/NSIAD-84-40, Feb 24, 1984)	A,C,E,G,H
Production of Some Major Weapon Systems Began With Only Limited Operational Test and Evaluation Results (GAO/NSIAD-85-68, June 19, 1985)	B,C,D,F,G
Individual weapon systems	
Need for Additional Test and Evaluation on the Major Caliber Lightweight Gun (GAO/PSAD-77-4, Nov 5, 1976)	B,D,E,H
M1 Tank's Reliability Is Still Uncertain (GAO/PSAD-80-20, Jan 29, 1980)	C,D,H
The Joint Tactical Information Distribution System—How Important Is It? (GAO/PSAD-80-22, Jan 30, 1980)	B,C,H
Inherent Risk in the Army's Acquisition Strategy Demands Particular Caution in Evaluating the Division Air Defense Gun System's Production Readiness (GAO/C-PSAD-80-9, Jan 31, 1980)	A,H
Concerns About the Army's Infantry Fighting Vehicle Program (B-196877, GAO/PSAD-80-27, Feb 5, 1980)	C,H
Status of the CAPTOR Mine Warfare Program (GAO/C-PSAD-80-11, Feb 11, 1980)	A,H
F/A-18 Naval Strike Fighter· Its Effectiveness Is Uncertain (GAO/PSAD-80-24, Feb 14, 1980)	B,D,F
Current Difficulties in Effectively Deploying Multiple Launch Rocket System Render Program's Concurrency Questionable (GAO/C-PSAD-80-20, Feb 26, 1980)	A,H
Cruise Missiles Status and Issues as They Near Production (GAO/C-PSAD-80-19, Feb 28, 1980)	B,C,D,H
Future Procurements of Army's Copperhead Projectile Should Be Contingent on Improvements in Performance and Reliability (GAO/C-PSAD-81-4, Nov 13, 1980)	D,F
Opportunities for Improving Management of the Navy's AEGIS Cruiser Program (GAO/C-MASAD-81-8, Feb 19, 1981)	E,F

Our Reports	Problem Areas Discussed in Reports[a]
Progress and Problems of the Advanced Medium Range Air-To-Air Missile Program (GAO/C-MASAD-81-6, Feb 23, 1981)	A,H
Review of the High Speed Anti-Radiation Missile (HARM) Program (GAO/C-MASAD-81-7, Feb 28, 1981)	D,H
Some Land Attack Cruise Missiles Acquisition Programs Need to Be Slowed Down (GAO/C-MASAD-81-9, Feb. 28, 1981)	A,B,C,D,E,H
Most Critical Testing Still Lies Ahead for Missiles in Theater Nuclear Modernization (GAO/MASAD-81-15, Mar 2, 1981)	D,H
The Army's Advanced Attack Helicopter Is Not Ready for Production (B-201273, GAO/MASAD-82-8, Dec 1, 1981)	B,C,D,H
Navy's F/A-18 Expected to be an Effective Performer but Problems Still Face the Program (GAO/MASAD-82-20, Feb 26, 1982)	C
Air Launched Cruise Missile Shows Promise but Problems Could Result in Operational Limitations (GAO/C-MASAD-82-13, Feb 26, 1982)	C,D,F,G
Defense Plans to Deploy Some Cruise Missiles Before They Are Ready (GAO/C-MASAD-82-15, Feb 26, 1982)	B,C,F,G
The Navy's Landing Craft Air Cushion—Uncertainty Over How It Will Be Used With Amphibious Forces (GAO/C-MASAD-82-9, Feb 26, 1982)	A,H
Test and Evaluations Still in Progress Should Indicate Division Air Defense Gun's Potential Effectiveness (GAO/C-MASAD-82-7, Feb 26, 1982)	A,H
Progress of the Light Armored Vehicle Program Should Be Closely Monitored (B-208521, GAO/MASAD-82-41, Aug 10, 1982)	D,H
Results of Production Testing Should Be Considered Before Increasing Patriot's Production (GAO/C-MASAD-83-7, Jan 26, 1983)	D,E,F,G
The Army's AH-64 Helicopter and Hellfire Missile Retain Risks as They Enter Production (GAO/C-MASAD-83-9, Jan 26, 1983	B,H
The Army Should Confirm Sergeant York Air Defense Gun's Reliability and Maintainability Before Exercising Next Production Option, Radar System Should Be Reevaluated (GAO/C-MASAD-83-14, Mar 15, 1983)	A,H
Status of the CG-47 Cruiser and DDG-51 Destroyer Shipbuilding Programs (GAO/C-MASAD-83-11, Feb 22, 1983)	C
Acquisition of the Over-The-Horizon Backscatter Radar System Should Be Reevaluated (GAO/C-MASAD-83-14, Mar 15, 1983)	D,H
The B-1 Bomber Program—A New Start (B-206613, GAO/MASAD-83-21, Apr 13, 1983)	A,C,D
Air Force and Navy Trainer Aircraft Acquisition Programs (GAO/MASAD-83-22, July 5, 1983)	D,H
Department of the Army's Competitive Test and Evaluation of Alternative Light Antiarmor Weapons (B-213544, GAO/NSIAD-84-57, Feb 14, 1984)	E
Results of Forthcoming Critical Tests Are Needed to Confirm Army Remotely Piloted Vehicle's Readiness for Production (GAO/NSIAD-84-72, Apr 4, 1984)	D,E
Status of the Peacekeeper (MX) Weapon System (GAO/NSIAD-84-112, May 9, 1984)	D,H
Army's Decision to Begin Production of the High Mobility Multipurpose Wheeled Vehicle Was Premature (GAO/NSIAD-84-136, June 12, 1984)	D,E,F
Army Has the Opportunity to Recompete DAS3 Purchases and Improve Automated Battlefield Support (B-216005, GAO/IMTEC-84-20, Sept 28, 1984)	D,G
Evaluation of Army's Mobile Subscriber Equipment Program (GAO/NSIAD-85-117, July 16, 1985)	D
DOD and other reports	
Report to the President and the Secretary of Defense on the Department of Defense by the Blue Ribbon Defense Panel, 1970	A,B,C,D,E

Our Reports	Problem Areas Discussed in Reports[a]
Test and Evaluation Considerations for the Naval Air Systems Command Program Manager, Kimble, Charles Donald, Commander, USN, DSMC,[b] Fort Belvoir, Va., May 1975	A,C,D,E,G,H
Testability Considerations for the Conceptual and Early Validation Phases of a Missile Weapon System, Smith, Richard Clark, DSMC, Fort Belvoir, Va., 1975	B,C,F,G,H
The Navy Test and Evaluation Process in Major Systems Acquisition, Hoivik, Thomas Harry, Lt. Commander, USN, DSMC, Fort Belvoir, Va., 1976	A,B,C,D,E,G
Logistical Ambushes in Operational Testing, Lytle, William L., LTC , USA, DSMC, Fort Belvoir, Va., 1977	B,F,G
Fielding Army Weapon Systems. Experiences and Lessons Learned, Lincoln, James B., LTC , USA, Industrial College of the Armed Forces, 1980	B,C,D,E,G,H
Lessons Learned Multiple Launch Rocket System, Murrow, Garcia E , et al , DSMC, and Information Spectrum, Inc., Fort Belvoir, Va , 1980	B,C,D
Report on the Review of DOD Systems Independent Test and Evaluation Program - Navy (U), by the Inspector General, No 81-100, May 29, 1981	B,C,D,E
Systems Independent Test and Evaluation Program - Air Force (U), by the Inspector General, No 83-031, Nov. 23, 1982	A,B,C,D,E,H
Lessons Learned Advanced Attack Helicopter, Department of Research and Information, DSMC, Ft. Belvoir, Va., 1983	A,B,D,F
MLRS: A Case Study of Manpower, Personnel and Training Requirements Determination, O'Conner, F E., et. al , Information Spectrum, Inc , Arlington, Va , 1984	B,C,D,E,H
Reverse Engineering of the Multiple Launch Rocket System (MLRS) Human Factors, Manpower, Personnel and Training in the Weapon System Acquisition Process, Bryden, William D , U S Army Research Institute for the Behavioral and Social Scientist, June 1984	A,B,C,F
The Test and Evaluation Evolution (Relative to the U.S Air Force), Bryden, William D , Jr , Major, DSMC, Ft. Belvoir, Va., 1984	A,D,E

[a]See table 1 (chapter 2) for description of problem areas

[b]Defense Systems Management College

Description of Weapon Systems Reviewed

Air Force Systems

GLCM

The GLCM is a tactical theater nuclear missile system designed to increase the North Atlantic Treaty Organization's capability to strike targets in the Soviet Union. The missile is guided by an inertial navigation system and a system of terrain contour matching, for making guidance corrections. As of April 1985, the total program cost was estimated at about $3.8 billion.

EF-111A Aircraft Tactical Jamming System

The EF-111A aircraft is a modified F-111A, a twin-engine, swing-wing fighter/bomber aircraft. The mission of the EF-111A is to provide support jamming for tactical aircraft operations. The delivery of the 42nd and final EF-111A is expected to take place before the end of 1986, at a total estimated program cost of $1.8 billion.

Army Systems

Bradley Fighting Vehicle

The Bradley system includes the infantry fighting vehicle, or the M-2, and the cavalry fighting vehicle, or the M-3. The M-2 is designed to carry a nine-person infantry squad, which includes a driver, a commander, and a gunner. The M-2 has six firing ports, positioned along the sides and back of the vehicle, through which the six operators in the rear of the vehicle can fire their weapons. The squad can, therefore, fight from within the vehicle, as well as dismounted.

The M-3, which carries five troops, serves the armored cavalry units as a scout vehicle for purposes of reconnaissance and security missions, using its firepower mainly to defeat the lightly armored vehicles ahead of the enemy's main tank force. Both versions have a 25-mm. chain gun, which can use either kinetic energy rounds or high explosive rounds; a TOW antitank guided missile launcher; and a coaxial machine gun. Both versions are protected with aluminum armor, which can withstand up to 14.5-mm. caliber ammunition. The Bradley began production in 1980, and first deployment was in December 1983. The current estimated total program acquisition cost is $10.8 billion.

MLRS

MLRS is a tracked, self-propelled launcher loader with disposable launch pads and fire control equipment. MLRS is intended to defeat enemy field artillery and air defense systems, command and control sites, logistic complexes, and personnel targets. MLRS began production in 1980, and first deployment was in April 1983. The current estimated total program acquisition cost is $4.3 billion.

Navy Systems

F/A-18 Aircraft

The F/A-18 naval strike fighter is a twin-engine, single-pilot aircraft designed to replace the F-4 and the A-7 and perform fighter and attack missions for the Navy and Marine Corps. The Navy plans to purchase 1,377 F/A-18s at an estimated cost of $40 billion. The Navy began deploying the F/A-18 aboard carriers in 1985.

HARM

HARM is a guided missile system designed to help tactical aircraft penetrate enemy defenses by destroying or suppressing the radar used to direct surface-to-air missiles and anti-aircraft guns. The joint Navy and Air Force HARM program will provide missiles for use on the F/A-18 and other Navy and Air Force planes. Navy and Air Force plans include the purchase of 16,189 missiles at a total estimated program cost of $5.2 billion.

Comments From the Director, OT&E

OPERATIONAL TEST AND EVALUATION

OFFICE OF THE SECRETARY OF DEFENSE

WASHINGTON D C 20301 1700

29 October 1986

Mr. Frank C. Conahan
Director, National Security and International
Affairs Division
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Conahan:

(U) This is the Department of Defense (DoD) response to the General Accounting Office (GAO) draft report entitled "Operational Test and Evaluation Can Contribute More to Decisionmaking," dated August 26, 1986, OSD Case 7115, GAO Code 396203.

(U) The Department agrees for the most part with the findings and recommendations of the draft report. The DoD would like to point out, however, that the period covered by the report predates the establishment (in 1984) of the Office of the Director, Operational Test and Evaluation (DOT&E). Since that time, the DoD and the Services have initiated several efforts--for example, the establishment of the DoD Test and Evaluation Council (DTEC), the revision of DoD Directive 5000.3, and the initiation of the Air Force Baseline Correlation Matrix process. These initiatives will significantly improve the DoD's planning for, conduct of, and reporting on operational test and evaluation.

(U) The enclosed comments address the findings and recommenations in greater detail. Several technical corrections have been separately provided to the GAO staff. The DoD appreciates the opportunity to comment on the draft report.

Sincerely,

JOHN E. KRINGS
Director

Note Portions of paragaraph were deleted to remove classified information.

DEPARTMENT OF DEFENSE COMMENTS
ON GAO DRAFT REPORT
(GAO CODE NO. 396203) - OSD CASE 7115

DATED AUGUST 26, 1986

"OPERATIONAL TEST AND EVALUATION CAN
CONTRIBUTE MORE TO DECISIONMAKING"

* * * * *

FINDINGS

(U) FINDING A: Historical Operational Test and Evaluation (OT&E) Problems and Concerns. The GAO analyzed 63 reports issued since 1970 and identified numerous instances involving OT&E problems. The GAO found the problems to include (1) incomplete and/or unclear OT&E planning, (2) limited or unavailable test resources, (3) unrealistic testing because the operational environment was not adequately replicated, (4) inadequate reporting of OT&E results, and (5) production approval before OT&E was completed. The GAO reported that over the years the DoD has implemented, or is planning to implement, many of the recommendations that have been made to improve OT&E. According to the GAO, some of the more significant actions are: (1) the establishment of an OT&E office within the Office of the Secretary of Defense in 1984, (2) the establishment of independent OT&E offices within the Services, and (3) increased emphasis by the DoD since 1980 on the importance of OT&E. While acknowledging these actions, the GAO nevertheless concluded that past recommendations need to be fully implemented to improve OT&E. (The GAO noted that it is currently evaluating the effectiveness of the DoD Office of the Director, OT&E, under a separate GAO review.) (p. 2, pp. 14-21, pp. 37-38/GAO Draft Report)

(U) DoD POSITION: Concur. The DoD agrees that these problems have existed historically. However, in addition to the improvements acknowledged by the GAO as implemented and planned by DoD, the Services have initiated a number of efforts to improve operational testing (OT), and the Director, Operational Test and Evaluation (DOT&E) has done likewise. For example: (1) The Air Force is instituting a process and procedure (the Baseline Correlation Matrix) that aligns requirements, specifications, and test criteria to preclude any disconnects in test planning and between decision makers and other participants. (2) The Army has put in place its Continuous Comprehensive Evaluation (C2E) program, requiring its Operational Test and Evaluation Agency (OTEA) to assess a system's status throughout the materiel acquisition process, reporting periodically on its operational effectiveness and suitability based on all evidence to date, not just operational test results. (3) The Navy has promulgated OT&E guidance

instructions setting forth requirements designed to ensure adequacy and improve realism of OT&E. Among these is Chief of Naval Operations (OPNAV) Instruction 3960.10 (August 1983), which currently is being updated and which, among other things, establishes requirements for initial operational test and evaluation (IOT&E) before low-rate initial production (LRIP) and completion of the final phase of IOT&E as a prerequisite for approval of full-rate production. (4) The authority granted to the DOT&E in November 1983 under 10 USC 136a to review and approve the adequacy of OT plans is being applied to good effect. In large measure as a result of the DOT&E's active participation, the newly revised DoD Directive 5000.3, "Test and Evaluation" (12 March 1986), lays particular emphasis on early OT involvement in the acquisition process, the requirement for OT&E before production decisions, long-range test resource planning, and realistic OT.

(U) Concerning the matter of production approvals before OT&E has been completed, 10 USC 136a requires the DOT&E to report to the Secretary of Defense and the Armed Services and Appropriations committees of the House and Senate on the adequacy of testing and the operational effectiveness and suitability of weapon systems before a decision to proceed beyond LRIP may be approved. This authority and the office of the DOT&E were not in place during the period covered by the GAO report. The current GAO review of the effectiveness of the DOT&E office should provide ample evidence that the exercise of this authority is having a positive impact.

(U) FINDING B: Impediments to Useful OT&E. The GAO identified three primary factors that hinder OT&E:

-- the belief that a weapon system must go into production and be deployed regardless of its readiness for production;

-- compressed acquisition cycles that limit the time available to perform OT&E before production; and

-- the diffusion of the responsibility for correction of deficiencies identified during OT&E.

(U) The GAO pointed out that the DoD acquisiton policy encourages concurrent development of major weapon systems because of the importance of shortening the acquisition cycle, but does not address how a shortened acquisition cycle can provide for planned OT&E results before production start-up. While recognizing the potential benefits of concurrency, the GAO concluded safeguards to minimize risk should be built into a program. The GAO further concluded that, at the very least, such safeguards should include one separate phase of OT&E and completion of any planned OT&E before production start-up. The GAO also concluded that weapon systems with planned concurrency should allow for special attention to OT&E so performance risks

resulting from a shortened acquisition time do not affect the planned deployment date. With regard to the diffusion of management responsibility, the GAO found that performance shortcomings identified during OT&E are frequently not corrected before production and deployment. The GAO concluded that this could be indicative of a lack of effective planning. The GAO further concluded that without sufficient management emphasis and accountability to ensure that test resources are sufficient to meet test goals and objectives, and deficiencies identified during OT&E are corrected before production and deployment, OT&E results will continue to be of limited usefulness. (pp. 2-3, pp. 21-24/GAO Draft Report)

(U) DoD POSITION: Concur. As discussed in the DoD position on Finding A, these problems are being appropriately addressed under recent initiatives by the Office of the Secretary of Defense (OSD) and the Services.

(U) FINDING C: Test Sites During OT&E Not Fully Representative of Operational Environment. For two of the weapon systems examined, the GAO found that the OT&E test sites were not fully representative of the operational environment. The GAO concluded this is one factor that limits the usefulness of OT&E in estimating a weapon system's performance. The GAO cited the Army's Multiple Launch Rocket System (MLRS), which had a significant problem with water intrusion. According to the GAO, because the OT&E was conducted in a desert environment, as a result of the test limitation this problem did not surface during OT&E. The second example cited by the GAO involved the Navy's F/A-18 Aircraft. According to the GAO, contractor support that was not fully representative of fleet intermediate level support was used during F/A-18 OT&E, which precluded a full evaluation of F/A-18 availability and logistic supportability. The GAO observed that as a result of this limitation, the GAO found that in-house intermediate level support developed for F/A-18 units has been inadequate, requiring the extensive use of contractor support. The GAO concluded that this test approach has required more spare parts than anticpated and degraded mission readiness. The GAO generally concluded that without testing realism, OT&E cannot fully identify operational problems that will be encountered once weapons are fielded and exposed to their actual operational environment. (pp. 3-4, 26, and 51-53, GAO Draft Report)

(U) DoD POSITION: Partially concur. Although it is true that operational test sites often do not fully represent expected operational environments, the MLRS OT is not an appropriate example of the limitations imposed on OT&E by such shortcomings. MLRS OT III was conducted at Fort Bliss, Texas, and White Sands Missile Range because a live-fire operational test was required and no other areas had been cleared for MLRS live fire.

GAO note: An earlier draft of this report contained a discussion of the test environment for the MLRS. We have deleted this discussion based on DOD comments.

(U) Moreover, the MLRS example is not entirely accurate. During the two 12-day field training exercises conducted during MLRS OT III, there were 11 days during which there was rain, ranging from a trace to 1.58 inches. The OT III test report cited 18 MLRS electronic unit failures (without, however, specifying moisture as the cause). Based on the OT III results, the materiel developer made modifications that corrected the problem.

(U) With regard to the F/A-18 example,

D E L E T E D

the testing was designed to be as close as possible to fleet representative given the funding constraints faced by the testers. It is worth noting that, today, the F/A-18 composite fleet squadron operational readiness rate exceeds Chief of Naval Operations readiness goals, and organic capability passed the ninetieth percentile in June 1986.

(U) The Deputy Secretary of Defense has established the DoD Test and Evaluation Council (DTEC), chaired by the DOT&E, to address improvement of T&E resources investment and management. Of the two high-priority issues currently being addressed by the DTEC, one is the development of a capability to do more realistic OT&E, to include the ability conduct OT in locations closely representative of operational environments. Results of this effort will be reported to the Deputy Secretary by November 1, 1986.

(U) FINDING D: OT&E Objectives and Criteria Not Clearly Defined. The GAO found that inadequate OT&E test objectives and criteria existed for two of the weapon systems examined, which the GAO cited as a second factor limiting OT&E usefulness. According to the GAO, DoD policy requires that (1) meaningful critical issues, test objectives, and evaluation criteria be established prior to testing, and (2) that dependence on subjective judgment be minimized. The GAO found, however, that criteria to evaluate the operational range of the F/A-18 in the attack configuration was not established before the start of OT&E, which adversely impacted the Navy's ability to assess the aircraft's performance capabilities. The GAO also found that misunderstandings on the meaning of requirements and criteria have been common among users, developers and testers of the Air Force Ground Launched Cruise Missile (GLCM), which has in turn caused confusion in interpreting test results. The GAO also noted that the Air Force has conducted a study to provide better linkage between requirements and test objectives, which resulted in development

Of a process to improve this linkage and ensure a common understanding among users, developers and testers. Overall, the GAO concluded that without clear objectives and criteria for evaluating mission capabilities, weapon systems may not be adequately tested or assessed. (p. 25, pp. 27-28, p. 47/GAO Draft Report)

(U) DoD POSITION: Partially concur. As previously discussed in the DoD position on Finding A, the period covered by the GAO predates the establishment of the Office of the DOT&E. Under the requirements of the newly revised DoDD 5000.3, clear, objective OT issues and criteria must be included in each Test and Evaluation Master Plan (TEMP), which must be reviewed and approved by both the DOT&E and the Deputy Under Secretary of Defense for Research and Engineering (Test and Evaluation) (DUSDR&E(T&E)). Moreover, each TEMP must be updated annually and the updates reviewed and approved by both the DOT&E and the DUSDR&E(T&E). The OT issues and criteria in current, approved TEMPS form the basis for OT test plans, which also must be approved by the DOT&E before OT may commence.

(U) With regard to the F/A-18 example,

D E L E T E D

in 1978, the DCP changed the Strike Mission Radius definition to indicate "Design Mission Profile using the fleet oriented rules." This 1978 DCP was the first time operational requirements were used in place of technical thresholds. In 1982 the Department of Defense Instruction on Major System Acquisition Procedures (Number 5000.2) for the first time specified the type of thresholds to be used by OT&E to determine performance--operational. Essentially, Commander, Operational Test and Evaluation Force (COMOPTEVFOR), the Navy's OTA, had been testing the F/A-18's operational performance well before the DoD instruction change. The problem of clearly defining the requirements for test objectives and criteria has been rectified by DoD and Navy instructions specifying operational thresholds for OT&E activities and as noted above.

(U) With regard to GLCM, many test objectives and criteria were well defined. However, there did exist some differences of opinion among participants as to the required level of detail and scope of testing. As was noted above in the DoD position on Finding A, the Air Force is initiating a procedure and process (the Baseline Correlation Matrix) which aligns requirements, specifications, and test criteria. It is expected that this new approach will alleviate confusion over criteria in the future.

(U) FINDING E: OT&E Test Resources Not Available or Not Fully Representative of Threats. According to the GAO, DoD policy requires that, to the extent practical, appropriate test instrumentation be used to provide data, and the test items be sufficiently representative of production items to insure valid assessments. For three weapon systems it reviewed, however, the GAO found problems with the OT&E test resources. The GAO cited, for example, that F/A-18 OT&E limitations included the lack of clearance to fly and test certain items, and the nonavailability of key subsystems such as radar warning receivers and defensive electronic countermeasures equipment. The GAO also found limitations during OT&E of the Air Force EF-111A Aircraft caused by the limited availability or quality of electronic warfare test resources. Even though the OT&E reports concluded that the overall operational effectiveness and suitability were satisfactory, the GAO reported EF-111A users identified performance limitations, that may reduce mission effectiveness, and testing limitations, which raise questions about the EF-111A's jamming capabilities. The GAO further found that the lack of adequate OT&E resources for the Navy's High Speed Anti-Radiation Missile (HARM) limited the evaluation of the missile's self-protect effectiveness and performance in a dense environment. The GAO pointed out that, although the OT&E report concluded HARM was potentially operationally effective and suitable, problems in the self-protect mode have been encountered since HARM was deployed. The GAO concluded the lack of adequate test resources is a third factor that has limited the usefulness of OT&E for recently deployed systems. (pp. 25, 29-30 and 47-51, GAO Draft Report)

(U) DoD POSITION: Partially concur. Shortfalls in test resources--with particular emphasis on threat-representative simulators and targets--are being addressed at both the OSD and Service levels. For example: (1) The DTEC has been established and will make recommendations to the Deputy Secretary of Defense by November 1, 1986, concerning improvement of space systems test and OT&E test capabilities. (2) At the direction of the Secretary of Defense, the DOT&E has carried out a study of DoD OT&E capabilities, techniques, and management and reported his findings and recommendations to the Secretary on September 24, 1986. (3) In June 1986 the Army formed the Army Instrumentation Development Council to manage, review, and prioritize test instrumentation, threat simulator, and test target development.

(U) With respect to the EF-111A,

D E L E T E D

operational flight testing is not permitted against unvalidated threats.

GAO Note: An earlier draft of this report contained a discussion of the EF-111A jamming capabilities. This discussion was deleted because it was classified.

(U) Concerning the HARM missile example,

D E L E T E D

recent actual combat experience with the HARM system in a very dense threat environment over Libya has demonstrated the operational effectiveness of the system.

(U) FINDING F: OT&E Problems Not Adequately Reported. According to the GAO, DoD policy requires each Service to report to decisionmakers the results of its independent evaluation and issues that adversely impact accomplishment of adequate OT&E. The GAO found, however, examples of inadequate reporting for five of the weapon systems reviewed. In the case of the MLRS, the GAO found that the OTEA report did not adequately describe the significance or impact of the problem with the ammunition resupply trailer. The GAO also found that information indicating problems with the vulnerability and night sight effectiveness of the Army's Bradley Fighting Vehicle was not adequately reported by either OTEA or the Director, Defense Test and Evaluation (DDT&E). In addition, the GAO found that the DDT&E did not identify the operational effect of adverse test results on three other systems--the F/A-18, EF-111A, and HARM--even though instructions require a detailed discussion of OT&E results. The GAO concluded that since decisionmakers rely on OT&E results to estimate weapon performance, increased emphasis should be placed on ensuring (1) test reports contain current, complete and accurate data, and (2) potential operational impacts of test limitations are fully disclosed. (p. 2, pp. 25-26, pp. 30-36/GAO Draft Report)

(U) DoD POSITION: Partially concur. As discussed in the DoD position on Finding A, initiatives have been undertaken that have and will continue to improve OT&E reporting. The advent of OSD-level oversight of OT&E reporting by the DOT&E, together with such Service initiatives as the Army's C2E program mentioned above in the DoD position on Finding A, has effectively underscored the requirement for clear, complete OT&E reporting to decision makers.

(U) With respect to the GAO's contention that the Army Operational Test and Evaluation Agency (OTEA) did not adequately describe problems with the MLRS ammunition resupply trailer, OTEA did identify these problems in its test report and its independent evaluation report on MLRS OT III. It is true that neither report specifically stated that the resupply trailer tipped over when carrying a full load of rocket pods and traveling at normal speeds. However, they did state that the trailer did not perform satisfactorily during the test and that it was incapable of operating on unimproved roads when carrying a full load of pods.

(U) With respect to the Bradley Fighting Vehicle, vulnerability/ballistic protection was a developmental test issue and thus not under the purview of and properly not reported by OTEA. The problem with the Bradley's night sight was identified by OTEA during OT III in 1983. The problem was caused by out-of-tolerance input voltages from the power control unit. Improved power control units were introduced on production models and successfully demonstrated during initial production tests and follow-on test and evaluation. Fielded systems demonstrate acceptable night sight availability and effectiveness.

(U) With respect to the assertion that the DDT&E did not identify the operational effect of adverse EF-111A test results, Air Force Operational Test and Evaluation Center (AFOTEC) files contain copies of DDT&E memoranda prepared for the EF-111A DSARC III production decision process. These memoranda identified test limitations and constraints, test results, and system deficiencies for the EF-111A.

RECOMMENDATIONS

(U) RECOMMENDATION 1: The GAO recommended that the Secretary of Defense enforce the requirement that operational testing be done in as operationally realistic an environment as possible. (p.4, p. 36/GAO Draft Report)

(U) DoD POSITION: Concur. As discussed in the DoD positions on Findings A and C, the DOT&E, working with the Services and the DUSDR&E(T&E), is pursuing a number of initiatives to improve OT, with particular emphasis on enhanced realism.

(U) RECOMMENDATION 2: The GAO recommended that the Secretary of Defense enforce the requirement that operational testing not be performed unless there are clearly stated test objectives and criteria. (p. 36/GAO Draft Report)

(U) DoD POSITION: Concur. The DOT&E, under the authority conferred upon him by 10 USC 136a, is currently doing this. For example, the recently revised (March 1986) DoDD 5000.3 and the new TEMP Guidelines Manual (DoDD 5000.3-M-1), which will be

published during the first quarter of FY 1987, set forth specific and detailed requirements for clearly and completely stated test objectives and criteria.

(U) RECOMMENDATION 3: The GAO recommended that the Secretary of Defense improve the usefulness of OT&E by the Services' OT&E agencies by requiring the agencies to: (1) state whether OT&E demonstrated that the system met operational requirements, (2) discuss the operational effect of significant test limitations and adverse test results on system performance, and (3) recommend whether a system is ready for production. (p. 36/GAO Draft Report)

(U) DoD POSITION: Partially concur. Both Service and DOT&E findings and recommendations on major acquisition programs are currently being provided to either or both the Joint Resources and Management Board (JRMB) and the cognizant Service Systems Acquisition Review Council, as appropriate. (The Office of the DOT&E was established in January 1984.)

(U) With regard to part (3) of the recommendation, that the Secretary of Defense require OT&E agencies to recommend whether systems are ready for production, while these agencies should not be precluded from making such recommendations should they deem it appropriate, requiring them to do so would be outside the scope of their charters and could compromise their independency by making them subject to internal pressures.

Requests for copies of GAO reports should be sent to:

U.S. General Accounting Office
Post Office Box 6015
Gaithersburg, Maryland 20877

Telephone 202-275-6241

The first five copies of each report are free. Additional copies are $2.00 each.

There is a 25% discount on orders for 100 or more copies mailed to a single address.

Orders must be prepaid by cash or by check or money order made out to the Superintendent of Documents.

United States
General Accounting Office
Washington, D.C. 20548

Official Business
Penalty for Private Use $300

Address Correction Requested

First-Class Mail
Postage & Fees Paid
GAO
Permit No. G100